


Enhancing Flood Resilience: Streamflow Forecasting and Inundation Modeling in Pakistan [†]

Maham Shehzadi ^{1,‡}, Raja Hashim Ali ^{1,2,*} , Zain ul Abideen ³, Ali Zeeshan Ijaz ¹ and Talha Ali Khan ^{2,*}

¹ AI Research Group, Faculty of Computer Science and Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi 23460, Khyber Pakhtunkhwa, Pakistan; u2021273@giki.edu.pk (M.S.); ali.zeeshan@giki.edu.pk (A.Z.I.)

² Department of Technology and Software Engineering, University of Europe for Applied Sciences, 14469 Berlin, Germany

³ BCMaterials Basque Center for Materials, Applications & Nanostructure, Buil. Martina Casiano, Pl. 3 Parque Científico UPV/EHU Barrio Sarriena, 48940 Leioa, Spain; zainul.abideen@bcmaterials.net

* Correspondence: hashim.ali@giki.edu.pk (R.H.A.); talhaali.khan@ue-germany.de (T.A.K.)

[†] Presented at the 4th International Electronic Conference on Applied Sciences, 27 October–10 November 2023; Available online: <https://asec2023.sciforum.net/>.

[‡] These authors contributed equally to this work.

Abstract: Climatic changes have increased the frequency of natural disasters, and Pakistan, as a developing nation, is facing severe challenges in coping with floods, which have devastatingly impacted people's livelihoods. In 2022, floods affected over 33 million people, resulting in more than 1730 deaths, according to the World Bank. Flood prediction is a critical research area which can aid in saving critical lives, crops, livestock, and money. This study employs machine learning techniques to provide accurate and reliable flood forecasts for Pakistan. Specifically, Support Vector Machine (SVM) and Artificial neural network (ANN) are utilized in this research for flood prediction. Historical data encompassing floods, rainfall, temperature, water level, topographic information, and land cover of Pakistan is collected and split into 75% for model training and 25% for testing. Additionally, topographic data and land cover information are employed to create inundation maps. The findings highlight three topographic factors that play a pivotal role in predicting flood-sensitive areas: slope, distance to the river, and river. The combined Support Vector Machine (SVM) and Artificial neural network (ANN) exhibited areas under the curve values of 0.94 and 0.95 for the training and testing phases, respectively. These results demonstrate the efficacy of the SVM and ANN integration for precise flood forecasting in Pakistan, contributing to enhancