remote sensing

MDPI

Article

A Novel Hybrid Deep-Learning Approach for Flood-Susceptibility Mapping

Abdelkader Riche [1,2,*], Ammar Drias [1], Mawloud Guermoui [3,4], Tarek Gherib [1], Tayeb Boulmaiz [5], Boularbah Souissi [6] and Farid Melgani [2]

[1] Faculty of Earth Sciences, Geography and Territorial Planning, University of Sciences and Technology Houari Boumediene, BP 32 Bab Ezzouar, Algiers 16111, Algeria; deriasammar@yahoo.fr (A.D.); tgherib@usthb.edu.dz (T.G.)
[2] Department of Information Engineering and Computer Science, University of Trento, Via Sommarive 9, I-38123 Trento, Italy; farid.melgani@unitn.it
[3] Unité de Recherche Appliquée en Energies Renouvelables, URAER, Centre de Développement des Energies Renouvelables, CDER, Zone Industrielle Bounoura, BP 88, Ghardaïa 47000, Algeria; m.guermoui@cder.dz
[4] Telecommunications and Smart Systems Laboratory, University of ZianeAchour, Djelfa 17000, Algeria
[5] Materials, Energy Systems Technology and Environment Laboratory, University of Ghardaia, Scientific Zone, P.O. Box 455, Ghardaia 47000, Algeria; boulmaiz.tayeb@univ-ghardaia.dz
[6] Faculty of Electrical Engineering, University of Sciences and Technology Houari Boumediene, BP 32 Bab Ezzouar, Algiers 16111, Algeria; boulerbah.souissi@usthb.edu.dz
* Correspondence: abdelkader.riche@usthb.edu.dz; Tel.: (+39)-3482911305

Abstract: Flood-susceptibility mapping (FSM) is crucial for effective flood prediction and disaster prevention. Traditional methods of modeling flood vulnerability, such as the Analytical Hierarchy Process (AHP), require weights defined by experts, while machine-learning and deep-learning approaches require extensive datasets. Remote sensing is also limited by the availability of images and weather conditions. We propose a new hybrid strategy integrating deep learning with the HEC–HMS and HEC–RAS physical models to overcome these challenges. In this study, we introduce a Weighted Residual U-Net (W-Res-U-Net) model based on the target of the HEC–HMS and RAS physical simulation without disregarding ground truth points by using two loss functions simultaneously. The W-Res-U-Net was trained on eight sub-basins and tested on five others, demonstrating superior performance with a sensitivity of 71.16%, specificity of 91.14%, and area under the curve (AUC) of 92.95% when validated against physical simulations, as well as a sensitivity of 88.89%, specificity of 93.07%, and AUC of 95.87% when validated against ground truth points. Incorporating a "Sigmoid Focal Loss" function and a dual-loss function improved the realism and performance of the model, achieving higher sensitivity, specificity, and AUC than HEC–RAS alone. This hybrid approach significantly enhances the FSM model, especially with limited real-world data.

Keywords: FSM; hybrid strategy; physical models; ground truth points; W-Res-U-Net

Citation: Riche, A.; Drias, A.; Guermoui, M.; Gherib, T.; Boulmaiz, T.; Souissi, B.; Melgani, F. A Novel Hybrid Deep-Learning Approach for Flood-Susceptibility Mapping. *Remote Sens.* **2024**, *16*, 3673. https://doi.org/10.3390/rs16193673

Academic Editor: Gabriel Senay

Received: 24 June 2024
Revised: 20 September 2024
Accepted: 27 September 2024
Published: 1 October 2024

1. Introduction

In an era marked by the escalating impacts of climate change, flooding has emerged as one of the most devastating and deadly natural hazards and a significant and recurring threat to communities worldwide. The devastating consequences of floods, ranging from loss of life and property damage to disruption of essential services and displacement of populations [1–3], underscore the urgent need for robust flood-detection methods. As climate change intensifies, the frequency and severity of extreme weather events, including storms and heavy rainfall, are expected to increase, amplifying the risk of flooding in vulnerable regions.

Addressing this pressing challenge involves a diverse range of methodologies for flood detection. These methods encompass traditional approaches rooted in physical processes, such as HEC–HMS/RAS (Hydrologic Engineering Center–Hydrologic Modeling

System/River Analysis System), SWAT (Soil and Water Assessment Tool), WetSpa (Water and Energy Transfer between Soil, Plants, and Atmosphere), and SOBEK Hydrodynamic Model, SWAT, WetSpa, and SOBEK, among others [4–6]. These models rely on intricate mathematical representations of physical phenomena. However, their widespread adoption faces challenges in acquiring diverse datasets due to various restrictions.

Recent advancements in remote-sensing technology have ushered in a new era of flood-mapping capabilities. By harnessing a diverse array of sensors mounted on satellites and aircraft, including synthetic aperture radar (SAR) and multispectral imaging, researchers can achieve both high-resolution detail and wide-area coverage, which are crucial for precise flood delineation. These sophisticated tools not only enable real-time monitoring but also facilitate timely disaster response and management efforts. Moreover, through seamless integration with geographic information systems (GIS), remote-sensing data empower comprehensive flood-risk assessment and the formulation of effective mitigation strategies. The effectiveness of this remote-sensing approach has been evidenced in numerous studies worldwide [7–10], affirming its pivotal role in flood mapping and management initiatives. Ongoing research endeavors focusing on algorithm refinement and sensor innovation promise to further elevate the accuracy and efficiency of these techniques. However, it is important to acknowledge that the method of FSM by remote sensing has limitations, such as the unavailability of SAR radar satellite images on days of flooding or the limitations of optical images due to the presence of clouds. These constraints highlight the need for continued advancements in remote-sensing technology to overcome such challenges and improve the reliability of flood-detection systems.

Remote sensing plays a crucial role in flood detection, yet it grapples with several challenges, notably limited spatial and temporal resolution, cloud-cover interference [11,12], and spectral sensitivity issues. These obstacles often lead to an underestimation of flood occurrences. Moreover, FSM faces its own set of hurdles, primarily stemming from data limitations. The scarcity of comprehensive flood inventory data derived from historical records and field surveys hampers the development of precise machine-learning models. This scarcity is compounded by challenges related to data accessibility, interpretation complexities, and the dynamic nature of flood events. Therefore, there is an urgent need to find an effective and accurate way of mapping floods, which can overcome these challenges and provide timely and reliable information for disaster management and mitigation efforts.

FSM and flood zone-determination systems offer crucial advantages in disaster mitigation. Accurate identification of flood zones enables timely warnings, which in turn reduces both casualties and property damage [13]. Early alerts empower communities to enact preventive measures, such as implementing evacuation plans and installing flood barriers, thus bolstering overall safety. Furthermore, precise flood zone determination facilitates optimal resource allocation, allowing emergency response teams to be deployed effectively. Ultimately, these systems play a pivotal role in saving lives, minimizing losses, and enhancing resilience against natural calamities.

2. State of the Art

Inundation of water in normally dry areas is the defining characteristic of floods. As a result, the study of flood management encompasses a wide range of considerations due to the inherent potential for impact that the presence of water entails. Several distinct approaches to flood management strategies have been identified and categorized [14]. River floods are the consequence of rivers exceeding their capacity, typically triggered by significant precipitation or snowmelt [13]. Flash floods, on the other hand, manifest swiftly, predominantly provoked by intense rainfall [13]. Coastal floods emerge from storm surges or tidal phenomena along coastal regions [13]. Similarly, dam breakages and dike breaches induce flooding events by surpassing infrastructure thresholds, causing water discharge beyond intended capacity. In our study, we will focus on investigating river floods and flash floods specifically. There are, therefore, several methods of flood-susceptibility mapping (FSM), as follows:

2.1. Physics-Based Models

Past research has utilized different physics-based flood models to simulate and evaluate flood susceptibility. Prominent software applications like HEC-HMS/RAS, in their ground-breaking research article, ref. [15] conducted a comprehensive study in New York City utilizing three key inputs, ultimately achieving a remarkable level of accuracy. Also in their seminal paper, ref. [16] undertook a rigorous user study in the Ohio River basin (USA), using SWAT and LISFLOOD-FP software and a diverse set of eight datasets, producing results deemed acceptable by the scientific community. Ref. [17] conducted a significant research study in Hyderabad city (India), employing the SWMM software with a focused strategy utilizing merely two inputs. So, these models have been employed for simulating flood progression in water channels and terrestrial flooding extents. Nevertheless, these conventional models have their limitations, including the need for intricate input data and extended computation durations due to their complex nature [3]. Additionally, they have restricted capabilities in simulating large-scale watershed flooding exceeding 1000 km^2 [3,18]. Consequently, it is crucial to devise more efficient and proficient approaches to tackle these deficiencies.

2.2. Multi-Criteria Decision Analysis (MCDA) and Statistical Models

In addition to the aforementioned considerations, the implementation of physics-based models often necessitates substantial computational resources and a thorough comprehension of the underlying principles. Spatial analyses, particularly employing multi-criteria decision analysis (MCDA) techniques, have been a focal point of research. Among these, the Analytic Hierarchy Process (AHP) emerges as a widely recognized and frequently utilized MCDA approach in the realm of FSM. Specifically, ref. [19] conducted their study utilizing eight inputs within the northern region of Tunisia, while ref. [20] employed seven inputs in Iran. Similarly, ref. [21] utilized eight inputs within Tanzania and has achieved acceptable accuracy, and ref. [22] employed eight inputs in Iran and has achieved good accuracy. These studies collectively underscore the diverse geographical contexts in which AHP-based FSM methodologies have been applied, highlighting their adaptability and effectiveness across different regions. This method facilitates systematic and structured evaluations of various criteria and their relative importance in flood susceptibility contexts. However, a significant limitation of MCDA is its vulnerability to bias due to reliance on expert knowledge. Another methodology in FSM involves utilizing bivariate and multivariate statistical analyses. Bivariate statistical analysis techniques play a crucial role in identifying the influential factors of each class within the modeling framework. Ref. [23] applied integrated bivariate-multivariate statistical models with 15 inputs in Korea, achieving acceptable accuracy. Ref. [24] utilized bivariate–multivariate statistical models with eight inputs in Jeddah (KSA), yielding good accuracy. Similarly, ref. [25] employed bivariate models with 10 inputs in Iran, achieving high accuracy. Ref. [26] utilized Bivariate Statistics-ML Models with 11 inputs in Romania, resulting in good accuracy. Furthermore, ref. [27] applied DEMATEL-ANP with 14 inputs in Slovakia, attaining high accuracy. It is worth noting that while statistical methods typically adopt a linear approach, which may have limitations, especially in addressing flood prediction with its nonlinear dynamics, these studies underscore the necessity for more advanced modeling techniques.

2.3. Remote Sensing

The existing literature on flood detection highlights the efficacy of Synthetic Aperture Radar (SAR) and optical imagery for flood-susceptibility mapping. Ref. [7] utilized TRS_LM_RGA on VV Polarization and TRS_KI_THR on VH Polarization with Sentinel-1 images in China, achieving very high accuracy. Ref. [28] applied the difference image index (DII) using Sentinel-1 images in the USA, resulting in acceptable accuracy. In Turkey, ref. [29] employed the MNDWI index with Sentinel-2 images. Moreover, research by [30] in Greece demonstrated the effectiveness of SAR and Optical Earth Observation with Hydraulic Simulation, utilizing Sentinel-1, Landsat 7, Landsat 8, and HEC–RAS modeling.

Similarly, ref. [31] conducted fusion techniques with Sentinel-1 and Sentinel-2 images in Turkey. These studies underscore the significance of SAR-optical fusion techniques in flood detection, aiding timely and precise assessments crucial for disaster management and mitigation efforts, especially in flood-prone regions. However, challenges persist, including the lack of imagery for relevant dates and interference from cloud cover, highlighting the need for advancements in remote-sensing methodologies to enhance the accuracy and efficacy of flood-detection systems. It is important to note that remote sensing is not a parallel comparison with flood-susceptibility mapping, while remote sensing provides real-time data and imagery for immediate flood detection and monitoring, and flood-susceptibility mapping involves predicting potential flood-prone areas based on various environmental and climatic factors, requiring a different set of analytical approaches and tools.

2.4. Machine and Deep Learning

In recent years, there has been a notable trend among researchers towards embracing advanced data-driven models, particularly machine-learning (ML) algorithms. This shift arises from the observed limitations in traditional physical flood models, analytical hierarchy processes, statistical models, and remote-sensing methods. Data-driven models depart from the reliance on physical processes, instead relying on mathematical equations derived from concurrent inputs and outputs. This approach enables the estimation of FSM directly from data without predefined assumptions. The application of ML techniques has shown promising results in various comparative studies. For instance, ref. [32] achieved high accuracy in China using Naive Bayes Tree (NBT) with 12 inputs. Ref. [33] obtained acceptable accuracy in Austria by employing Ensemble AHP-ANP_RF-SVM with 11 inputs. Then, ref. [34] achieved acceptable accuracy in Bangladesh through a hybrid model (Dagging models) with 12 inputs. Ref. [35] achieved acceptable accuracy in India using XGBoost with eight inputs. Ref. [36] attained good accuracy in India, utilizing AB-RF with 12 inputs. Ref. [3] achieved good accuracy in China with RF using 15 inputs. Ref. [37] obtained good accuracy in Iran employing the Cascade Forest Model (CFM) with 20 inputs. Refs. [38,39] achieved good and very high accuracy in Egypt, respectively, using RF with 15 inputs. Also, ref. [40] achieved high accuracy in India utilizing Stochastic Gradient Boosting (SGB) with 16 inputs.

In addition to traditional methodologies, a plethora of studies have delved into the efficacy of employing deep-learning techniques for enhancing flood-susceptibility mapping. Notable among these endeavors are the works of [41], who employed DLNN with nine inputs in Vietnam, achieving commendable accuracy. Similarly, ref. [42] utilized Social Spider Optimization (SSO) with 11 inputs, attaining high accuracy in Vietnam. Ref. [43] leveraged autoencoder-MLP with nine inputs, yielding high accuracy in Iran. Ref. [44] employed deep belief network, backpropagation, and genetic algorithm (DBPGA) with 11 inputs, achieving very high accuracy in Iran. Ref. [45] utilized Iterative Classifier Optimizer—Deep-Learning Neural Network—Frequency Ratio (ICO-DLNN-FR) with 14 inputs, demonstrating good accuracy in Romania. Ref. [46] applied CNN with 17 inputs, showcasing good accuracy in India. Ref. [47] compared CNN versus GMDH with 10 inputs, resulting in acceptable accuracy in Mozambique. Ref. [48] employed Filtered Classifier—Deep Learning (FC-DL) with 10 inputs, achieving high accuracy in China. Ref. [49] utilized DLNN-ICO with 18 inputs, demonstrating good accuracy in Bangladesh. Lastly, ref. [50] applied 2D-CNN with 10 inputs, yielding acceptable accuracy in Kenya. We provide detailed and numerical results for each study in Table 1.

Despite its efficiency and robustness, ML still grapples with limitations. One such constraint lies in the scarcity and inadequacy of flood inventory data derived from historical records and field surveys in existing studies. This scarcity poses a significant hurdle in developing accurate ML models [3], especially when deep-learning (DL) techniques are employed. Moreover, in the realm of FSM, prior research utilizing ML or DL techniques has predominantly focused on utilizing target points and geospatial factors as inputs, and

there are no studies considering geospatial factors as inputs and images as Target "Images to Image".

Table 1. State-of-the-art and comparison of previous studies in FSM.

Method	Reference	Model	Input	Location (Area)	Performance Metrics
Physics-Based Model	[15]	Physics-based Model HEC-RAS	* Topography (LIDAR) * LULC * Hydrological point data	New York, NY, USA	OA = 89.77% RMSE = 0.320
	[16]	Physics-based Model SWAT and LISFLOOD-FP	* DEM * LULC * Soil texture * Precipitation, temperature, solar radiation, relative humidity and wind speed	Ohio River Basin, OH, USA (500,000 km^2)	F = 0.75 C = 0.85
	[17]	Physics-based Model SWMM	* DEM * Rainfall Data	Hyderabad City, India (90.87 km^2)	
AHP and MCDA	[19]	GIS-AHP	*LULC * Elevation *lithology *Rainfall *Drainage density *slope *Soil *Ground water level	North-East of Tunisia (524.4 km^2)	
	[20]	MCDA	* Runoff * Elevation * Slope * Drainage Distance * Rainfall Intensity * Erosion * LULC	Mashhad Plain Basin, Iran (9762 km^2)	
	[21]	GIS-MCDA	* Slope *Geology *Flow accumulation *LULC *Drainage density *Elevation *stream order *Soil.	Dodoma Region, Tanzania (41.05 km^2)	AUC = 87.24%
	[22]	MCDA	* Rainfall *Distance from rivers *Slope *Soil type *Geology *Land use *Elevation *Drainage density.	Babolroud Watershed, Iran (2344 km^2)	AUC = 90.5%
Statistical Models	[23]	Integrated bivariate–multivariate statistical models	*Curvature *DEM *Geology *Greenfarm *Rainfall * River * Slope *Soil drain * Soil effect *Soil texture *SPI *Timber age *Timber density *Timber diameter *Timber type	Busan, Republic of Korea	AUC = 82.3
	[24]	Bivariate–multivariate statistical models	*Slope *Elevation *curvature *geological units * Land use *Soil drain *Distance from streams	Jeddah City, Saudi Arabia (219 km^2)	AUC = 90.4%
	[25]	Bivariate models	*Slope angle *plan curvature *altitude * TWI * SPI *Distance from river *Rainfall *Geology* Land use *NDVI	Haraz Watershed (4015 km^2)	AUC = 98.72%
	[26]	Bivariate Statistics-ML Models	*Plan curvature *Slope angle *Hydrological *Soil Group *Lithology *Land use *Convergence index *Elevation *Distance from river *Rainfall *TWI	Basin of Prahova River, Romania (2600 km^2)	AUC = 91.1%
	[27]	DEMATEL-ANP	*Elevation *Curve number * SPI*Slope *Stream density *TRI*Distance from river *STI *TWI *Rainfall *Lithology *Soil texture * NDVI * Land cover	Topľa River Basin, Slovakia (1548 km^2)	AUC = 97.5%

Table 1. *Cont.*

Method	Reference	Model	Input	Location (Area)	Performance Metrics
Remote Sensing	[7]	TRS_LM_RGA on VV Polarization TRS_KI_THR on VH Polarization	* Sentinel 1 SAR * Landsat 8 OLI/TIRS for Validation	Jialing River, China	VV OA = 98.82% VH OA = 98.59%
	[28]	Difference image index (DII)	* Sentinel-1 * Sentinel-2 and Landsat 8 imagery for validation	U.S. New Orleans (2500 km^2) and Hurricane Harvey (7500 km^2)	agreement percentages = 80% in the both area
	[29]	MNDWI	* Sentinel-2	Basin of Mersin, Turkey (24,940 hectares)	
	[30]	SAR and Optical Earth Observation with Hydraulic Simulation	* Sentinel-1 * Landsat 7 * Landsat 8 * HEC-RAS modelling for comparison	Sperchios River basin, Greece (1823 km^2)	
	[31]	Fusion Sentinel-1 and Sentinel-2	* Sentinel-1 * Sentinel-2	Ordu Province, Northern of Turkey	
Machine Learning	[32]	Naive Bayes Tree (NBT)	* Slope *Curvature *Altitude * Distance from the river *NDVI map *Lithology *Land use *Soil type *Rainfall *TWI *SPI *STI	Southeast Jiangxi Province, China (4053.16 km^2)	OA = 96.8% AUC = 0.97 Kappa = 0.95 RMSE = 0.115 MAE = 0.12
	[33]	Ensemble AHP-ANP_RF-SVM	* Elevation *Slope *Aspect *NDVI *TWI *SPI *Geology, *Land cover *Distance to drainage *distance to roads *Rainfall	Salzburg, Austria (7156.03 km^2)	AUC = 89.3%
	[34]	hybrid model (Dagging models)	*Elevation *Curvature *Aspect and slope *topographic roughness index *topographic wetness index *Stream power index *Stream transport index *land use, land cover *Distance to river *soil types *Rainfall	Teesta Sub-Catchment, Bangladesh	AUC = 0.873 RMSE = 0.189 MAE = 0.084 R^2 = 0.852
	[35]	XGBoost	* Rainfall * CN * Elevation * Slope * Distance from stream * Evapotranspiration * LST * NDVI	Telangana State, India	AUC = 0.83
	[36]	AB-RF	*Elevation *Aspect *Slope * TRI *TWI *SPI *Distance from rivers *Rainfall *NDVI *soil *geomorphology * Lithology	West Coast of India (3177 km^2)	AUC = 0.940
	[3]	RF	*Slope * Slope aspect * Profile curvature * Planar curvature * TP I* TRI * TWT * DTR * Flow accumulation * SPI * STI * Stream order * NDVI * NDISI * RDPI	Xinluo Sub-Watershed, China (600 km^2)	AUC = 91.2%

Table 1. *Cont.*

Method	Reference	Model	Input	Location (Area)	Performance Metrics
Machine Learning	[37]	Cascade Forest Model (CFM)	* Soil type * LCLU * NDVI * DEM * Slope * Aspect * Curvature * SCS * LC, ci * VD * LS * FA* TRI * TPI * MCA * SPI * TWI * HOFD * VOFD	Gorganrud (11,290 km^2) and Karun Basin (67,297 km^2), Iran	Karun basin: OA = 93.94% F1-Score = 90.70% Gorganrud: OA = 92.40% F1-Score = 91.60%
	[38]	RF	*Elvation *Soil *TPI *CI *LS *TWI * PrC *PlC *DD *GEO *NDVI *LULC *DTR *DTW *Rainfall	El-Matulla Catchment, Egypt (7231 km^2)	AUC = 0.93 OA = 0.88 Kappa = 0.85 RMSE = 0.34 MAE = 0.12
	[39]	RF	*Elevation *Slope *Aspect *Plan curvature *hillshade *Flow accumulation *TWI *STI *Horizontal flowdistance *Vertical flow distance *Rainfall *Flow distance *LU/land cover *lithology *NDVI.	Hurghada, Egypt (138 km^2)	AUC = 98.2%
	[40]	Stochastic Gradient Boosting (SGB) and GBC	*Elevation *Slope *Aspect *STI * TRI *TWI *SPI *Curvature * D2R*D2S *Geology *Geomorphology * LULC *Soil *NDVI *Rain	Idukki District, South India (4358 km^2)	AUC = 0.92
Deep Learning	[41]	DLNN	*Elevation *Slope *Aspect *curvature *stream density *NDVI *Soil type* Lithology *Rainfall	Lao Cai Province, Vietnam (1465.07 km^2)	SEN = 90.04% SPE = 94.26% AUC = 0.960
	[42]	Social Spider Optimization (SSO)	*DEM* Slope *Aspect * Curvature * TWI * SPI * DTR * River density * NDVI * NDBI * Rainfall	Lai Chau Province, Vietnam	AUC = 97.003% MAE = 0.033 OA = 95.926%
	[43]	Autoencoder-MLP	*Slope * Aspect * Altitude * Plan curvature * TWI * Lithology * Rainfall * Land use * Distance to drainage	Golestan Province, Iran (12,050 km^2) and Ganga Basins, India	AUC = 0.974
	[44]	Deep-belief network, back propagation, and genetic algorithm (DBPGA)	*Slope angle *elevation *Curvature *TWI *SPI * Distance to river * River density *rainfall *Lithology * Land use *NDVI	Haraz Watershed, Iran (4014 km^2)	AUC = 0.985
	[45]	Iterative Classifier Optimizer—Deep Learning Neural Network—Frequency Ratio (ICO-DLNN-FR)	* Lithology* Slope angle *Profile curvature *Hydrological Soil Groups * TWI *Land use *Convergence Index * Elevation* Distance from river *Plan curvature *Rainfall *Aspect * SPI *TPI	Putna River Basin, Romania (2509 km^2)	AUC = 0.911
	[46]	CNN	*Elevation* Aspect *Distance from river *LULC *TWI * Slope *TRI *TPI * Soil type *NDVI *Geomorphology *rainfall intensity *CN *SPI * Convergence index *CI *Geology	Kunur River, India (646 km^2)	AUC = 0.931 MAE = 0.143 RMSE = 0.378

Table 1. *Cont.*

Method	Reference	Model	Input	Location (Area)	Performance Metrics
Deep Learning	[47]	CNN versus GMDH	*Aspect * Slope *Altitude *Plan curvature *TWI *Profile curvature *valley depth LS-*factor or slope lengths *Land use *Distance of the river	Coastal City of Beira Central Mozambique	AUC CNN = 0.90 AUC GMDH = 0.87
	[48]	Filtered Classifier–Deep learning (FC-DL)	*Altitude *Slope angle *aspect *TWI *SPI *STI *Distance to rivers * Rainfall *NDVI *Land use	Jiangxi Province of China	AUC = 0.963
	[49]	DLNN-ICO	*Aspect *Elevation *Slope angle*Curvature *Plan curvature *Profile curvature *flow direction *Flow accumulation *LULC *NDVI *Distance to the rivers * Soil *Mean annual rainfall * river density *SPI*TWI *STI *Geology	Padma River, Bangladesh (2562 km^2)	AUC = 0.917
	[50]	2D-CNN	*Rainfall * LULC *Altitude *Slope *Flow direction *Flow accumulation *SPI *TWI *DTR*Soils	Eldoret Municipality, Kenya (247.7 km^2)	OA = 82.5% AUC = 0.809

3. Contribution of Paper

It is worth noting that while many machine-learning (ML) and deep-learning (DL) methods have primarily focused on classification and regression tasks, our research represents a recent advancement in exploring the potential of segmentation models for flood-susceptibility mapping (FSM). This approach has rarely been used, with a few exceptions in recent articles focusing on synthetic aperture radar (SAR) image segmentation using radar-based water indices as masks, rather than on geospatial factors. The main contributions of this paper are as follows:

3.1. Utilization of 2D Models

Traditionally, flood detection and classification have been approached using 1D models, focusing on point-based analysis. In contrast, our methodology leverages 2D modeling techniques, treating the problem as image classification. This paradigm shift allows for a more comprehensive analysis of flood patterns and characteristics.

3.2. Input Features

We introduce a comprehensive set of 19 input features for flood detection, including a novel feature termed "flow discharge", which captures critical information previously overlooked in the literature. By incorporating a diverse range of inputs, our methodology enhances the discriminative power of flood detection algorithms, leading to improved accuracy and reliability.

3.3. Data-Scarcity Challenges

One significant challenge in flood-detection research is the limited availability of real-ground truth data, often necessitating the use of physical models for training. However, relying solely on physical models presents its own challenges, including potential biases and inaccuracies in the generated data. To address this issue, we propose a deep classification framework employing two focal-loss functions designed to prioritize learning from ground

truth data obtained from physical models. By doing so, we aim to mitigate the effects of data scarcity and enhance the generalization capabilities of our models.

3.4. Introduction of Weighted Residual U-Net Architecture (W-Res-U-Net)

We introduce a novel architecture termed "Weighted Residual U-Net" (W-Res-U-Net), which combines the expressive power of residual connections with the versatility of the U-Net architecture. Additionally, we integrate two focal loss functions into the W-Res-U-Net framework, aimed at enhancing attention to ground truth data while placing less emphasis on data from physical models. This approach facilitates more effective gradient propagation and feature learning, ensuring that the model prioritizes accurate real-world data. To the best of our knowledge, this is the first instance of such an architecture being applied to flood detection and classification tasks.

3.5. Validation Strategy

To evaluate the effectiveness of our proposed methodology, we employ a rigorous validation strategy. We partition the study basin into sub-basins and train our models on subsets while validating them on independent sub-basin data from different years. This cross-validation approach ensures the robustness and generalizability of our models across various temporal and spatial contexts.

Overall, our contributions advance state-of-the-art flood detection and classification by leveraging innovative techniques and addressing key challenges inherent to the field.

4. Materials and Methods

Our methodology begins with an in-depth exploration of the studied area to establish a solid foundation. We then examine the theoretical underpinnings of the physical model HEC-HMS/RAS, ensuring a clear understanding of its principles. Concurrently, we engage deeply with the theoretical frameworks underpinning the employed deep-learning models, leveraging their power for enhanced analysis. Finally, we synthesize these methodologies into a cohesive and innovative hybrid approach, poised to deliver nuanced insights and solutions within the realm of our study.

4.1. Study Area

This study focuses on the Wadi El Harrach sub-watershed, a major river system running through the Algiers region of Algeria. This area was chosen for analysis due to several key factors: the availability of pertinent data, historical instances of flooding, and notable changes in land use and land cover (LULC) [51]. Originating from the Atlas Blidéen and the Sahel, this river traverses the Mitidja plain before entering the region of Algiers, covering an expansive area exceeding 800 km^2. Positioned in North–Central Algeria, its coordinates span latitudes from 2°56′34″ to 3°19′39″E and longitudes from 36°25′11″ to 36°46′40″N. The topography of the region exhibits a diverse range of elevations, from −1 to 1256 m above sea level, as depicted in Figure 1, using the one-arcsecond Shuttle Radar Topography Mission Digital Elevation Model (SRTM DEM). The climatic conditions feature rainfall return periods between 60 mm and 166 mm, with flow discharge in the rivers ranging from 24 m^3/s to 1690 m^3/s. The watershed incorporates nine rainfall-monitoring stations and one flow-monitoring station, covering four of the most densely populated provinces in Algeria: Algiers, Blida, Boumerdes, and Médéa.

The Wadi El Harrach has experienced numerous hydrological events of extreme intensity over the years, resulting in catastrophic floods that have caused significant loss of life and property. Additionally, the authorities' desire to develop and modernize the wadi has added to the urgency of studying this area. It is due to these factors, alongside the availability of data, that this study area was selected for investigation.

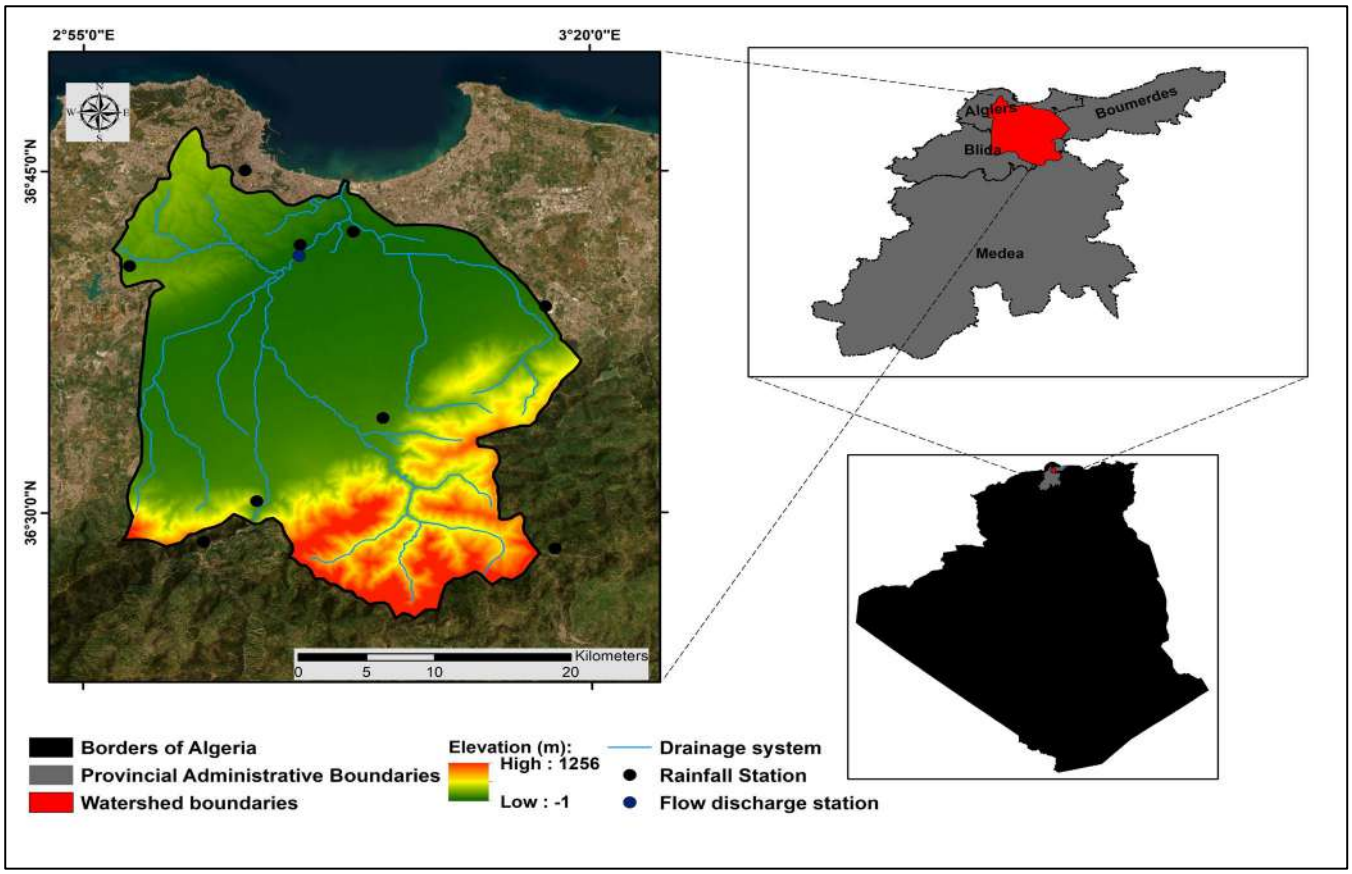

Figure 1. Localization of study area [51].

4.2. Theory of the Physical Model HEC-HMS/RAS

The physical model HEC-HMS/RAS integrates the Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS) and River Analysis System (HEC-RAS) to provide a comprehensive approach for simulating hydrological processes and hydraulic responses in river systems. This model employs physical principles to simulate rainfall–runoff processes using HEC-HMS, which calculates hydrographs for input into HEC-RAS, a hydraulic modeling tool that simulates flow profiles and inundation extents within river channels.

HEC-HMS is a comprehensive hydrological modeling system used to simulate rainfall–runoff processes. It incorporates various factors such as precipitation, land-surface characteristics, soil properties, and land use [52] to predict how rainfall is transformed into runoff over time [53]. The model typically follows these steps:

(a) Precipitation Input: Rainfall data are provided as the input to the model, either in the form of historical records or synthetic rainfall generated based on statistical methods.
(b) Rainfall–Runoff Transformation: HEC-HMS uses hydrological methods, such as the Soil Conservation Service (SCS) Curve Number method or the Soil Moisture Accounting (SMA) method, to simulate how rainfall infiltrates into the soil, is stored, and eventually becomes runoff.
(c) Hydrograph Generation: Based on the rainfall–runoff transformation, HEC-HMS generates hydrographs representing the flow of water over time at various points in the watershed.

HEC-HMS offers a thorough grasp of watershed responses to rainfall, enabling diverse hydrological scenario evaluations. However, its reliance on simplifications and uncertainties in input data and parameters may impact accuracy.

HEC-RAS is a hydraulic modeling tool designed for simulating flow profiles and inundation extents within river channels. It utilizes principles of fluid mechanics to predict water surface profiles, flow velocities, and floodplain inundation [54]. The typical workflow of HEC-RAS involves the following steps:

(a) Geometry Input: River channel geometry, including cross-sectional profiles and hydraulic structures, is input into the model.
(b) Flow-Boundary Conditions: Flow data, such as upstream discharges or boundary conditions, are specified.
(c) Hydraulic Analysis: HEC-RAS solves the equations of continuity and momentum to compute water-surface elevations, flow velocities, and depths along the river channel. The continuity equation ensures mass conservation along the flow path:

$$Q = AV \tag{1}$$

where: Q = Flow rate (m^3/s), A = Cross-sectional area of flow (m^2), V = Flow velocity (m/s).

The momentum equation is derived from Newton's second law and is used to compute flow velocities and depths:

$$\frac{\partial(AV)}{\partial x} + gA\frac{\partial Z}{\partial x} + gAi = 0 \tag{2}$$

where: x = Spatial coordinate (m), g = Acceleration due to gravity (m/s^2), Z = Elevation head (m), i = Energy slope (m/m).

Additionally, Manning's equation may be used to calculate flow velocities based on channel roughness and slope:

$$V = \frac{1}{n} \times R^{2/3} \times S^{1/2} \tag{3}$$

where: n = Manning's roughness coefficient ($s/m^{1/3}$), R = Hydraulic radius (m), S = Channel slope (m/m).

(d) Floodplain Mapping: Based on the hydraulic analysis, HEC-RAS can generate floodplain maps showing areas prone to inundation under different flow conditions.

HEC-RAS provides complex hydraulic analysis for river systems, encompassing floodplain mapping and flood-risk assessment. Nevertheless, it requires precise input data such as the definition of cross-sections and implies substantial computational requirements, particularly for large or complex river systems.

4.3. Machine/Deep Learning Models Employed in This Study

In this research, we employed a diverse set of machine- and deep-learning models to explore their efficacy in flood detection and generalization. Specifically, our methodology incorporates the random forest (RF) algorithm alongside three deep-learning architectures: Convolutional Neural Network (CNN), U-Net, and Residual U-Net. By leveraging this array of models, we aim to comprehensively evaluate their individual performance and comparative strengths within the context of FSM.

4.3.1. Random Forest (RF)

Random forest (RF) stands out as an ensemble learning technique utilized in both classification and regression tasks. Through the construction of numerous decision trees during training, it derives the mode of classes for classification or the average prediction for regression. This method significantly bolsters accuracy and robustness by amalgamating the predictions of individual trees, thereby mitigating overfitting concerns. RF algorithms have gained widespread adoption across diverse domains, owing to their efficacy in tackling large datasets characterized by high dimensionality and mixed variable types. Originally conceptualized by [55], RFs, as evidenced by a multitude of

flood-susceptibility mapping studies like those conducted by [3,38,39], have emerged as a favored methodology for predictive modeling and flood-analysis tasks. Equation (4) presents the RF model:

$$\hat{y} = \frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{n} W_j(x_i, x')y_i = \sum_{i=1}^{n}\left(\frac{1}{m}\sum_{j=1}^{m} W_j(x_i, x')\right) y_i \quad (4)$$

where $\hat{y}$ is the predicted output, m is the number of trees in the forest, n is the number of training samples, $W_j(x_i, x')$ is the weight assigned to the i^{th} training sample by the j^{th} tree, and y_i is the output of the i^{th} training sample.

The weight assigned is provided by:

$$W_j(x_i, x') = \begin{cases} 1, \text{ if } x_i \text{ is in the same leaf as } x' \text{ in the } j^{th} \text{ tree} \\ 0, \text{ other wise} \end{cases} \quad (5)$$

where x' is the input sample to be predicted.

4.3.2. Convolutional Neural Network (CNN) Model

CNNs are a class of deep neural networks particularly well-suited for tasks involving images and spatial data. They have revolutionized various fields including computer vision, pattern recognition, and image classification. CNNs are characterized by their ability to automatically learn spatial hierarchies of features from the input data through the application of convolutional filters (as shown in Figure 2). The CNN model is thus used in the field of FSM like the study by [46,47,50]. Its mathematical formulation can be explained as follows:

Let I be the input image with dimensions $H \times W \times C$, where H is the height of the image, W is the width of the image, and C is the number of channels.

Due to a convolutional filter K with dimensions $K_h \times K_w \times C$ and a stride S, the convolution operation at position (i, j) in the output feature map is computed as follows:

$$O_{i,j} = \sum_{m=0}^{K_{h-1}}\sum_{n=0}^{K_{w-1}}\sum_{k=0}^{C-1} I_{(i\times S+m),(j\times S+n),k} \times K_{m,n,k} + b \quad (6)$$

where:

$O_{i,j}$ is the value at position (i, j) in the output feature map,

$I_{(i\times S+m),(i\times S+n),k}$ is the pixel value at position $(i \times S + m), (i \times S + n)$ in the input image I at channel k,

$K_{m,n,k}$ is the value of the convolutional kernel at position (m, n) or channel k,

b is the bias term, and K_h and K_w are the height and width of the kernel, respectively.

Here, (m, n) represent the indices of the spatial dimensions of the convolutional kernel, indicating the offset within the kernel at each position.

After the convolution operation, an activation function f is usually applied element-wise to introduce non-linearity. The most common activation function is the Rectified Linear Unit (ReLU), which is defined as:

$$f(x) = max(0, x) \quad (7)$$

where: x is the input to the activation function. This function sets all negative values to zero, allowing the model to learn complex patterns and relationships in the data.

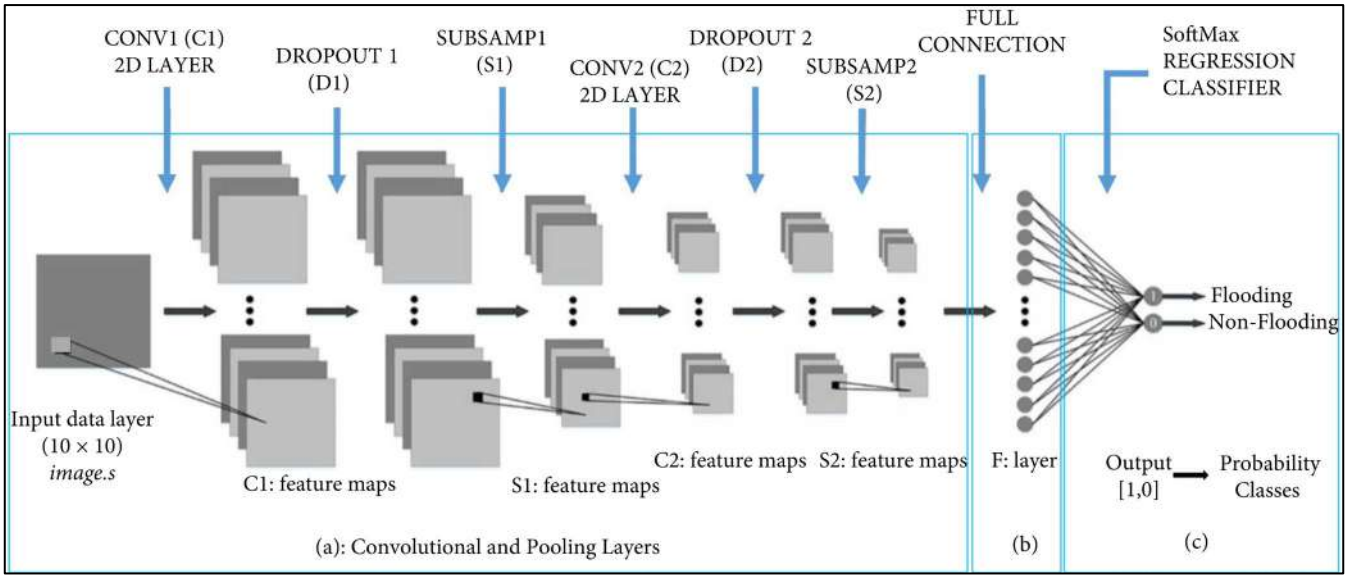

Figure 2. Topological architecture of the 2D feedforward CNN. (**a**) represents the feature extraction phase; (**b**) represents the extracted features; (**c**) is the food and non-food classification processes of an output [50].

After the activation function, pooling layers are often used to reduce the spatial dimensions of the feature maps. Max pooling is the most common type of pooling; the maximum value within each pooling window is selected. If the pooling window has dimensions $P_h \times P_w$, and a stride P_s, then the output of the max pooling operation at position (i, j) in the feature map is computed as:

$$O_{i,j} = \max_{m=0}^{P_{h-1}} \max_{n=0}^{P_{w-1}} I_{(i\times P_s+m),(j\times P_s+n)} \tag{8}$$

where $O_{i,j}$ is the value at position (i, j) in the pooled feature map, and $I_{(i\times P_s+m),(j\times P_s+n)}$ is the input value at position $(i \times P_s + m), (j \times P_s + n)$ in the feature map.

These formulas describe the basic operations of CNN for multichannel images. The network typically consists of several such layers, followed by fully connected layers and an output layer for classification or regression tasks.

4.3.3. U-Net Model

U-Net is a convolutional neural network architecture originally designed for biomedical image segmentation tasks. It was introduced by [56]. U-Net architecture (as shown in Figure 3) consists of a contracting path followed by an expansive path. The contracting path is responsible for capturing the context, while the expansive path enables precise localization. And since then, it has found applications in various fields. For instance, ref. [57] utilized it for the detection of mathematical expressions in scientific documents. Additionally, ref. [58] employed it for nuclei segmentation in histology images. Furthermore, it has been applied in the field of flood detection, as demonstrated by [59] in their study for near-real-time flood detection using SAR images.

The contracting path follows the typical architecture of a convolutional neural network with repeated applications of convolutional layers followed by down-sampling operations such as max pooling. Each convolutional layer is typically followed by batch normalization and a rectified linear unit (ReLU) activation function. The expansive path consists of up-sampling operations followed by convolutional layers. The up-sampling is usually performed using transposed convolutions or interpolation techniques like bilinear up-sampling. Skip connections from the contracting path are concatenated with feature maps at each step in the expansive path to enable precise localization.

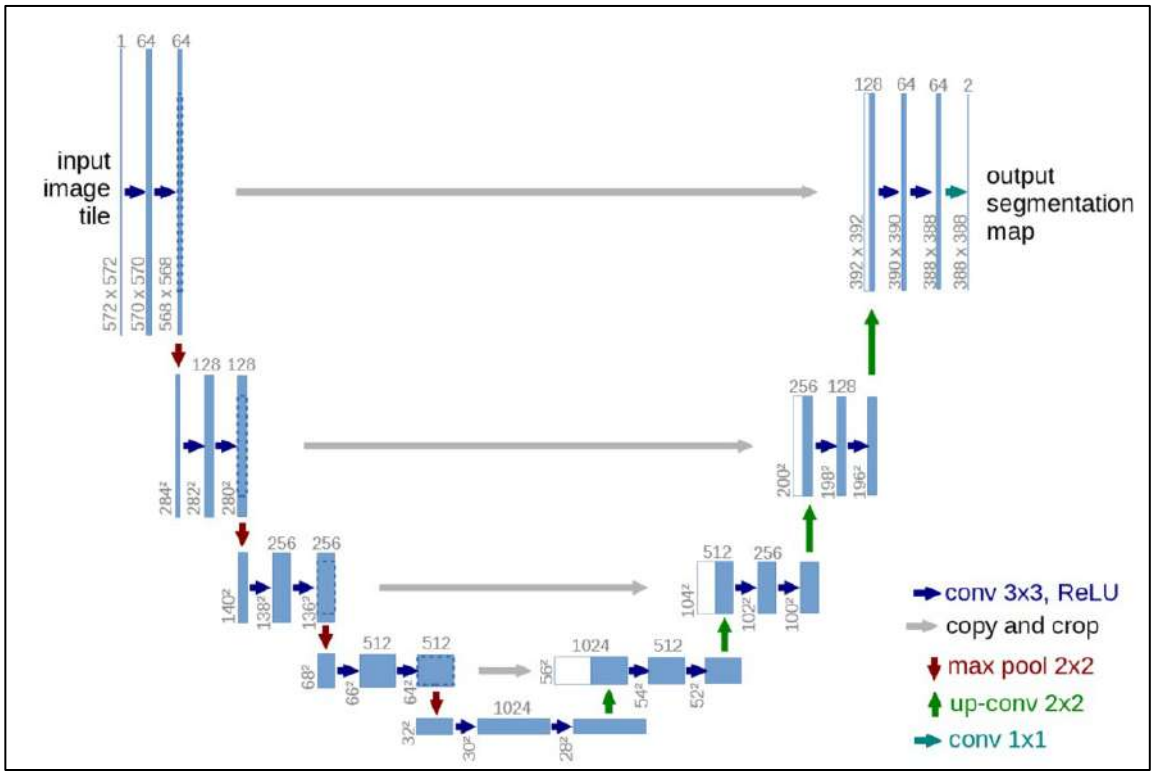

Figure 3. U-net architecture (example for 32 × 32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x–y size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations [56].

Let us denote the input image as X and the output segmented image as Y. The U-Net architecture can be described by the following mathematical formulas:

(a) Contracting Path (Down-Sampling):

$$C_i = Convolution(C_{i-1}) \tag{9}$$

$$P_i = MaxPooling(C_i) \tag{10}$$

where C_i represents the feature map at the i^{th} layer of the contracting path and P_i represents the pooled feature map.

(b) Expansive Path (Up-Sampling):

$$U_{i=}UpSampling(U_{i-1}) \tag{11}$$

$$M_i = Concatenate(C_{n-1}, U_i) \tag{12}$$

$$E_i = Convolution(M_i) \tag{13}$$

where U_i represents the up-sampled feature map at the i^{th} layer of the expansive path, M_i represents the concatenated feature map, and E_i represents the output feature map.

$$\mathrm{Y} = \mathrm{Sigmoid}(\mathrm{E}_0) \tag{14}$$

where E_0 represents the final output feature map, and the Sigmoid activation function is applied to map the values to the range [0, 1] for segmentation.

4.3.4. Residual U-Net Model

In a Residual U-Net, residual blocks are added to the contracting and/or expansive paths of the U-Net architecture. A residual block consists of skip connections (as shown in Figure 4) that bypass one or more convolutional layers and directly connect input and output tensors. This enables the network to learn residual mappings, making it easier to train deeper networks without encountering degradation in performance [60]. Notably, the study conducted by [61] implemented RMS-UNet, a variant of UNet featuring multiscale residuals for liver and lesion segmentation. Here, the input and output tensors are denoted as X and Y, respectively.

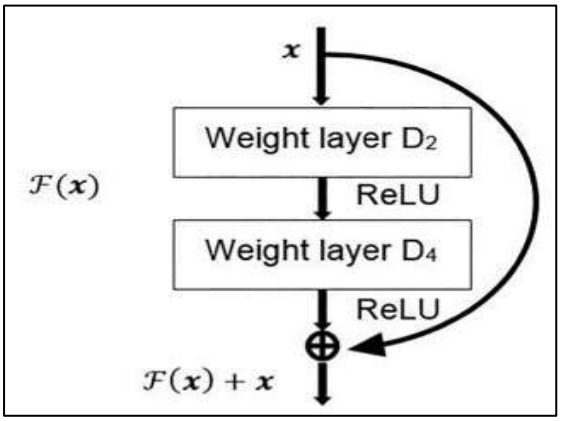

Figure 4. Residual learning: a building block [59].

Mathematically, a residual block can be represented as follows:

$$Y = F(X, W_i) + X \tag{15}$$

where $F(X, W_i)$ is the residual mapping that is to be learned, and X and Y are the input and output of the considered layer. In our case, F is equivalent to:

$$F(X, W_i) = rW_2r(W_1X) \tag{16}$$

where r is the rectified linear activation function and W_i represents the dilated convolution layer with weights.

4.4. Dataset Used

In this section, to better understand the methodology used for FSM, we first present the data used, and then the detailed approach that includes all the scenarios.

4.4.1. Flood Factors (Geospatial Database)

Flood occurrences are influenced by a myriad of factors spanning physical, environmental, meteorological, and hydrological domains. In our investigation, we comprehensively assessed 19 such factors as Rasters of 30 m, drawing from previous research confirmations [62–66] and data availability within our designated study area. These factors were systematically categorized into three overarching groups:

(a) Topographical factors:

This category encompassed crucial variables such as the Digital Elevation Model (DEM), slope, flow accumulation, flow direction, Topographic Wetness Index (TWI), topographic roughness Index (TRI), and Topographic Position Index (TPI). We utilized the Shuttle Radar Topography Mission Digital Elevation Model (SRTM DEM) with a 30 m resolution to derive the topographical factors.

- Digital Elevation Model (DEM): Provides the elevation data, which are crucial for understanding the terrain. The SRTM DEM data acquired through remote-sensing

techniques were fundamental in this regard. Elevation is a fundamental factor in understanding how water moves across landscapes. Low-lying areas or depressions are more prone to water accumulation and flooding, while higher elevations are less likely to flood but may influence downstream areas.

- Slope: Calculated from the SRTM DEM data to determine the steepness of the terrain, which influences runoff. The slope directly influences the speed and volume of surface runoff. Steeper slopes tend to result in faster runoff, reducing the chance of water infiltrating the soil and leading to greater flood potential in downstream areas.
- Flow Accumulation: Uses SRTM DEM to determine the accumulated flow to each cell, aiding in identifying potential flood paths. Flow accumulation helps identify areas where water is likely to gather, such as valleys or river channels, making these areas more prone to flooding.
- Flow Direction: Uses SRTM DEM to model the direction of water flow across the terrain. Understanding the flow direction is essential for predicting water movement during rainfall events. It allows for the modeling of drainage patterns, helping to forecast where water will travel and which areas it might affect downstream.
- Topographic Wetness Index (TWI): originally introduced by [67], calculated using the formula:

$$TWI = \ln\left(\frac{a}{tan(\beta)}\right) \tag{17}$$

where α is the local upslope area draining through a point per unit contour length and β is the local slope. TWI helps identify areas prone to saturation and potential flooding and has been extensively utilized by many researchers in hydrological modeling, including FSM [3,39,65].

- Topographic Roughness Index (TRI) and Topographic Position Index (TPI): are two key morphological factors closely linked to flooding, as noted by [68]. TPI represents the average elevation difference between a target cell and its surrounding cells, while TRI is the average squared difference in elevation, providing a measure of terrain ruggedness. Both indices, derived from SRTM-DEM, are crucial for understanding landform position and terrain roughness, which significantly influence water-flow and accumulation patterns.

(b) Environmental factors:

Including land-use and land-cover (LULC) classification, lithology, normalized difference vegetation index (NDVI), normalized difference built-up index (NDBI), and modified normalized difference water index (MNDWI). We utilized Landsat images for LULC classification, NDVI, NDBI, and MNDWI. Specifically, we employed the following Landsat data: Landsat –5 TM USGS: 25 February 2000, 21 January 2005, 16 March, 2010, Landsat-8 OLI USGS: 10 February 2015, 24 February 2020. These dates were selected to represent the winter rainy season in the study area, with a cloud cover limitation of 10% or less.

- Land Use and Land Cover (LULC): Classification through the SVM method using Landsat imagery to determine different land types and their impact on flood behavior. Different land uses and covers influence the hydrological response of an area. For example, forests and wetlands promote water infiltration and storage, reducing runoff, while urban and built-up areas increase impervious surfaces [51], leading to greater runoff and flood potential.
- Lithology: Geological composition of the area, affecting water infiltration and runoff. Lithology impacts how much water can infiltrate the ground versus how much will run off. Permeable rocks (like sandstones) allow water to infiltrate, reducing surface runoff and flood risk, whereas impermeable rocks (such as clays or granites) limit infiltration [69], increasing runoff and flood potential. Therefore, understanding lithology is essential for determining areas more prone to flooding based on the underlying geology.

- Normalized Difference Vegetation Index (NDVI) was first introduced by [70] in 1974 as a measure of vegetation density using red and near-infrared bands. It became widely adopted for monitoring vegetation health and assessing land cover.

$$NDVI = (NIR - RED/NIR + RED) \quad (18)$$

where NIR is the near-infrared reflectance and RED is the red reflectance. NDVI, derived from Landsat images, is used to assess vegetation density, influencing evapotranspiration and runoff. Higher NDVI values indicate denser vegetation, which can reduce flood risk by increasing evapotranspiration, enhancing water absorption, and slowing down runoff. Low NDVI values (indicative of bare soil or sparse vegetation) can lead to higher runoff rates, increasing flood risk.

- Normalized Difference Built-up Index (NDBI): was proposed by [71] as a method for identifying built-up areas in Landsat imagery.

$$NDBI = (SWIR - NIR/SWIR + NIR) \quad (19)$$

where SWIR is the shortwave infrared reflectance. NDBI, calculated from Landsat data, helps in identifying built-up areas, affecting impervious surfaces and runoff. Built-up areas often have a high percentage of impervious surfaces (such as roads, buildings, and concrete), which prevent water from infiltrating into the soil, increasing surface runoff and flood risk.

- Modified Normalized Difference Water Index (MNDWI) was developed by [72] in 2006 to improve water feature extraction in satellite imagery, by modifying the Normalized Difference Water Index (NDWI) using the green and shortwave infrared bands.

$$MNDWI = (Green - SWIR/Green + SWIR) \quad (20)$$

where Green is the green reflectance. MNDWI, using Landsat imagery, is used to enhance open water features and distinguish them from built-up areas. MNDWI is useful in distinguishing between permanent water bodies and areas that are temporarily flooded, allowing for real-time assessment of flood conditions.

(c) Meteorological and hydrological factors:

Encompassing precipitation, flow discharge (a novel factor introduced in our study), curve number (CN), hydrologic soil group (HSG), stream power index (SPI), stream transportation index (STI), and distance to rivers (DTR).

- Precipitation: The rainfall data are from our nine rainfall stations interpolated using the IDW method to obtain raster maps.
- Flow Discharge: Measurement of water flow in rivers using the HEC–HMS model, crucial for understanding flood potential.
- Curve Number (CN): Derived from LULC and soil data to estimate direct runoff potential.
- Hydrologic Soil Group (HSG): Classification based on soil infiltration rates, affecting runoff.
- Stream Power Index (SPI):

$$SPI = A \times tan(\beta) \quad (21)$$

where: A is the upslope contributing area (flow accumulation) per unit contour length. $tan(\beta)$ is the local slope of the terrain, which quantifies the erosive power of flowing water based on the flow accumulation and slope of the terrain [73].

- Stream Transportation Index (STI): estimates the sediment transport capacity of flowing water and reflects sediment mobility, with higher STI values associated with an increased frequency of floods [74]. The STI is calculated using the formula:

$$STI = \left(\frac{A}{a_0}\right)^m \times \left(\frac{sin(\beta)}{b_0}\right)^n \quad (22)$$

where A is the upslope contributing area (flow accumulation). a_0 is a normalization constant for the area. β is the local slope of the terrain. b_0 is a normalization constant for slope (often taken as 1). m and n are empirical constants, commonly taken as 1, but can vary depending on the study.

- Distance to Rivers (DTR): Calculated using GIS techniques to understand the influence of proximity to rivers on flooding risk.

Importantly, our study accounted for the temporal dynamics inherent in certain factors (as shown in Figure 5), notably LULC, CN (aligned with the study of [51] in the study area), NDVI, NDBI, and MNDWI, leveraging Landsat imagery to capture temporal variations (2000, 2005, 2010, 2015, and 2020). Additionally, we incorporated rainfall and simulated flow discharge (5, 10, 20, 50, and 100 years), employing the HEC–HMS model. The discharges were strategically selected to build a robust dataset, essential for training our deep-learning model, which requires extensive data to accurately capture spatial and temporal flood dynamics and recognize their temporal variability and significance in flood dynamics. In our study, we relied on the rainfall and flow discharge return periods of 5, 10, 20, 50, and 100 years with the years 2000, 2005, 2010, 2015, and 2020, respectively, because the use of the data from these years helped generate the mentioned return periods; for example, LULC and CN for the year 2000 were used to generate the flow discharge return period of 5 years.

4.4.2. Flood-Inventory Data

The initial phase of FSM entails the acquisition of flood-inventory data, which often poses challenges due to limited availability and the absence of satellite images on specific dates. Insufficient data can hinder the effective training of machine-learning models, particularly deep-learning architectures. In addressing this challenge, ref. [75] employed a long-short-term local spatial sequential memory (LSS-LSTM) neural network to enhance their dataset. Originally comprising 152 data points, with an equal distribution of 76 flooded and non-flooded instances for training and 64 (32 flooded and 32 non-flooded) for testing, their approach aimed to mitigate data sparsity. They tackled this issue through augmentation techniques, including rotating each extracted image patch by 90°, 180°, and 270°, and generating four additional samples by horizontally flipping the four orientations of the image patches. This augmentation strategy significantly bolstered the training set, expanding it eight-fold to a final count of 1216 samples. Nonetheless, while data augmentation is a pragmatic method to enhance model performance, its applicability in hydrological contexts warrants scrutiny. Repetitive training with mere orientation alterations may not sufficiently enhance the model generalization capabilities.

In our study, we have a dataset comprising 63 data points obtained from civil-protection authorities. Of these, 37 were associated with river floods, while 26 were linked to flash floods. Our research deviates from traditional approaches by integrating these ground truth points into HEC-RAS modeling, akin to the way an image incorporates pixels representing flooded and non-flooded areas (coded as "1" and "0", respectively). This integration was crucial as the limited number of 63 truth points rendered training a machine-/deep-learning model unfeasible. Moreover, physical modeling, despite its advantages, presents challenges such as the requirement for specialized expertise and the delineation of hydrological boundaries, along with defining cross-sections in HEC-RAS.

To address these limitations, our investigation emphasizes the temporal aspect of the factors mentioned above. We augment our analysis by simulating various return periods (5, 10, 20, 50, and 100 years) within the HEC-RAS framework as shown in Figure 6. By doing so, we not only base our study on physical modeling but also ensure that ground truth points are not overlooked, signifying a relevant advancement. This innovative "images-to-image" methodology for FSM leverages ground truth points in conjunction with physical modeling, representing a novel approach in the field.

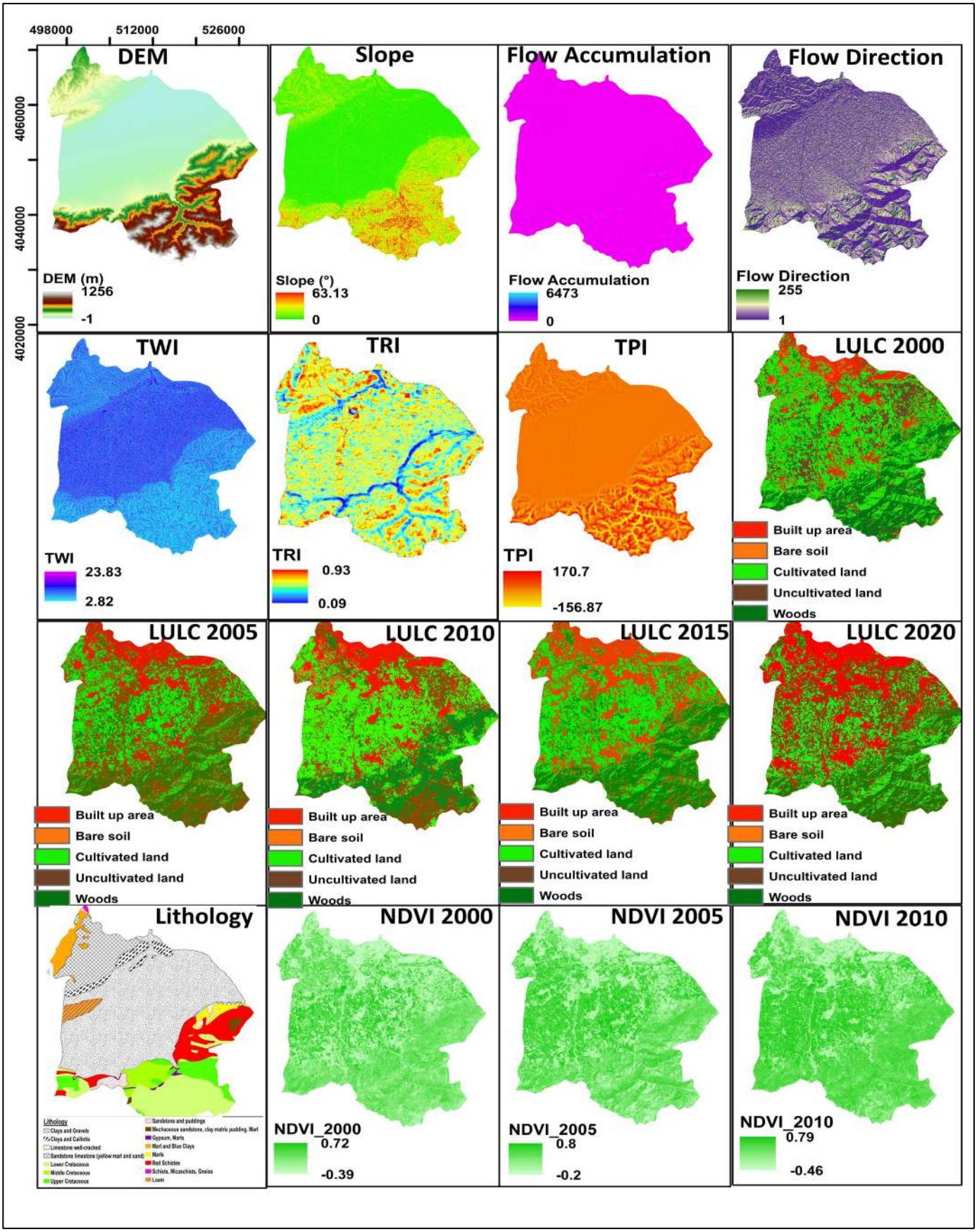

Figure 5. *Cont.*

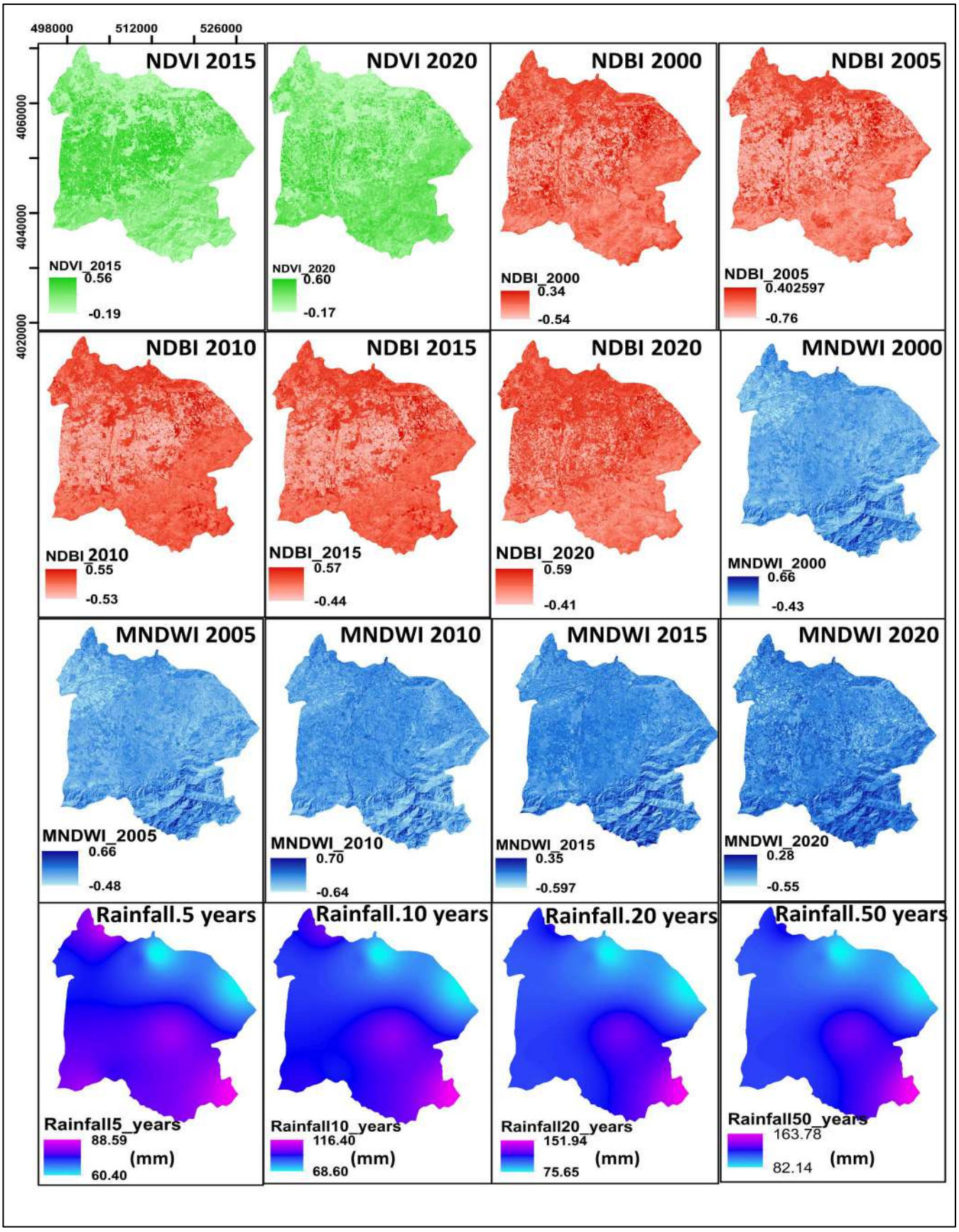

Figure 5. *Cont.*

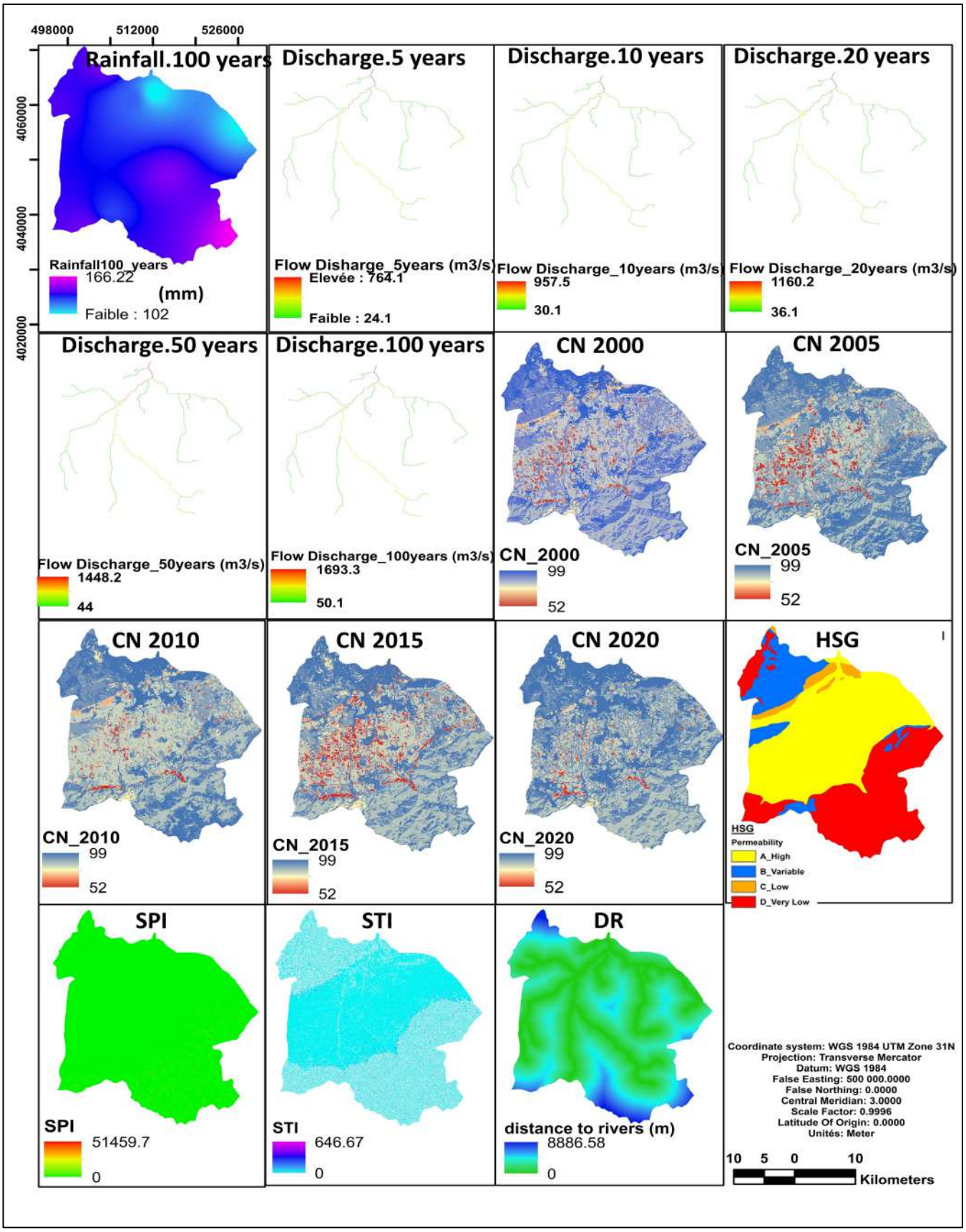

Figure 5. Geospatial database for flood factors.

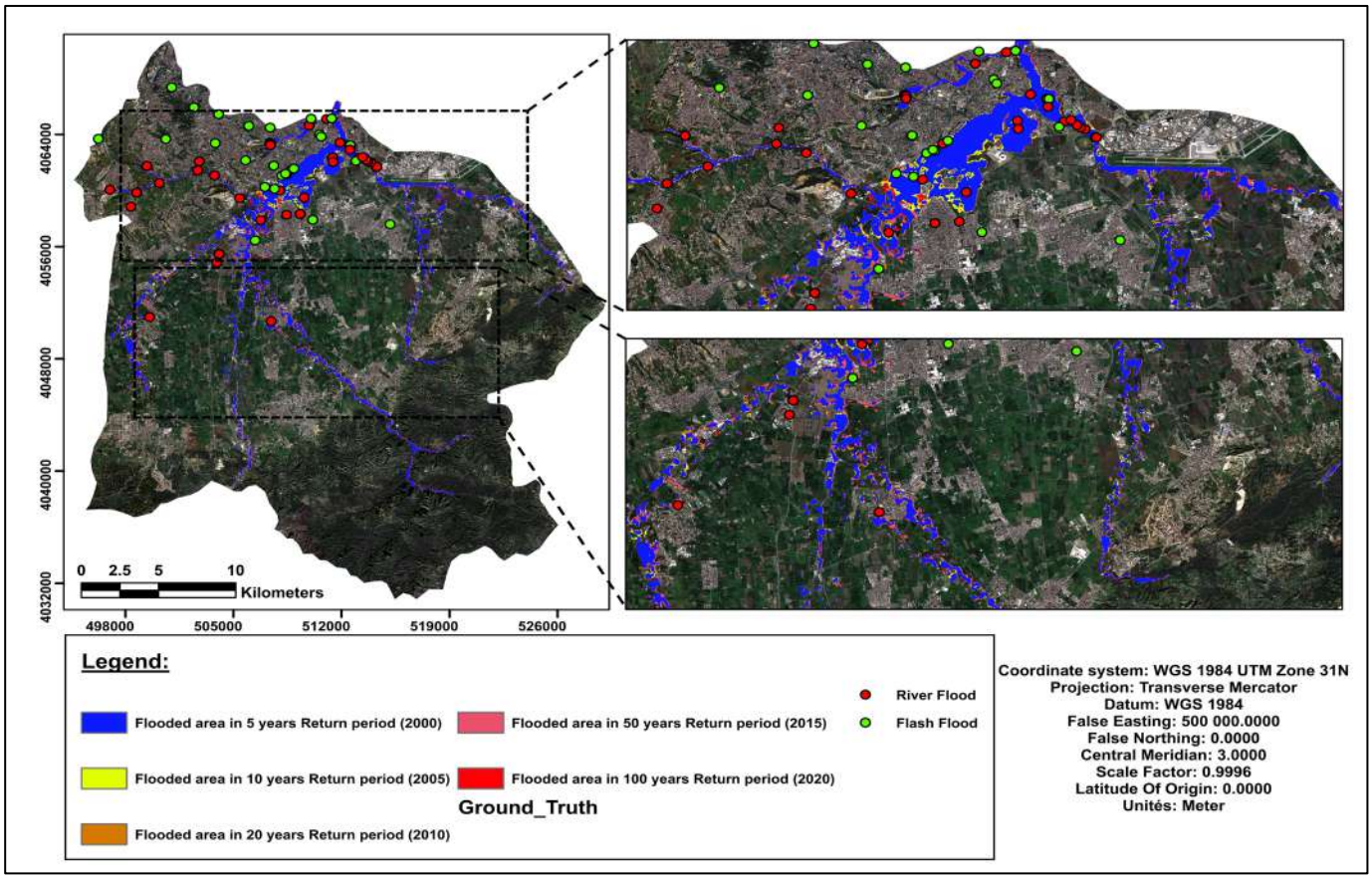

Figure 6. Flood-inventory data: Flooded area simulated by HEC-RAS on multiple scenarios and ground truth points.

5. Proposed Approach

In the present study, we introduce an advanced deep Res-Unet model designed to improve FSM utilizing a novel learning strategy and the integration of geospatial imagery. This approach shifts from traditional FSM techniques that predominantly rely on point-based data for target representation, moving towards a more holistic image-based conceptualization of both input data and targets. This enhancement allows for a comprehensive analysis of spatial information, surpassing the state of the art in feature space modulation. Moreover, our approach incorporates principles of physical modeling by fusing input data derived from the HEC–HMS/RAS models with our precisely defined ground truth for more accurate training. This methodology advances beyond the work of [3], who employed 1000 data points from HEC–RAS for training a random forest model. By emphasizing a ground truth-focused "images-to-images" modeling paradigm, we introduce a novel perspective in the field of FSM research.

We conducted numerous experiments across various sub-basins for training and testing in other sub-basins as shown in Figure 7, knowing that data outside each sub-basin was considered as No Data, employing different machine- and deep-learning algorithms, and experimenting with various loss functions to validate the efficacy of the proposed approach. Initially, we applied a feature selection in order to identify the factors contributing the most to the prediction. This was done by thresholding their influence values estimated by the random forest method. However, since it was observed that the model performances did not improve significantly, we decided to rely on all factors in the subsequent steps. Our experiments can be divided into three main steps, as illustrated in the flowchart in Figure 8.

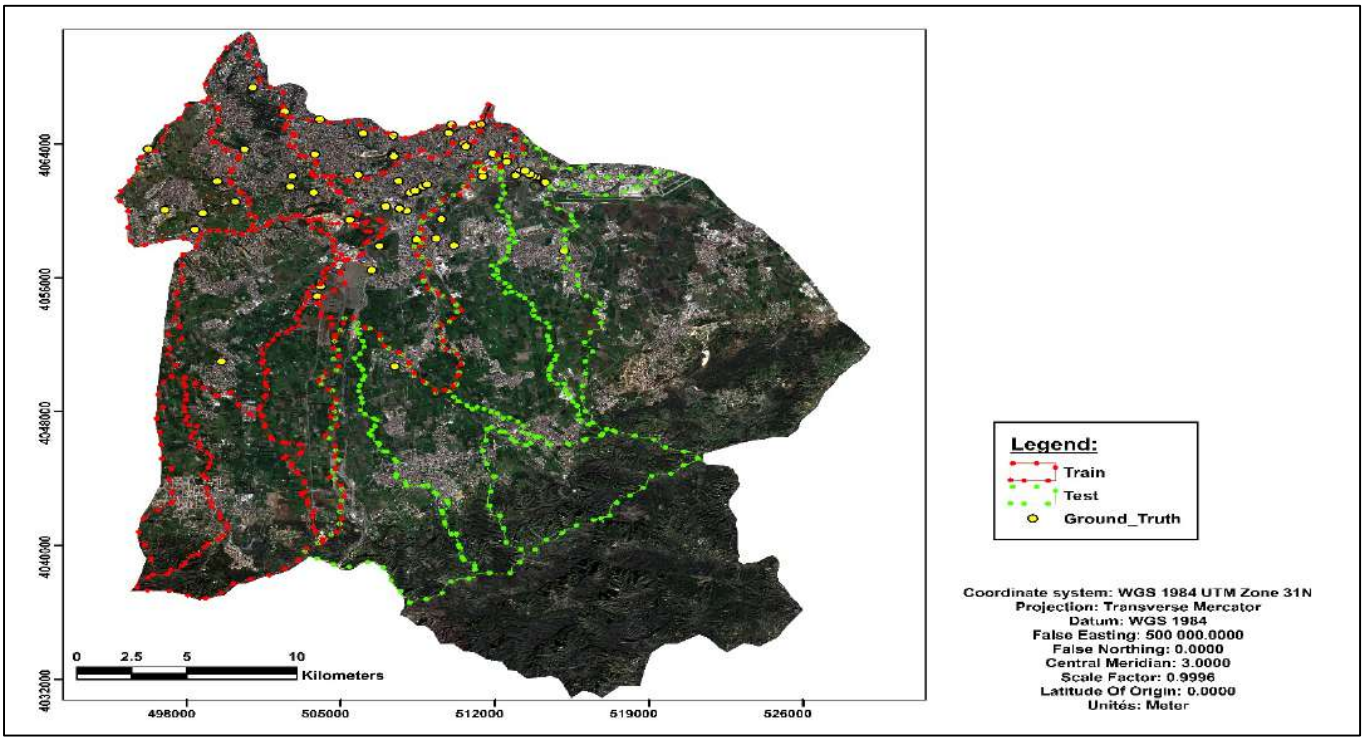

Figure 7. Sub-Basins division for training and test.

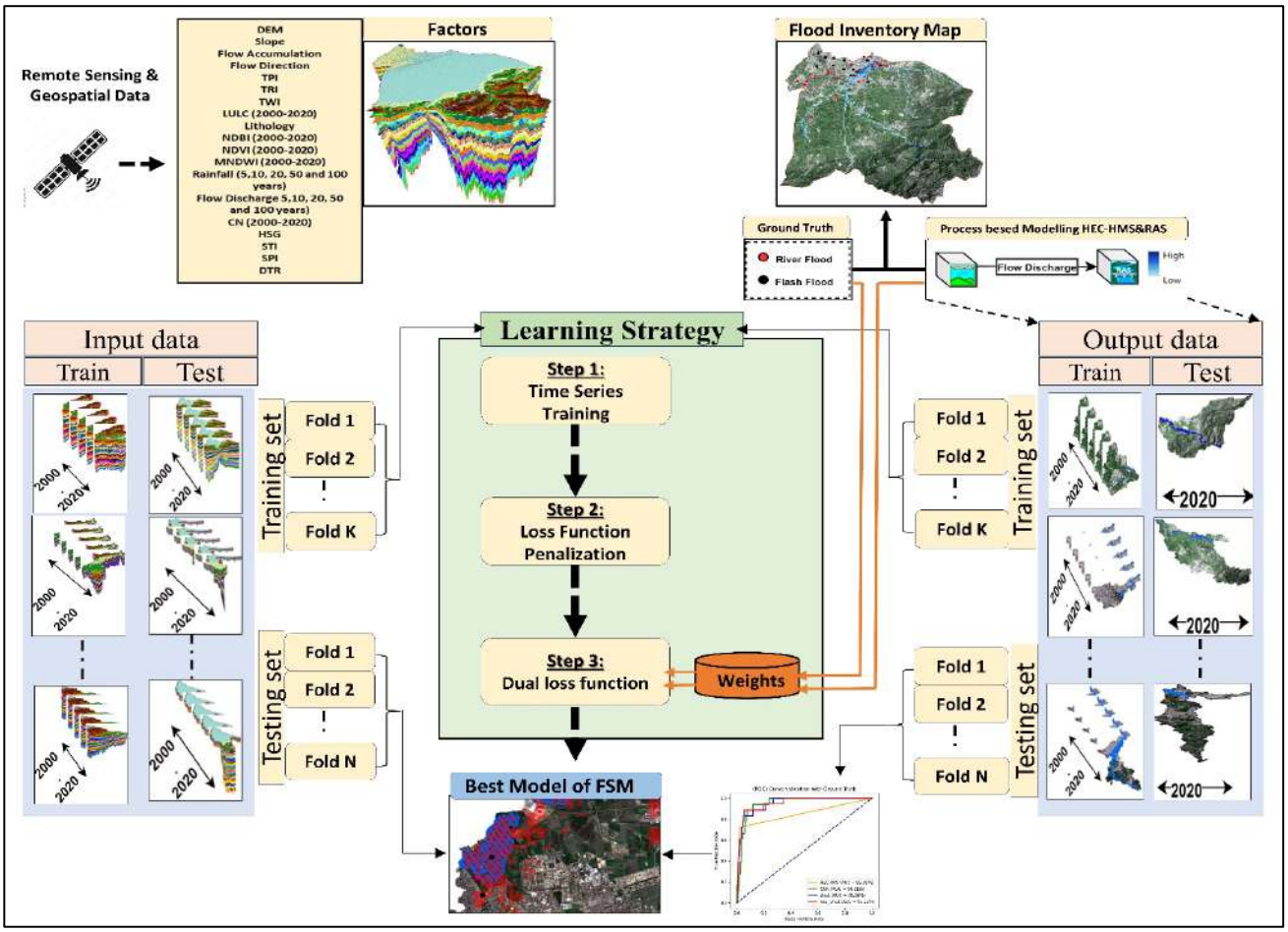

Figure 8. Flow chart of the detailed steps of this study.

5.1. Step 1: Time Series Modeling

This step involves conducting multiple training periods, each spanning different years such as 2000, 2005, 2010, 2015, and 2020. The objective of this phase is to assess the model performance across various time frames to identify which period yields the highest performance. This analysis is based solely on targets derived from the physical models HEC-HMS/RAS. The best-performing period is then selected for use in subsequent experiments.

5.2. Step 2: Loss-Function Comparison

In this phase, we compare two distinct training methodologies, both focused exclusively on the physical model target (HEC-HMS/RAS). The first methodology utilizes the "binary cross-entropy" loss function, and the second employs "sigmoid focal loss" to address data imbalance within the HEC-HMS/RAS model. The mathematical formulations for these loss functions are detailed below:

5.2.1. Binary Cross-Entropy Loss

To calculate the loss for binary classification tasks, we use the Binary Cross-Entropy Loss function as follows:

$$L(y, \hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}[y_i log(\hat{y}_i) + (1 - y_i) log(1 - \hat{y}_i)] \quad (23)$$

Here, y represents the ground truth label (either 0 or 1), and $\hat{y}$ denotes the predicted probability of class 1. The term $y_i log(\hat{y}_i)$ penalizes the model more if the prediction diverges from the true label, whereas $(1 - y_i) log(1 - \hat{y}_i)$ penalizes incorrect predictions of the negative class.

5.2.2. Sigmoid Focal Loss

The Sigmoid Focal Loss function, which introduces additional hyper parameters, is defined as follows:

$$L(y, \hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}\left[(1 - \hat{y}_i)^{\gamma} \times y_i \, log(\hat{y}_i) + (1 - y_i)^{\alpha} \times \hat{y}_i \, log(1 - \hat{y}_i)\right] \quad (24)$$

The "sigmoid focal loss" introduces two additional hyperparameters:

α and γ. Here, α controls the balance between the easy and hard examples, and γ focuses the model attention on the hard, misclassified examples. It is important to note that in our initial scenario, we exclusively utilized the "binary cross-entropy" loss function to train the model.

5.3. Step 3: Dual-Loss Function

In this scenario, we simultaneously employ two combined loss functions to optimize both physical modelling (HEC-HMS/RAS) and ground truth points. This approach ensures a balanced consideration between the outputs of the physical model and empirical data, thereby reducing the likelihood of neglecting crucial ground truth points. The weighted loss function integrates two branches, with an adjustable Alpha variable that varies with iterations, as depicted in the Figure 9 and the following equation:

$$Loss_{total} = (1 - Alpha) \times Loss_1 + Alpha \times Loss_2 \quad (25)$$

where $Loss_{total}$ is the weighted loss function, $Loss_1$ is the loss function for the ground truth based on Equation (24), $Loss_2$ is the loss function for the physical model also based on Equation (24). $Alpha$ is a scaler value that balances the contribution of the two losses.

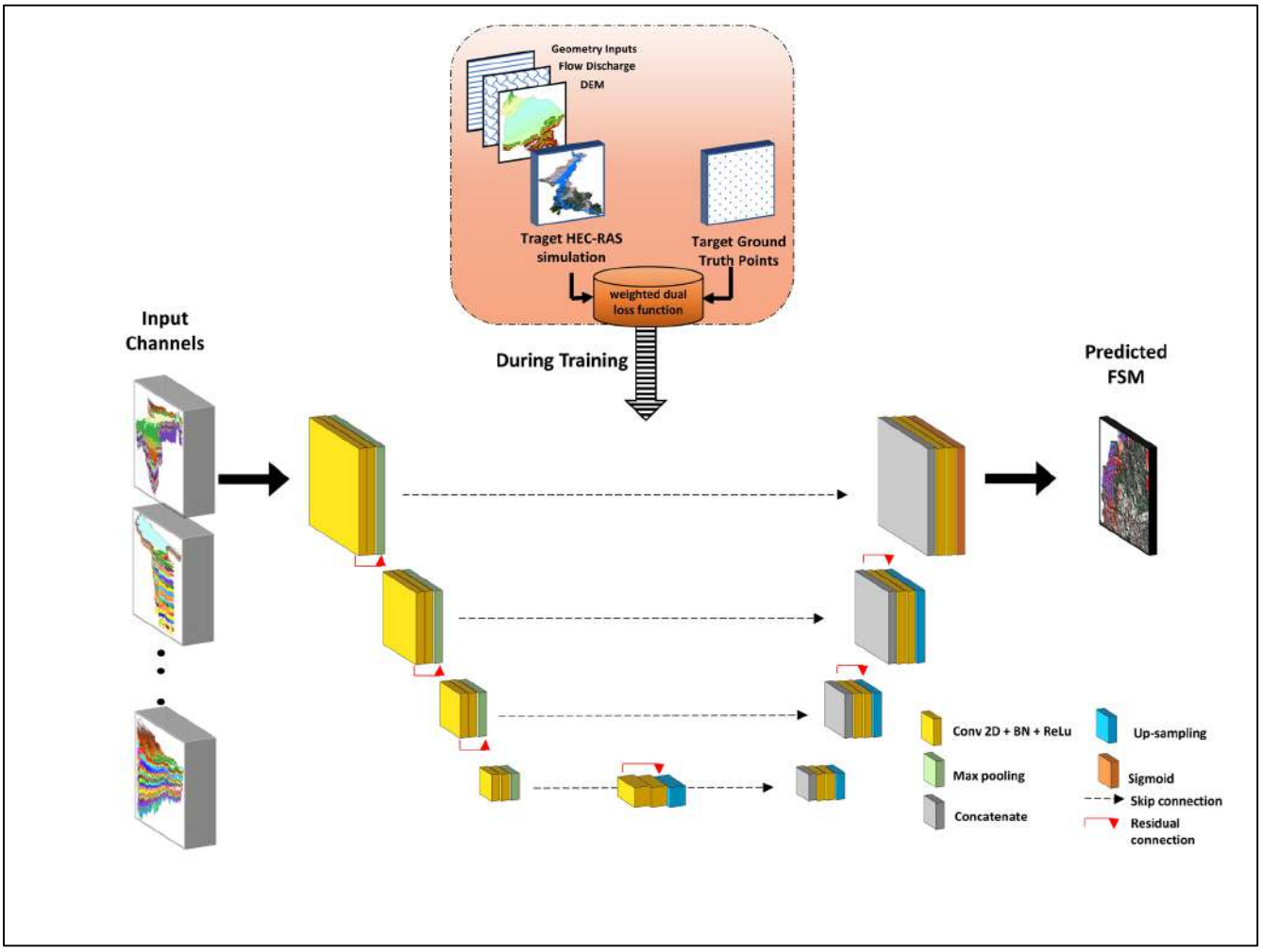

Figure 9. Architecture of our proposed weighted residual U-Net (W-Res-U-Net).

6. Results

Following our proposed FSM methodology, this section details and analyzes the results obtained at each step. For Scenarios 1 and 2, validation is conducted against the HEC–HMS/RAS physical model, while for Scenario 3, validation is conducted against both the physical model and ground truth points. We evaluate measures such as accuracy, specificity, sensitivity, and the AUC (ROC) curve, which represents the rate of true positives (sensitivity) versus the rate of false positives (1—specificity) at different thresholds. In this study, we considered a threshold of 0.5 for calculating performance, where <0.5 presents non-flooding and >0.5 presents flooding, and we focus on sensitivity values as they signify the detection of flooded areas. These measures are calculated using the following equations:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{26}$$

$$Specificity = \frac{TN}{TN + FP} \tag{27}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{28}$$

$$AUC = \frac{\sum TP + \sum TN}{P + N} \tag{29}$$

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives.

6.1. Step 1: Time Series Modelling Results

During this phase, the model underwent formation at various time intervals within specific sub-basins. This facilitated testing within the same sub-basins at alternate times and across different sub-basins. Validation was conducted against the HEC–RAS physical model, serving as our benchmark.

6.1.1. Scenario 1 Training and Test on the Same Sub-Basins

Initially, all models were trained within respective sub-basins and subsequently tested within the same sub-basins. As per Table 2, our analysis primarily focused on sensitivity values, emphasizing the detection accuracy of flooded areas. Notably, our findings indicate that models exhibit optimal performance when trained over multiple years (2000 to 2020), particularly evident when testing on earlier years. However, sensitivity diminishes in subsequent years due to higher return periods, indicating increased flood-prone areas, where we can notice:

- The RF model consistently underperforms when it comes to detecting flooded areas, with sensitivity levels ranging from 50.14% (during training in 2000 and testing in 2000) to 58.42% (during training from 2000 to 2020) during testing in 2000 and falling to 52.16% during testing in 2020.
- The CNN model demonstrates improved sensitivity, rising from 92.37% to 96.14% during multi-year training and tests in 2000, though it decreases to 94.62% in the 2020 test.
- The Unet model also exhibits sensitivity improvements, increasing from 92.12% to 96.51% with multi-year training and tests in 2000, followed by a reduction to 95.01% in the 2020 test.
- The Res–Unet model showcases the highest sensitivity gains, improving from 92.23% to 97.76% when training spans from the year 2000 to 2020 and tests in 2000, with a slight drop to 96.08% in the 2020 test.

6.1.2. Scenario 2 Training and Test on Different Sub-Basins

Subsequently, all models were trained on one set of sub-basins and tested on different ones. As shown in Table 3, the sensitivity analysis confirmed previous findings, demonstrating robust performance during multi-year training. However, sensitivity decreased in the following years due to increasing return periods.

- The RF model consistently fails to detect flooded areas in other sub-basins, maintaining a sensitivity of 0% across all training periods.
- The CNN model shows optimal performance with multi-year training, achieving a sensitivity of 62.71% when trained and tested in 2000, and improving to 63.52% and 64.45% from 2000 to 2020. However, sensitivity declines in subsequent years, with a rate of 61.23% in the 2020 test.
- The U-net model exhibits increased sensitivity with training from 2000 to 2020, achieving 66.9% in the 2000 test. However, there is a sensitivity decline in later years, dropping to 65.11% in the 2020 test.
- The Res–Unet model shows improved sensitivity with multi-year training, achieving 67.97% in the 2000 test. Nonetheless, it experienced a slight decline in the 2020 test, with sensitivity dropping to 67.36%, indicating a performance decrease compared to other models in subsequent years.

Our analysis reveals that the RF model faces challenges in generalizing across different sub-basins, underscoring its limitations. On the other hand, training over multiple years improves the model's generalization by capturing temporal variations, especially in precipitation and flow discharge patterns. Sensitivity tends to decline in later years due to longer return periods produced by HEC-RAS, resulting in more widespread flooding. However, the Res–UNet model demonstrates better sensitivity retention under larger return periods compared to other models.

Table 2. Training and test on the same sub-basins.

Period Training	Model	Test on 2000 (%)				Test on 2005 (%)				Test on 2010 (%)				Test on 2015 (%)				Test on 2020 (%)			
		Acc	Spe	Sen	AUC	Acc	Spe	Sen	AUC	Acc	Spe	Sen	AUC	Acc	Spe	Sen	AUC	Acc	Spe	Sen	AUC
Train on 2000	**RF**	96.26	97.31	50.14	73.12	96.08	96.82	49.64	72.38	95.82	96.64	49.14	71.42	96.48	96.3	48.64	70.92	96.32	94.12	48.34	70.56
	CNN	96.52	97.06	**92.37**	**97.53**	96.35	96.56	**91.37**	**96.57**	96.21	96.38	**90.75**	95.68	96.21	**90.38**	95.42	96.36	96.13	**90**	**95.16**	94.74
	U-net	**97.31**	97.32	92.12	97.21	**97.1**	96.98	91.21	96.23	**96.85**	96.36	90.11	**95.75**	96.24	89.88	**95.53**	**96.53**	95.99	89.42	95.2	**94.85**
	Res-Unet	97	**97.86**	92.23	97.47	96.85	**97.67**	91.09	96.89	96.23	**97.51**	90.56	95.29	**97.31**	90.12	95.04	96.14	**97.14**	89.68	94.85	94.69
Train from 2000 to 2005	**RF**	96.42	97.46	51.35	73.89	96.29	96.99	50.35	73.29	95.98	96.25	49.85	72.59	96.09	96.45	49.37	71.75	95.92	94.89	49.08	71.4
	CNN	96.62	98.23	**93.62**	**97.62**	96.46	**97.75**	**92.82**	**97**	96.25	**97.53**	**91.92**	**96.22**	**97.34**	**91.46**	**95.88**	**95.41**	**97.14**	**91.02**	**95.46**	**95.13**
	U-net	97.15	98.64	93.47	97.41	96.78	97.46	92.22	96.75	96.65	97.25	91.09	95.99	96.98	91	95.62	95.23	96.84	90.56	95.24	94.97
	Res-Unet	**97.23**	**98.82**	93.59	97.59	**97.03**	97.68	92.26	96.91	**96.81**	97.13	91.28	96.15	96.89	90.96	95.68	95.36	96.66	90.42	95.36	94.89
Train from 2000 to 2010	**RF**	97.02	97.87	51.98	74.08	96.47	96.68	50.43	73.52	96.21	96.42	49.82	72.24	96.2	97.13	49.45	71.34	96.01	95.06	49.08	71.02
	CNN	97.14	98.15	93.86	97.7	96.58	97.33	**93.31**	**96.99**	96.31	96.93	92.02	95.77	96.63	91.77	95.44	95.04	96.53	91.35	95.18	95.16
	U-net	97.53	98.76	94.12	97.84	**97.04**	**97.57**	93.22	96.89	96.57	**96.94**	**92.11**	95.88	**96.85**	**91.88**	**95.56**	**95.26**	**96.64**	**91.46**	95.3	**95.38**
	Res-Unet	**97.67**	**98.8**	**95.19**	**97.94**	96.99	97.38	93.14	96.95	**96.66**	96.8	91.95	**95.99**	96.72	91.79	95.49	95.23	96.51	91.32	**95.24**	95.29
Train from 2000 to 2015	**RF**	97.29	98.12	56.19	74.81	96.57	97.17	55.19	74.16	96.34	96.85	54.19	73.06	**96.58**	97.41	53.64	72.16	96.38	95.87	53.22	71.84
	CNN	97.39	98.37	94.89	97.82	96.8	**97.89**	94.39	**97.45**	96.52	**97.68**	93.39	96.17	97.55	92.05	95.66	95.32	**97.43**	91.65	95.34	95.22
	U-net	97.67	98.82	96.32	98.11	96.99	97.69	**95.32**	97.22	96.66	97.29	94.22	96.52	97.09	92.72	**95.91**	**95.64**	97.02	92.18	95.58	95.62
	Res-Unet	**97.81**	**98.91**	**97.04**	**98.23**	**97.12**	97.68	96.02	97.36	**96.78**	97.19	**94.56**	**96.65**	97.02	**92.88**	95.84	95.58	96.97	**92.32**	**95.69**	**95.81**
Train from 2000 to 2020	**RF**	97.48	98.32	58.42	75.26	97.25	97.69	57.42	74.57	96.89	97.09	56.42	73.32	96.98	97.87	55.32	72.48	96.88	96.15	55.02	72.16
	CNN	97.54	98.48	95.04	97.93	97.28	97.99	94.55	97.63	97.04	97.64	93.22	96.54	97.45	92.32	95.78	96.04	97.35	91.72	95.62	95.41
	U-net	97.84	**98.95**	96.51	98.41	97.6	98.32	95.91	97.98	97.3	97.94	94.51	96.89	97.73	92.81	96.12	96.37	97.6	92.48	95.98	95.78
	Res-Unet	**97.91**	97.78	**97.76**	**98.56**	**97.64**	**98.51**	**96.24**	**98.17**	**97.46**	**98.12**	**95.24**	**97.07**	**97.99**	**93.14**	**96.29**	**96.89**	**97.87**	**92.72**	**96.08**	**95.96**

Table 3. Training and testing on the other sub-basins.

Period Training	Model	Test on 2000 (%)				Test on 2005 (%)				Test on 2010 (%)				Test on 2015 (%)				Test on 2020 (%)			
		Acc	Spe	Sen	AUC	Acc	Spe	Sen	AUC	Acc	Spe	Sen	AUC	Acc	Spe	Sen	AUC	Acc	Spe	Sen	AUC
Train on 2000	**RF**	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50
	CNN	96.84	95.6	**62.71**	**95.69**	97.63	95.89	**62.12**	**95.42**	97.02	97.12	**61.29**	**95.96**	96.9	95.22	**61.14**	**95.76**	97.03	96.13	**60.24**	**95.74**
	U-net	97.08	96.16	62.47	95.62	96.91	95.71	61.87	95.22	97.5	97.98	60.92	95.23	97.04	97.08	60.89	95.27	96.96	96.73	60.21	95.19
	Res-Unet	97	96.02	62.61	96.64	95.62	95.57	61.92	95.36	97.12	97.8	61.14	95.42	96.98	97.02	60.95	95.31	96.69	96.58	60.23	95.01
Train from 2000 to 2005	**RF**	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50
	CNN	96.51	95.41	**63.52**	**95.62**	97.51	96.05	**62.21**	**96.01**	96.23	96.96	**61.57**	**95.35**	95.08	95.42	**61.21**	**95.74**	96.27	95.63	**60.97**	**94.86**
	U-net	96.81	95.85	63.17	95.47	96.87	96.13	61.90	95.69	97.78	97.96	61.13	94.89	96.75	97.15	61.01	95.51	97.25	96.87	60.54	94.54
	Res-Unet	96.74	96.31	63.44	96.54	95.54	96.21	62.11	95.74	97.56	97.91	61.25	94.97	96.67	97.11	61.13	95.59	96.98	96.93	60.72	94.62
Train from 2000 to 2010	**RF**	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50
	CNN	96.87	95.79	63.72	95.73	97.42	96.12	63.17	96.62	94.69	95.82	61.84	95.81	94.61	95.37	61.61	94.98	94.31	95.13	61.26	94.65
	U-net	97.01	96.12	64.87	95.96	96.82	96.21	64.01	97.16	96.51	97.71	65.85	96.13	96.24	97.02	65.89	95.96	96.05	96.63	65.04	95.87
	Res-Unet	97.15	96.96	**65.19**	**96.26**	95.92	96.38	**64.17**	**97.87**	96.63	97.23	**66.87**	**96.41**	96.23	97.05	**67.19**	**96.68**	96.11	96.59	**66.89**	**96.7**
Train from 2000 to 2015	**RF**	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50
	CNN	97.27	96.12	64.12	96.82	97.12	96.8	63.29	96.42	94.87	95.77	62	96.14	94.74	95.56	61.87	95.17	94.63	95.32	61.49	94.97
	U-net	97.51	97.42	66.52	96.14	96.86	96.73	63.78	96.97	96.86	97.25	66.12	96.17	96.51	97.11	66.07	96.21	96.21	96.85	65.27	96.06
	Res-Unet	97.58	97.61	**67.04**	**96.87**	96.74	96.05	**64.93**	**98.09**	97.01	97.07	**67.73**	**96.98**	96.44	97.15	**67.53**	**96.89**	96.33	96.94	**67.02**	**96.7**
Train from 2000 to 2020	**RF**	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50	**97.96**	**100**	0	50
	CNN	97.31	96.24	64.45	96.74	97.25	96.11	64.35	96.23	94.97	95.83	62.54	95.84	94.93	95.73	62.03	95.36	94.81	95.52	61.23	95.02
	U-net	97.64	97.63	66.9	97	97.59	97.45	66.86	96.89	97.1	97.34	66.28	96.52	96.87	97.23	66.96	96.39	96.78	97.01	65.29	95.23
	Res-Unet	97.69	97.78	**67.97**	**97.98**	97.62	97.58	**67.81**	**97.95**	97.32	97.42	**67.74**	**97.12**	96.61	97.29	**67.54**	**97.02**	96.52	97.24	**67.36**	**95.67**

Our validation, focused on the period encompassing maximum flooded zones in 2020 (or 100-year return period), is illustrated by the ROC curve in Figure 10 and the maps in Figures 11 and 12. The maps show that our target (HEC-RAS) has a significant number of "0" pixels representing non-flooded areas, which are larger than the "1" pixels representing flooded areas. This imbalance may influence our prediction results, as the binary cross-entropy loss function was used in this step. Therefore, in the next step, we will consider a loss function that accounts for this target imbalance.

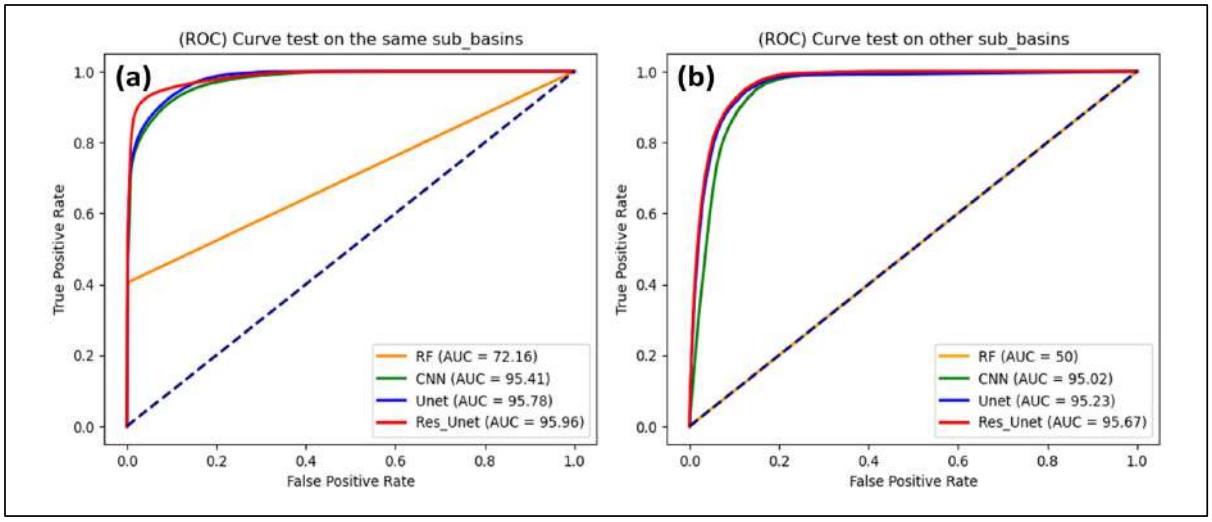

Figure 10. Roc curve train from 2000 to 2020, test in 2020 of: (**a**) same sub-basins; (**b**) other sub-basins.

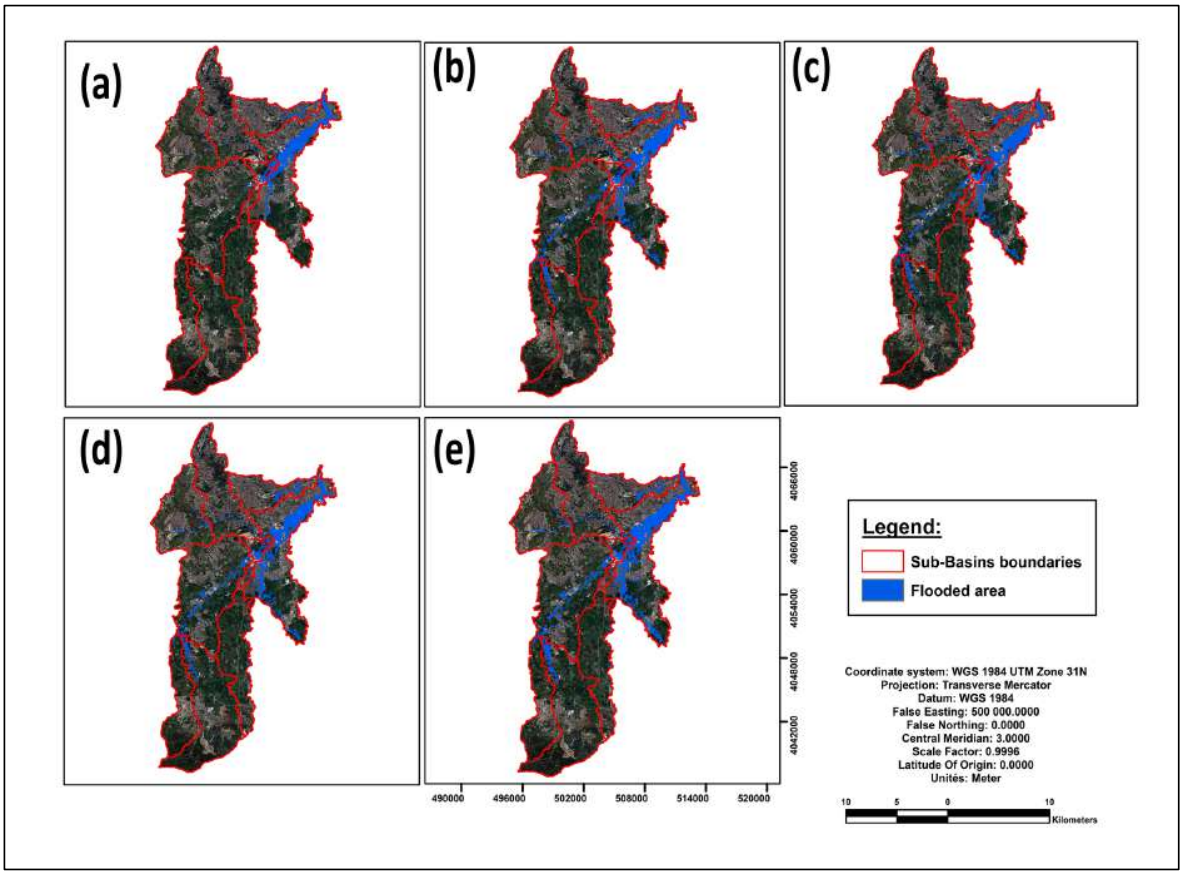

Figure 11. Flooding maps Train from 2000 to 2020, test on 2020 in the same sub-basins for: (**a**) RF; (**b**) CNN; (**c**) Unet; (**d**) Res-Unet; and (**e**) HEC-HMS and RAS simulation.

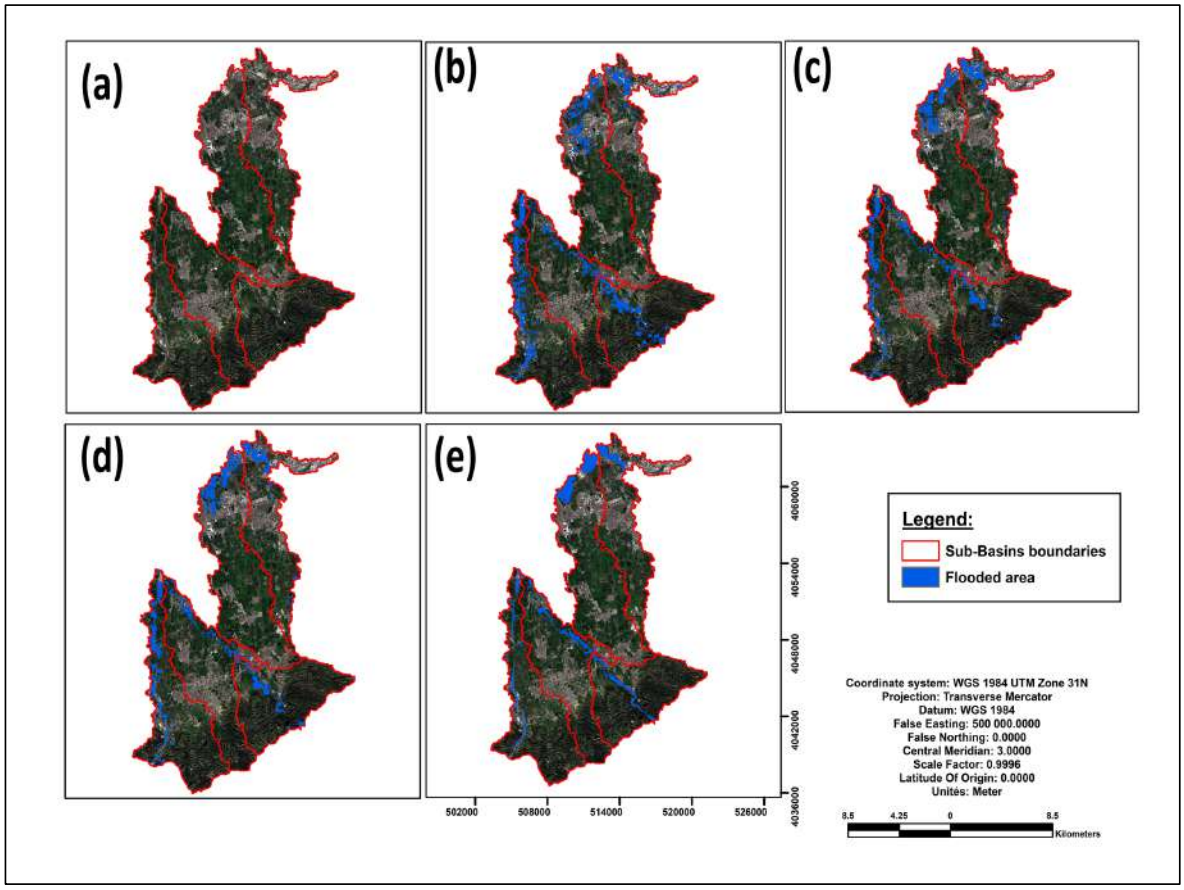

Figure 12. Flooding maps train from 2000 to 2020, test in 2020 in other sub-basins for: (**a**) RF; (**b**) CNN; (**c**) Unet; (**d**) Res-Unet; and (**e**) HEC-HMS and RAS simulation.

6.2. Step 2: Lost-Function Comparison

Based on the results of Step 1, as previously mentioned, the target generated by the HEC–RAS simulation contains a higher number of non-flooded pixels ("0") compared to flooded pixels ("1"). This imbalance can skew deep-learning models towards focusing on the more frequent "0" values, reducing attention to the less frequent "1" values that represent flooded areas. To address this, we compared the "Sigmoid Focal Loss" function with the "Binary Cross-Entropy Loss" function, using parameters $\alpha = 0.75$ and $\gamma = 0.25$. As shown in Table 4 and the ROC curve in Figure 13, this adjustment slightly improved sensitivity: from 61.23% to 61.46% for the CNN model, from 65.29% to 65.39% for the U-Net model, and from 67.36% to 67.63% for the Res–UNet model. These improvements are also visually evident in the maps presented in Figure 14.

Table 4. Metrics improvement using Sigmoid focal loss.

Period Training	Model	Test in 2020 in Other Sub-Basins			
		Acc	Spe	Sen	AUC
From 2000 to 2020	CNN	95.09 (+0.28)	95.73 (+0.21)	61.46 (+0.23)	95.34 (+0.32)
	Unet	96.92 (+0.14)	97.42 (+0.41)	65.39 (+0.10)	95.38 (+0.15)
	Res-Unet	**97.07** (+0.55)	**97.69** (+0.45)	**67.63** (+0.27)	**95.84** (+0.17)

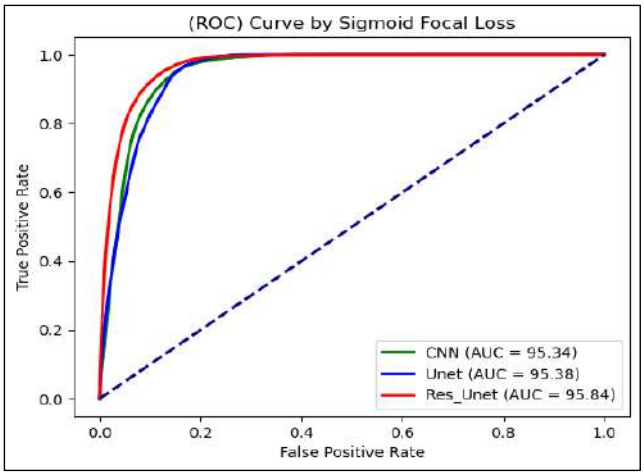

Figure 13. ROC curve trains from 2000 to 2020 and test on other sub-basins in 2020 using Sigmoid focal loss function.

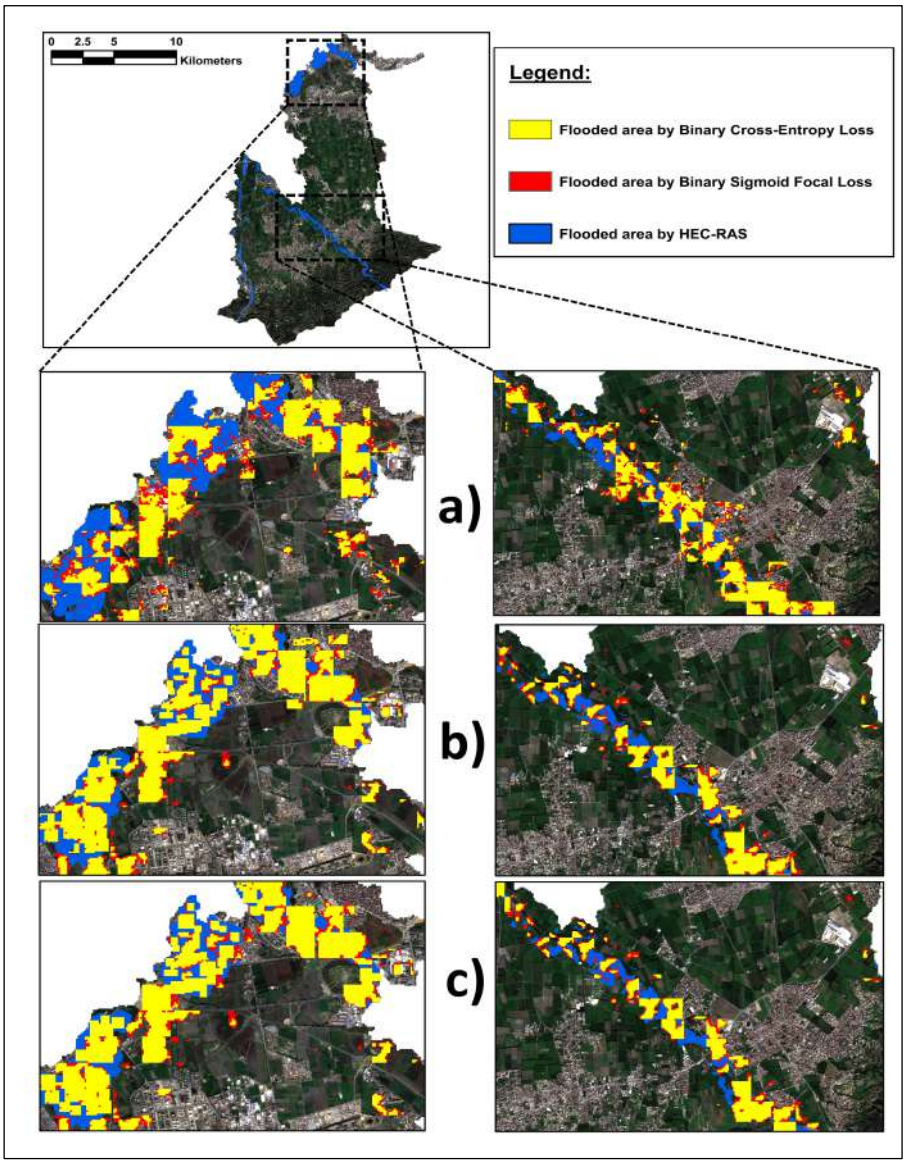

Figure 14. Flood maps by comparing the "Sigmoid Focal Loss" function with the "Binary Cross-Entropy Loss" function for: (**a**) CNN, (**b**) Unet, (**c**) Res-Unet.

6.3. Step 3: Dual-Loss Function

Based on the results obtained in the previous step, we validated our models using the HEC–RAS simulation. However, this validation remains somewhat unclear because our deep-learning models were trained using the HEC–RAS simulation as a target. To enhance the realism of our modeling, we used 63 ground truth points: 41 for training our models and 22 for testing in other sub-basins, in combination with the HEC–RAS physical simulation. This combination was achieved by applying two weighted loss functions simultaneously, with various Alpha values to balance the contributions of the two losses, ensuring that the ground truth points were not neglected. From the analysis of Table 5 and the graphs in Figure 15, we observed the following:

Table 5. Dual-loss function metrics performance with multi-Alpha scenarios.

Model	Alpha	Validation with HEC-RAS				Validation with Ground Truth			
		Acc	Spe	Sen	AUC	Acc	Spe	Sen	AUC
CNN	**0**	54.46	53.52	**99.45**	75.59	52.44	52.44	**100**	71.12
	0.1	95.03	95.81	57.55	95.32	94.72	94.72	66.67	94.91
	0.2	96.29	97.42	42.14	95.25	96.61	96.61	55.56	96.27
	0.3	96.23	97.47	36.72	95.08	96.77	96.77	44.44	94.28
	0.4	96.61	98	30.06	94.89	97.43	97.43	44.25	94.98
	0.5	96.66	98.09	28.17	94.88	97.55	97.56	50	95.04
	0.6	96.64	97.9	36.16	95.02	97.2	97.2	50	**96.72**
	0.7	**97.08**	98.47	30.36	**95.87**	97.88	97.88	38.89	96.22
	0.8	97	**98.53**	23.62	94.91	**98.08**	**98.08**	33.33	95.56
	0.9	96.84	98.14	34.27	95.35	97.48	97.48	38.89	94.25
	1	95.09	95.73	61.46	95.34	96.72	96.72	50	95.95
Unet	**0**	54.29	53.34	99.89	77.5	52.26	52.26	**100**	74.35
	0.1	92.72	93.31	64.41	93.28	92.14	92.14	83.33	95.06
	0.2	94.41	95.16	58.03	94.31	94.08	94.08	72.11	93.51
	0.3	93.31	93.96	62.08	93.93	92.82	92.82	69.01	93.04
	0.4	96.19	97.38	38.74	93.01	96.65	96.65	55.56	93.27
	0.5	95.24	96.36	41.17	91.89	95.6	95.6	44.44	93.56
	0.6	96.43	97.8	30.45	92.22	97.23	97.23	61.11	95.55
	0.7	96.66	98.03	30.82	93.28	97.44	97.44	44.44	94.444
	0.8	**96.92**	**98.34**	29.18	94.63	**97.77**	**97.77**	55.56	**96.94**
	0.9	96.42	97.79	31.12	92.62	97.19	97.2	38.89	91.89
	1	**96.92**	97.42	**65.39**	**95.38**	97.22	97.22	72.22	96.72
W-Res-U-Net	**0**	54.29	53.34	99.89	77.57	52.26	52.26	**100**	76.56
	0.1	90.73	91.14	**71.16**	92.95	93.07	93.07	88.89	95.13
	0.2	94.17	95.02	59.24	93.32	94.04	94.04	72.42	93.83
	0.3	95.44	96.64	38.22	91.63	95.92	95.93	44.44	92.93
	0.4	93.57	94.07	69.81	93.99	92.76	92.76	46.65	93.14
	0.5	96.38	97.89	23.84	91.75	97.44	97.45	16.67	92.18
	0.6	96.22	97.43	38.15	93.19	96.7	96.7	72.22	95.5
	0.7	96.32	97.61	36.97	91.74	96.91	96.91	50	94.26
	0.8	**97.35**	**98.96**	20.4	92.83	**98.56**	**98.56**	27.78	94.28
	0.9	94	94.65	62.59	93.22	93.49	93.49	81.84	**96.21**
	1	97.07	97.69	67.63	**95.84**	97.28	97.28	77.78	96.12

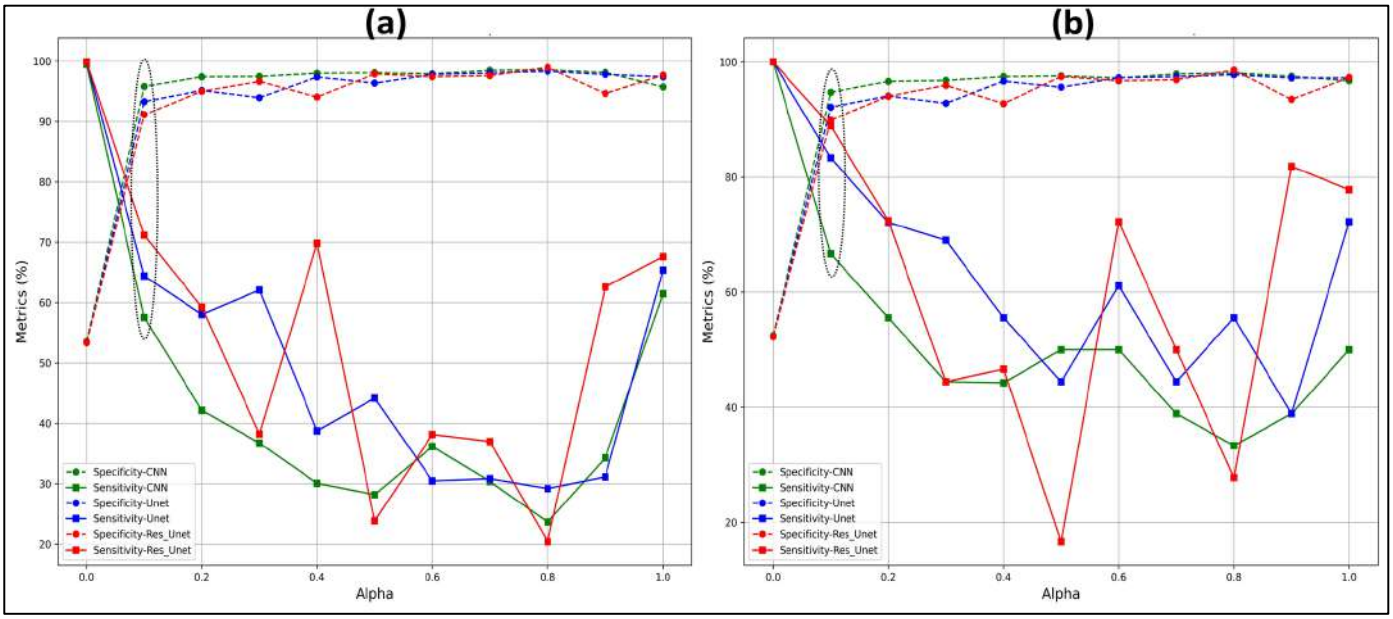

Figure 15. Metrics validation (Sensitivity and Specificity) versus Alpha values with: (**a**) HEC–RAS simulation. (**b**) Ground truth points.

6.3.1. CNN Model

When validating with HEC-RAS, the highest sensitivity (99.45%) was observed at **Alpha** = 0, but with a specificity of 53.52%, indicating flooding across the entire test region without considering non-floodable areas, which is unrealistic. At **Alpha** = 1, focusing solely on HEC-RAS, the sensitivity was 61.46%. The most applicable sensitivity was at **Alpha** = 0.1, with 57.55% sensitivity and 95.81% specificity.

When validating against ground truth points, the sensitivity was 100% at **Alpha** = 0, focusing solely on the ground truth, with a specificity of 52.44%. At **Alpha** = 1, the sensitivity was 50%. The optimal sensitivity was achieved at **Alpha** = 0.1, with 66.67% sensitivity and 94.72% specificity.

6.3.2. U-Net Model

When validating with HEC-RAS, the highest sensitivity (99.89%) was at **Alpha** = 0, but with a specificity of 53.34%. Focusing only on HEC-RAS (**Alpha** = 1) yielded a sensitivity of 65.39%. The most logical balance was achieved at **Alpha** = 0.1, with 64.41% sensitivity and 93.31% specificity.

When validating against ground truth points, the best sensitivity was at **Alpha** = 0.1, with 83.33%.

6.3.3. W-Res-U-Net Model

When validating with HEC-RAS, the highest sensitivity (71.16%) was at **Alpha** = 0.1, with a specificity of 91.14%. Focusing only on HEC-RAS (**Alpha** = 1) yielded a sensitivity of 67.02% while focusing only on ground truth points resulted in a sensitivity of 99.89% but with a specificity of 53.34%.

When validating against ground truth points, the best sensitivity was at ***Alpha*** = 0.1, with 83.33%.

In summary, for all models, the optimal ***Alpha*** value was consistently around 0.1, providing a balanced approach that incorporates both the HEC–RAS simulation and the ground truth points effectively (as shown in Figure 16). So, we can see the performance and dominance of model Res-Unet either for validation with the simulation generated by HEC-RAS or with ground truth points.

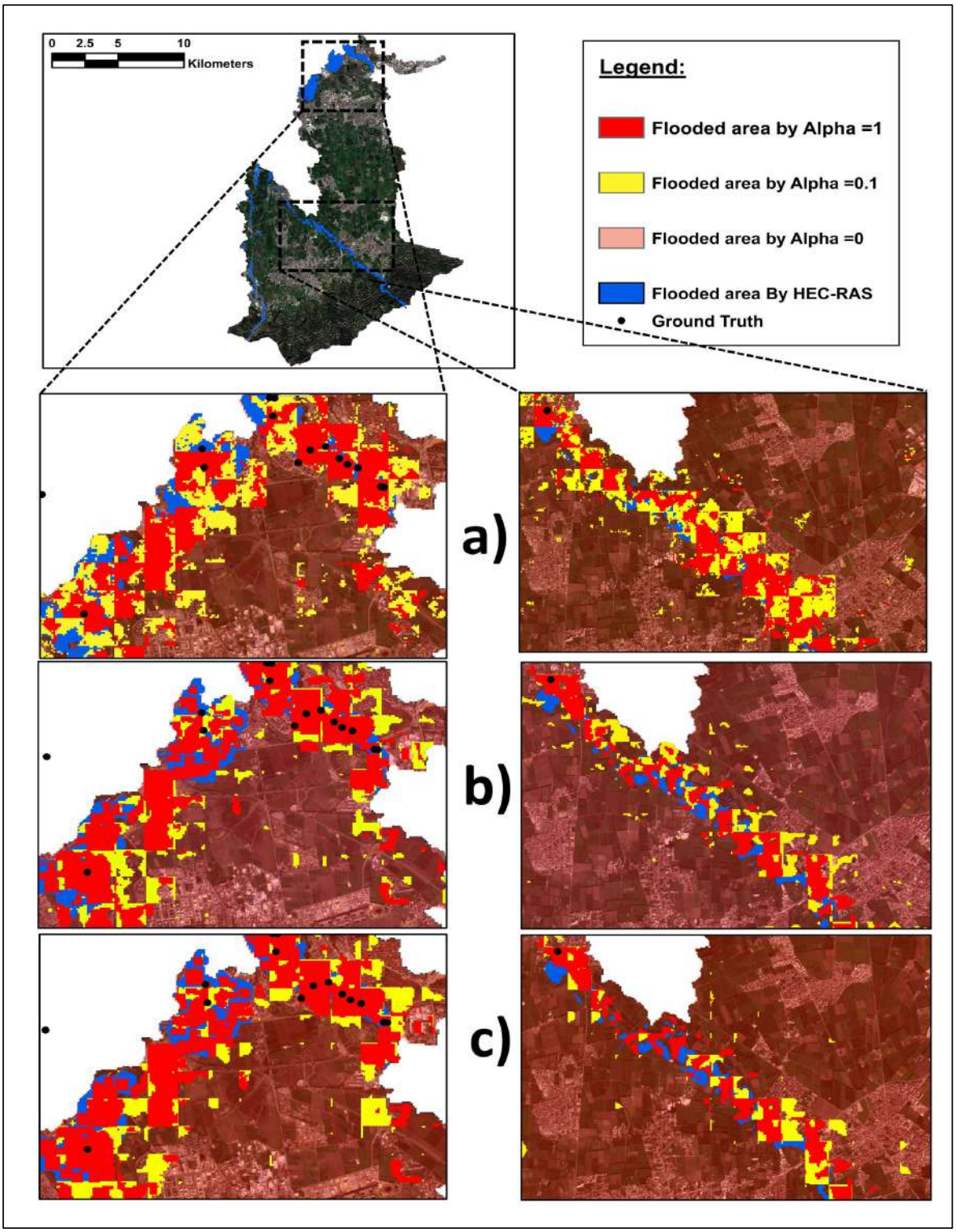

Figure 16. Flood maps by comparing the Alpha values "0", "0.1", "1" for: (**a**) CNN, (**b**) Unet, (**c**) W-Res-U-Net.

According to the last result, which combines two loss functions at the same time, based on the physical modeling of HEC-RAS without neglecting the ground truth points, also we evaluated the performance of the HEC-RAS simulation against these ground truth points. The findings are presented in Table 6.

Table 6. Validation of HEC-RAS simulation with the Ground Truth points.

Simulation	Validation on Ground Truth			
	Acc	Spe	Sen	AUC
HEC-RAS	97.96	97.96	72.22	85.09

Our research employed HEC-RAS physical simulation combined with ground truth points for modeling. The validation results revealed that the U-net and W-Res-U-Net models demonstrated higher sensitivity than the HEC–RAS simulation, with values of 83.33% and 88.89%, respectively, compared to 72.22% for the HEC–RAS simulation. Conversely, the CNN model exhibited a lower sensitivity of 66.67%. In terms of specificity, U-net and W-Res-U-Net achieved 92.14% and 93.07%, respectively, while the HEC–RAS simulation and CNN model attained 97.96% and 94.72%, respectively. These findings indicate that the Res–U-net model excels in detecting flooding points, with the highest AUC of 95.13% (as shown in Figure 17), followed by U-net at 95.06 and CNN at 94.91%, whereas the HEC–RAS simulation has an AUC of 85.09%.

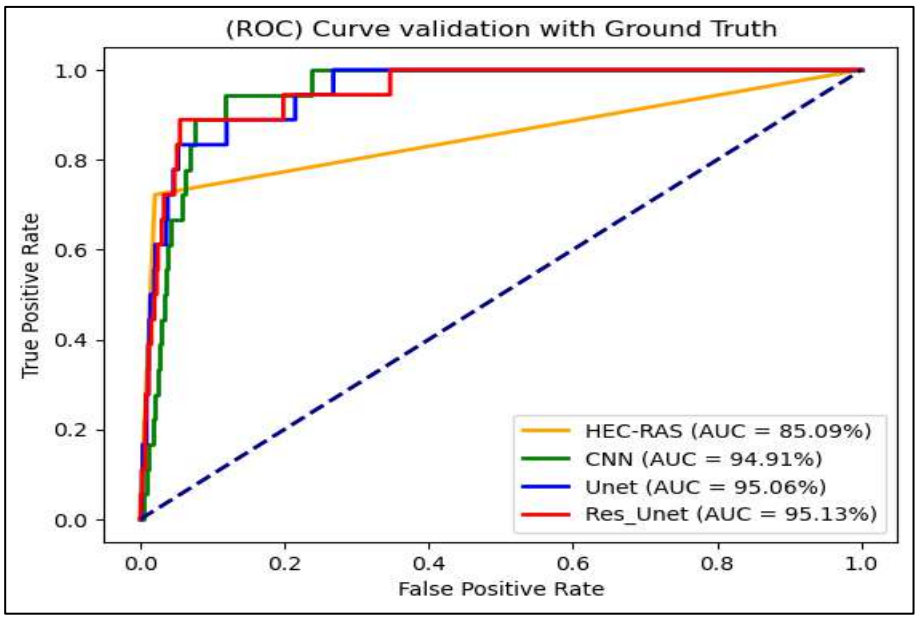

Figure 17. ROC curve validation with ground truth points.

Based on the latest results and Figure 18, it is evident that the U-net and W-Res-U-Net models outperform the HEC–RAS physical simulation in detecting flood points. Despite being trained on HEC–RAS data and 41 ground truth points, U-net and Res-U-net demonstrate superior performance when tested on 22 points in other sub-basins. This improvement can be attributed to the focused incorporation of ground truth points during training and the inherent limitations of HEC-RAS, particularly in defining cross-sections and riverbanks. The W-Res-U-Net model architecture significantly contributes to this enhanced performance. It employs convolutional layers, each followed by batch normalization and an activation function (ReLU). The convolutional layers use filters of varying sizes (e.g., 5×5, 3×3, 2×2) to capture spatial features at different scales. Batch normalization helps stabilize and accelerate the training process by normalizing the output of each layer. The model use of residual connections, where the input to a layer is added to its output, aids in training deeper networks by mitigating the vanishing gradient problem. These residual blocks are pivotal in maintaining the integrity of features through the network's depth. Additionally, the W-Res-U-Net model integrates multiple filters with increasing numbers (16, 32, 64, 128, 256, 512) as the network goes deeper, enhancing its ability to capture complex patterns. The up-sampling and concatenation

blocks ensure that high-resolution features from the encoder are effectively combined with the decoder's up-sampled features, leading to precise segmentation outputs. Conversely, the CNN model identifies ground truth points outside the cross-sections defined by the HEC-RAS simulation but fails to detect certain points within these cross-sections.

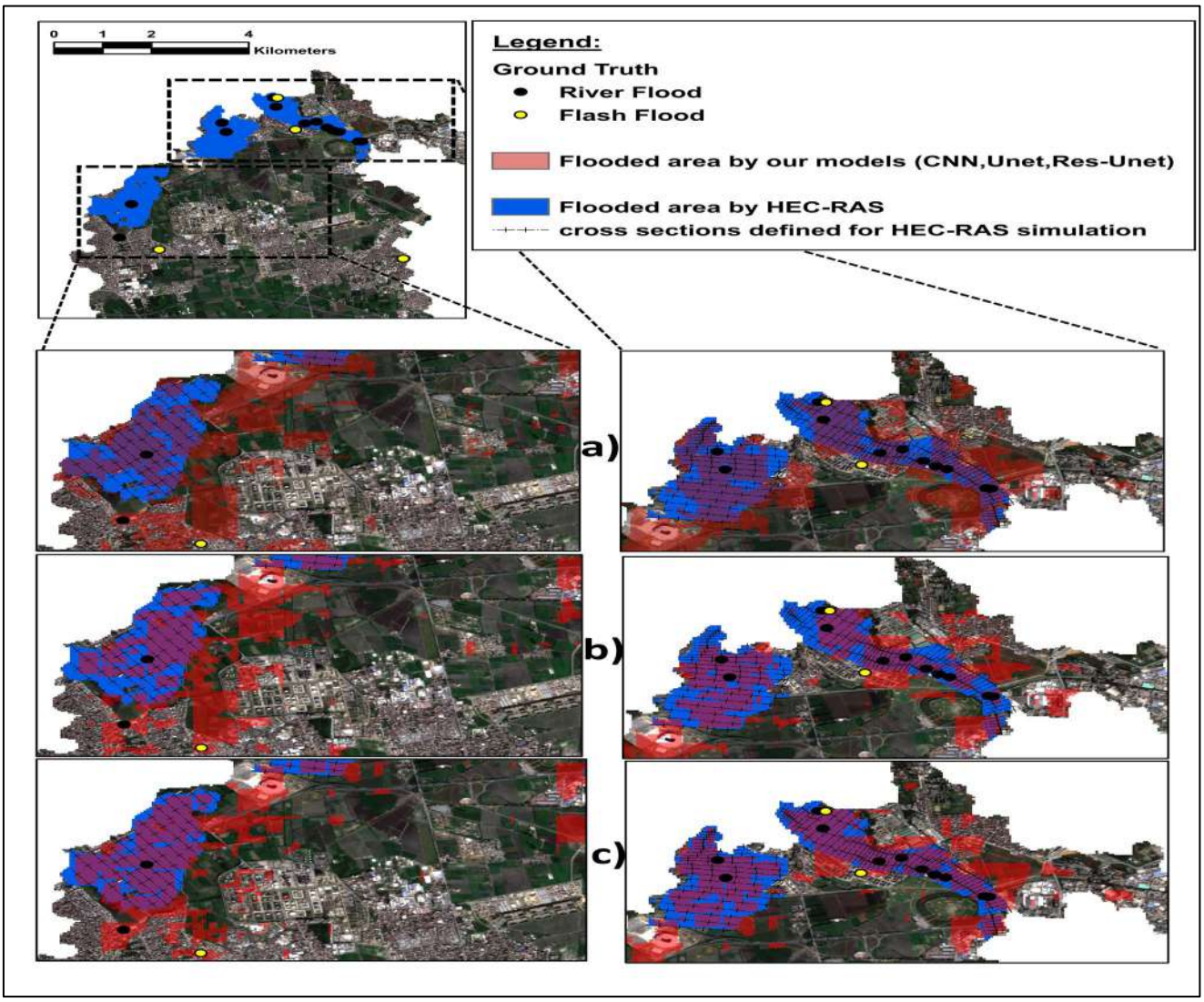

Figure 18. Validation with ground truth points by comparing the limits of HEC–RAS simulation for: (**a**) CNN, (**b**) Unet, (**c**) W-Res-U-Net.

7. Discussion

The Res–UNet and W–Res–UNet models consistently outperformed the other models in terms of sensitivity, demonstrating their robustness in detecting flooded areas, even under challenging conditions with high return periods. The incorporation of residual blocks in these models likely played a critical role in their ability to capture and retain important features, leading to better segmentation and flood detection. Additionally, the use of ground truth points in the validation process further reinforced the generalizability and accuracy of these models.

In our study, we addressed the limitation of ground truth data by proposing a weighted loss function in our deep-learning models. This approach mitigated the imbalance between flooded and non-flooded areas in the training data, enhancing the model's ability to detect flood-prone regions accurately. Unlike traditional state-of-the-art methods that rely on point-to-point flood-susceptibility mapping, our approach is based on image-to-

image segmentation for flood-susceptibility mapping. This method allows for a more comprehensive spatial analysis, capturing complex spatial dependencies and patterns, which is particularly beneficial for modeling intricate flood phenomena. Our focus on sensitivity metrics provided precise flood susceptibility assessments, further enhancing the practical utility of the models.

Overall, our results indicate that while physical models like HEC-RAS provide valuable insights into hydrological processes, the integration of advanced deep-learning models—particularly those utilizing residual connections and tailored loss functions—can significantly enhance flood-prediction accuracy. However, challenges remain in generalizing the models across spatially diverse regions and varying hydrological conditions. Addressing these challenges through the use of advanced loss functions and real-world validation data, as demonstrated in this study, can help improve the performance of flood-prediction models in future applications.

The performance of the Res–UNet and W–Res–UNet models, in particular, points to the potential of using hybrid modeling approaches that combine physical models with advanced deep-learning techniques. These models demonstrated superior flood-detection capabilities, especially when tested on unseen data, underscoring their potential for real-world applications. However, further research is needed to explore how these models can be optimized for different hydrological and geographic conditions and to address the inherent challenges in comparing image-based segmentation methods with point-based approaches.

Strengths and Weaknesses:

- **Strengths:** A key strength of this study is the integration of physical models with deep-learning techniques, leading to improved flood-prediction accuracy. By employing an image-to-image segmentation approach, we moved beyond traditional point-to-point methods, allowing for more detailed spatial analysis and capturing the complex spatial patterns of flooding. The use of a weighted loss function addressed the imbalance in the target data, enhancing the model's sensitivity to flooded areas. Additionally, the inclusion of real-world ground truth points alongside simulated data enabled a more comprehensive evaluation of the model's performance and reinforced its generalizability.
- **Limitation:** A limitation of this study is the difficulty in generalizing model performance across sub-basins with diverse hydrological characteristics. The comparison between our image-based segmentation approach and traditional point-based methods is also challenging due to fundamental differences in methodology and data representation. Furthermore, despite the improvements made through loss function adjustments, the imbalance in the target data (flooded versus non-flooded areas) poses an ongoing challenge. Future work is needed to further optimize the models for different hydrological and geographic conditions and to develop standardized methods for comparing different flood-susceptibility mapping approaches.

8. Conclusions

FSM is a critical strategy for flood forecasting and disaster prevention. Traditional FSM approaches often rely on unsupervised methods, such as the AHP, which require expert knowledge to define factor weights. Supervised machine-learning and deep-learning methods, while promising, demand extensive datasets for model training. Remote sensing, another common technique, is limited by the availability of imagery and weather conditions, such as cloud cover during short-term events. To address these challenges, we propose a novel hybrid strategy that integrates deep learning with physical modeling using HEC-HMS and HEC-RAS. This approach aims to mitigate the limitations of HEC–RAS modeling, which relies heavily on precise cross-section and riverbank definitions, and the scarcity of ground truth data.

Our study involved training the model on eight sub-basins and testing on five others. The Res–UNet model, when trained over several years, demonstrated superior performance with sensitivity rates of 67.36% and 95.67% AUC, outperforming HEC–RAS simulations.

Conversely, the machine-learning model (RF) could not be generalized across other test sub-basins, highlighting its limitations.

Additionally, employing the "Sigmoid Focal Loss" function in training the W–Res–U–Net model increased its sensitivity by 0.27% and AUC by 0.17% compared to HEC–RAS simulations. To enhance realism, our model incorporated a dual-loss function that accounted for physical simulations by HEC-RAS and ground truth data. This hybridization resulted in the W–Res–U–Net model achieving a sensitivity of 71.16%, specificity of 91.14%, and AUC of 92.95% when validated against physical simulations, and a sensitivity of 88.89%, specificity of 93.07%, and AUC of 95.87% when validated against ground truth points.

The hybrid approach demonstrated that the dual-loss function significantly improves model performance compared to training solely on physical models or ground truth data. This is particularly beneficial due to the limited availability of sufficient real-world data and the constraints of HEC–RAS simulations.

Moving forward, we aim to enhance the performance of our multi-hybridization approach by incorporating diverse targets and extending the study to various global regions to better generalize our findings.

Author Contributions: A.R.: Conceptualization, data curation, methodology, software and programming, writing—original draft. A.D.: Supervision, review, and editing. M.G.: Conceptualization, data curation, methodology, software and programming, Writing—original draft, Review and editing. T.G.: Data curation. T.B.: Review and editing. B.S.: Conceptualization, supervision, review, and editing. F.M.: Conceptualization, methodology, supervision, review, and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Statement: The dataset and the code for the models used in this study are openly available at the following link: https://disi.unitn.it/~melgani/datasets.html (accessed on 24 September 2024). This will allow for full reproducibility of the research and facilitate further studies.

Acknowledgments: This research was conducted with the support of the University of Sciences and Technology Houari Boumediene, Algeria, and the University of Trento, Italy. The authors extend their gratitude to the Algerian National Water Resources Agency for providing access to rainfall and river flow data. We would also like to thank the Algerian Civil Protection Services for supplying data on flooded points in the study area. Additionally, we would like to thank the National Delegation for Major Risk Management for encouraging and supporting this research.

Conflicts of Interest: The authors state that they have no known financial interests or personal relationships that could have influenced the work reported in this paper.

References

1. Chapi, K.; Singh, V.P.; Shirzadi, A.; Shahabi, H.; Bui, D.T.; Pham, B.T.; Khosravi, K. A Novel Hybrid Artificial Intelligence Approach for Flood Susceptibility Assessment. *Environ. Model. Softw.* **2017**, *95*, 229–245. [CrossRef]
2. Chen, W.; Li, Y.; Xue, W.; Shahabi, H.; Li, S.; Hong, H.; Wang, X.; Bian, H.; Zhang, S.; Pradhan, B.; et al. Modeling Flood Susceptibility Using Data-Driven Approaches of Naïve Bayes Tree, Alternating Decision Tree, and Random Forest Methods. *Sci. Total Environ.* **2020**, *701*, 134979. [CrossRef]
3. Fang, L.; Huang, J.; Cai, J.; Nitivattananon, V. Hybrid Approach for Flood Susceptibility Assessment in a Flood-Prone Mountainous Catchment in China. *J. Hydrol.* **2022**, *612*, 128091. [CrossRef]
4. Khalil, U.; Khan, N.M. Floodplain Mapping for Indus River: Chashma—Taunsa Reach. *Pak. J. Eng. Appl. Sci.* **2017**, *20*, 30–48.
5. Gharakhanlou, N.M.; Perez, L. Flood Susceptible Prediction through the Use of Geospatial Variables and Machine Learning Methods. *J. Hydrol.* **2023**, *617*, 129121. [CrossRef]
6. Ntanganedzeni, B.; Nobert, J. Flood Risk Assessment in Luvuvhu River, Limpopo Province, South Africa. *Phys. Chem. Earth Parts ABC* **2021**, *124*, 102959. [CrossRef]
7. Cao, H.; Zhang, H.; Wang, C.; Zhang, B. Operational Flood Detection Using Sentinel-1 SAR Data over Large Areas. *Water* **2019**, *11*, 786. [CrossRef]
8. Liang, J.; Liu, D. A Local Thresholding Approach to Flood Water Delineation Using Sentinel-1 SAR Imagery. *ISPRS J. Photogramm. Remote Sens.* **2020**, *159*, 53–62. [CrossRef]

9. Nhangumbe, M.; Nascetti, A.; Ban, Y. Multi-Temporal Sentinel-1 SAR and Sentinel-2 MSI Data for Flood Mapping and Damage Assessment in Mozambique. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 53. [CrossRef]
10. Tamiru Haile, A.; Worku Bekele, T.; Rientjes, T. Interannual Comparison of Historical Floods through Flood Detection Using Multi-Temporal Sentinel-1 SAR Images, Awash River Basin, Ethiopia. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *124*, 103505. [CrossRef]
11. Notti, D.; Giordan, D.; Caló, F.; Pepe, A.; Zucca, F.; Galve, J. Potential and Limitations of Open Satellite Data for Flood Mapping. *Remote Sens.* **2018**, *10*, 1673. [CrossRef]
12. Shastry, A.; Carter, E.; Coltin, B.; Sleeter, R.; McMichael, S.; Eggleston, J. Mapping Floods from Remote Sensing Data and Quantifying the Effects of Surface Obstruction by Clouds and Vegetation. *Remote Sens. Environ.* **2023**, *291*, 113556. [CrossRef]
13. Marzukhi, S.; Sidik, M.A.S.M.; Nasir, H.M.; Zainol, Z.; Ismail, M.N. Flood Detection and Warning System (FLoWS). In Proceedings of the 12th International Conference on Ubiquitous Information Management and Communication, Langkawi, Malaysia, 5–7 January 2018; pp. 1–4.
14. Bentivoglio, R.; Isufi, E.; Jonkman, S.N.; Taormina, R. Deep Learning Methods for Flood Mapping: A Review of Existing Applications and Future Research Directions. *Hydrol. Earth Syst. Sci.* **2022**, *26*, 4345–4378. [CrossRef]
15. Renschler, C.S.; Wang, Z. Multi-Source Data Fusion and Modeling to Assess and Communicate Complex Flood Dynamics to Support Decision-Making for Downstream Areas of Dams: The 2011 Hurricane Irene and Schoharie Creek Floods, NY. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *62*, 157–173. [CrossRef]
16. Rajib, A.; Liu, Z.; Merwade, V.; Tavakoly, A.A.; Follum, M.L. Towards a Large-Scale Locally Relevant Flood Inundation Modeling Framework Using SWAT and LISFLOOD-FP. *J. Hydrol.* **2020**, *581*, 124406. [CrossRef]
17. Seenu, P.Z.; Venkata Rathnam, E.; Jayakumar, K.V. Visualisation of Urban Flood Inundation Using SWMM and 4D GIS. *Spat. Inf. Res.* **2020**, *28*, 459–467. [CrossRef]
18. Afshari, S.; Tavakoly, A.A.; Rajib, M.A.; Zheng, X.; Follum, M.L.; Omranian, E.; Fekete, B.M. Comparison of New Generation Low-Complexity Flood Inundation Mapping Tools with a Hydrodynamic Model. *J. Hydrol.* **2018**, *556*, 539–556. [CrossRef]
19. Hammami, S.; Zouhri, L.; Souissi, D.; Souei, A.; Zghibi, A.; Marzougui, A.; Dlala, M. Application of the GIS Based Multi-Criteria Decision Analysis and Analytical Hierarchy Process (AHP) in the Flood Susceptibility Mapping (Tunisia). *Arab. J. Geosci.* **2019**, *12*, 653. [CrossRef]
20. Shadmehri Toosi, A.; Calbimonte, G.H.; Nouri, H.; Alaghmand, S. River Basin-Scale Flood Hazard Assessment Using a Modified Multi-Criteria Decision Analysis Approach: A Case Study. *J. Hydrol.* **2019**, *574*, 660–671. [CrossRef]
21. Msabi, M.M.; Makonyo, M. Flood Susceptibility Mapping Using GIS and Multi-Criteria Decision Analysis: A Case of Dodoma Region, Central Tanzania. *Remote Sens. Appl. Soc. Environ.* **2021**, *21*, 100445. [CrossRef]
22. Shahiri Tabarestani, E.; Afzalimehr, H. A Comparative Assessment of Multi-Criteria Decision Analysis for Flood Susceptibility Modelling. *Geocarto Int.* **2022**, *37*, 5851–5874. [CrossRef]
23. Tehrany, M.S.; Lee, M.-J.; Pradhan, B.; Jebur, M.N.; Lee, S. Flood Susceptibility Mapping Using Integrated Bivariate and Multivariate Statistical Models. *Environ. Earth Sci.* **2014**, *72*, 4001–4015. [CrossRef]
24. Youssef, A.M.; Pradhan, B.; Sefry, S.A. Flash Flood Susceptibility Assessment in Jeddah City (Kingdom of Saudi Arabia) Using Bivariate and Multivariate Statistical Models. *Environ. Earth Sci.* **2016**, *75*, 12. [CrossRef]
25. Khosravi, K.; Pourghasemi, H.R.; Chapi, K.; Bahri, M. Flash Flood Susceptibility Analysis and Its Mapping Using Different Bivariate Models in Iran: A Comparison between Shannon's Entropy, Statistical Index, and Weighting Factor Models. *Environ. Monit. Assess.* **2016**, *188*, 656. [CrossRef] [PubMed]
26. Costache, R. Flood Susceptibility Assessment by Using Bivariate Statistics and Machine Learning Models—A Useful Tool for Flood Risk Management. *Water Resour. Manag.* **2019**, *33*, 3239–3256. [CrossRef]
27. Ali, S.A.; Parvin, F.; Pham, Q.B.; Vojtek, M.; Vojteková, J.; Costache, R.; Linh, N.T.T.; Nguyen, H.Q.; Ahmad, A.; Ghorbani, M.A. GIS-Based Comparative Assessment of Flood Susceptibility Mapping Using Hybrid Multi-Criteria Decision-Making Approach, Naïve Bayes Tree, Bivariate Statistics and Logistic Regression: A Case of Topl'a Basin, Slovakia. *Ecol. Indic.* **2020**, *117*, 106620. [CrossRef]
28. Hamidi, E.; Peter, B.G.; Munoz, D.F.; Moftakhari, H.; Moradkhani, H. Fast Flood Extent Monitoring With SAR Change Detection Using Google Earth Engine. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 4201419. [CrossRef]
29. Güvel, Ş.P.; Akgül, M.A.; Aksu, H. Flood Inundation Maps Using Sentinel-2: A Case Study in Berdan Plain. *Water Supply* **2022**, *22*, 4098–4108. [CrossRef]
30. Psomiadis, E.; Diakakis, M.; Soulis, K.X. Combining SAR and Optical Earth Observation with Hydraulic Simulation for Flood Mapping and Impact Assessment. *Remote Sens.* **2020**, *12*, 3980. [CrossRef]
31. Tavus, B.; Kocaman, S.; Nefeslioglu, H.A.; Gokceoglu, C. A Fusion Approach for Flood Mapping Using Sentinel-1 and Sentinel-2 Datasets. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *XLIII-B3-2020*, 641–648. [CrossRef]
32. Khosravi, K.; Shahabi, H.; Pham, B.T.; Adamowski, J.; Shirzadi, A.; Pradhan, B.; Dou, J.; Ly, H.-B.; Gróf, G.; Ho, H.L.; et al. A Comparative Assessment of Flood Susceptibility Modeling Using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. *J. Hydrol.* **2019**, *573*, 311–323. [CrossRef]
33. Gudiyangada Nachappa, T.; Tavakkoli Piralilou, S.; Gholamnia, K.; Ghorbanzadeh, O.; Rahmati, O.; Blaschke, T. Flood Susceptibility Mapping with Machine Learning, Multi-Criteria Decision Analysis and Ensemble Using Dempster Shafer Theory. *J. Hydrol.* **2020**, *590*, 125275. [CrossRef]

34. Towfiqul Islam, A.R.M.; Talukdar, S.; Mahato, S.; Kundu, S.; Eibek, K.U.; Pham, Q.B.; Kuriqi, A.; Linh, N.T.T. Flood Susceptibility Modelling Using Advanced Ensemble Machine Learning Models. *Geosci. Front.* **2021**, *12*, 101075. [CrossRef]
35. Madhuri, R.; Sistla, S.; Srinivasa Raju, K. Application of Machine Learning Algorithms for Flood Susceptibility Assessment and Risk Management. *J. Water Clim. Chang.* **2021**, *12*, 2608–2623. [CrossRef]
36. Prasad, P.; Loveson, V.J.; Das, B.; Kotha, M. Novel Ensemble Machine Learning Models in Flood Susceptibility Mapping. *Geocarto Int.* **2022**, *37*, 4571–4593. [CrossRef]
37. Seydi, S.T.; Kanani-Sadat, Y.; Hasanlou, M.; Sahraei, R.; Chanussot, J.; Amani, M. Comparison of Machine Learning Algorithms for Flood Susceptibility Mapping. *Remote Sens.* **2022**, *15*, 192. [CrossRef]
38. Youssef, A.M. Optimal Flood Susceptibility Model Based on Performance Comparisons of LR, EGB, and RF Algorithms. *Nat. Hazards* **2023**, *115*, 1071–1096. [CrossRef]
39. Saber, M.; Boulmaiz, T.; Guermoui, M.; Abdrabo, K.I.; Kantoush, S.A.; Sumi, T.; Boutaghane, H.; Nohara, D.; Mabrouk, E. Examining LightGBM and CatBoost Models for Wadi Flash Flood Susceptibility Prediction. *Geocarto Int.* **2022**, *37*, 7462–7487. [CrossRef]
40. Saravanan, S.; Abijith, D.; Reddy, N.M.; Kss, P.; Janardhanam, N.; Sathiyamurthi, S.; Sivakumar, V. Flood Susceptibility Mapping Using Machine Learning Boosting Algorithms Techniques in Idukki District of Kerala India. *Urban Clim.* **2023**, *49*, 101503. [CrossRef]
41. Tien Bui, D.; Hoang, N.-D.; Martínez-Álvarez, F.; Ngo, P.-T.T.; Hoa, P.V.; Pham, T.D.; Samui, P.; Costache, R. A Novel Deep Learning Neural Network Approach for Predicting Flash Flood Susceptibility: A Case Study at a High Frequency Tropical Storm Area. *Sci. Total Environ.* **2020**, *701*, 134413. [CrossRef]
42. Bui, Q.-T.; Nguyen, Q.-H.; Nguyen, X.L.; Pham, V.D.; Nguyen, H.D.; Pham, V.-M. Verification of Novel Integrations of Swarm Intelligence Algorithms into Deep Learning Neural Network for Flood Susceptibility Mapping. *J. Hydrol.* **2020**, *581*, 124379. [CrossRef]
43. Ahmadlou, M.; Al-Fugara, A.; Al-Shabeeb, A.R.; Arora, A.; Al-Adamat, R.; Pham, Q.B.; Al-Ansari, N.; Linh, N.T.T.; Sajedi, H. Flood Susceptibility Mapping and Assessment Using a Novel Deep Learning Model Combining Multilayer Perceptron and Autoencoder Neural Networks. *J. Flood Risk Manag.* **2021**, *14*, e12683. [CrossRef]
44. Shahabi, H.; Shirzadi, A.; Ronoud, S.; Asadi, S.; Pham, B.T.; Mansouripour, F.; Geertsema, M.; Clague, J.J.; Bui, D.T. Flash Flood Susceptibility Mapping Using a Novel Deep Learning Model Based on Deep Belief Network, Back Propagation and Genetic Algorithm. *Geosci. Front.* **2021**, *12*, 101100. [CrossRef]
45. Costache, R.; Arabameri, A.; Costache, I.; Crăciun, A.; Md Towfiqul Islam, A.R.; Abba, S.I.; Sahana, M.; Pham, B.T. Flood Susceptibility Evaluation through Deep Learning Optimizer Ensembles and GIS Techniques. *J. Environ. Manag.* **2022**, *316*, 115316. [CrossRef]
46. Saha, S.; Gayen, A.; Bayen, B. Deep Learning Algorithms to Develop Flood Susceptibility Map in Data-Scarce and Ungauged River Basin in India. *Stoch. Environ. Res. Risk Assess.* **2022**, *36*, 3295–3310. [CrossRef]
47. Ramayanti, S.; Nur, A.S.; Syifa, M.; Panahi, M.; Achmad, A.R.; Park, S.; Lee, C.-W. Performance Comparison of Two Deep Learning Models for Flood Susceptibility Map in Beira Area, Mozambique. *Egypt. J. Remote Sens. Space Sci.* **2022**, *25*, 1025–1036. [CrossRef]
48. Li, Y.; Hong, H. Modelling Flood Susceptibility Based on Deep Learning Coupling with Ensemble Learning Models. *J. Environ. Manag.* **2023**, *325*, 116450. [CrossRef]
49. Mia, M.U.; Chowdhury, T.N.; Chakrabortty, R.; Pal, S.C.; Al-Sadoon, M.K.; Costache, R.; Islam, A.R.M.T. Flood Susceptibility Modeling Using an Advanced Deep Learning-Based Iterative Classifier Optimizer. *Land* **2023**, *12*, 810. [CrossRef]
50. Ouma, Y.O.; Omai, L. Flood Susceptibility Mapping Using Image-Based 2D-CNN Deep Learning: Overview and Case Study Application Using Multiparametric Spatial Data in Data-Scarce Urban Environments. *Int. J. Intell. Syst.* **2023**, *2023*, 5672401. [CrossRef]
51. Riche, A.; Drias, A.; Ricci, R.; Souissi, B.; Melgani, F. Predicting LULC Changes and Assessing Their Impact on Surface Runoff with Machine Learning and Remote Sensing Data. *Res. Sq.* **2023**. [CrossRef]
52. US Army Corps of Engineers Institute of Water Resources Hydrologic Engineering Center. *CEIWR-HEC Hydrologic Modeling System HEC-HMS Quick Start Guide*; US Army Corps of Engineers Institute of Water Resources Hydrologic Engineering Center: Davis, CA, USA, 2016.
53. Yilma, Z.L.; Kebede, H.H. Simulation of the Rainfall–Runoff Relationship Using an HEC-HMS Hydrological Model for Dabus Subbasin, Blue Nile Basin, Ethiopia. *H_2Open J.* **2023**, *6*, 331–342. [CrossRef]
54. Razi MA, M.; Deli MA, Z.; Yusoff MA, M.; Ahmad, M.A.; Adnan, M.S. Flood Modelling Studies Using River Analysis System (HEC-RAS) For Flood Plain Area in Muar City. *Int. J. Integr. Eng.* **2022**, *14*, 38–47. [CrossRef]
55. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
56. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2015; Volume 9351, pp. 234–241. ISBN 978-3-319-24573-7.
57. Ohyama, W.; Suzuki, M.; Uchida, S. Detecting Mathematical Expressions in Scientific Document Images Using a U-Net Trained on a Diverse Dataset. *IEEE Access* **2019**, *7*, 144030–144042. [CrossRef]

58. Zeng, Z.; Xie, W.; Zhang, Y.; Lu, Y. RIC-Unet: An Improved Neural Network Based on Unet for Nuclei Segmentation in Histology Images. *IEEE Access* **2019**, *7*, 21420–21428. [CrossRef]
59. Wu, X.; Zhang, Z.; Xiong, S.; Zhang, W.; Tang, J.; Li, Z.; An, B.; Li, R. A Near-Real-Time Flood Detection Method Based on Deep Learning and SAR Images. *Remote Sens.* **2023**, *15*, 2046. [CrossRef]
60. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; IEEE: Las Vegas, NV, USA, 2016; pp. 770–778.
61. Khan, R.A.; Luo, Y.; Wu, F.-X. RMS-UNet: Residual Multi-Scale UNet for Liver and Lesion Segmentation. *Artif. Intell. Med.* **2022**, *124*, 102231. [CrossRef]
62. Askar, S.; Zeraat Peyma, S.; Yousef, M.M.; Prodanova, N.A.; Muda, I.; Elsahabi, M.; Hatamiafkoueieh, J. Flood Susceptibility Mapping Using Remote Sensing and Integration of Decision Table Classifier and Metaheuristic Algorithms. *Water* **2022**, *14*, 3062. [CrossRef]
63. Samanta, S.; Pal, D.K.; Palsamanta, B. Flood Susceptibility Analysis through Remote Sensing, GIS and Frequency Ratio Model. *Appl. Water Sci.* **2018**, *8*, 66. [CrossRef]
64. Amiri, A.; Soltani, K.; Ebtehaj, I.; Bonakdari, H. A Novel Machine Learning Tool for Current and Future Flood Susceptibility Mapping by Integrating Remote Sensing and Geographic Information Systems. *J. Hydrol.* **2024**, *632*, 130936. [CrossRef]
65. Liu, J.; Liu, K.; Wang, M. A Residual Neural Network Integrated with a Hydrological Model for Global Flood Susceptibility Mapping Based on Remote Sensing Datasets. *Remote Sens.* **2023**, *15*, 2447. [CrossRef]
66. Burayu, D.G.; Karuppannan, S.; Shuniye, G. Identifying Flood Vulnerable and Risk Areas Using the Integration of Analytical Hierarchy Process (AHP), GIS, and Remote Sensing: A Case Study of Southern Oromia Region. *Urban Clim.* **2023**, *51*, 101640. [CrossRef]
67. Beven, K.J.; Kirkby, M.J. A Physically Based, Variable Contributing Area Model of Basin Hydrology/Un Modèle à Base Physique de Zone d'appel Variable de l'hydrologie Du Bassin Versant. *Hydrol. Sci. Bull.* **1979**, *24*, 43–69. [CrossRef]
68. Werner, M.G.F.; Hunter, N.M.; Bates, P.D. Identifiability of Distributed Floodplain Roughness Values in Flood Extent Estimation. *J. Hydrol.* **2005**, *314*, 139–157. [CrossRef]
69. Wu, X.; Meng, Z.; Dang, X.; Wang, J. Effects of Rock Fragments on the Water Infiltration and Hydraulic Conductivity in the Soils of the Desert Steppes of Inner Mongolia, China. *Soil Water Res.* **2021**, *16*, 151–163. [CrossRef]
70. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Deering. Monitoring Vegetation Systems in the Great Plains with ERTS. *NASA Spec. Publ.* **1974**, *351*, 309.
71. Zha, Y.; Gao, J.; Ni, S. Use of Normalized Difference Built-up Index in Automatically Mapping Urban Areas from TM Imagery. *Int. J. Remote Sens.* **2003**, *24*, 583–594. [CrossRef]
72. Xu, H. Modification of Normalised Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033. [CrossRef]
73. Moore, I.D.; Grayson, R.B. Terrain-based Catchment Partitioning and Runoff Prediction Using Vector Elevation Data. *Water Resour. Res.* **1991**, *27*, 1177–1191. [CrossRef]
74. Billi, P. Flash Flood Sediment Transport in a Steep Sand-Bed Ephemeral Stream. *Int. J. Sediment Res.* **2011**, *26*, 193–209. [CrossRef]
75. Fang, Z.; Wang, Y.; Peng, L.; Hong, H. Predicting Flood Susceptibility Using LSTM Neural Networks. *J. Hydrol.* **2021**, *594*, 125734. [CrossRef]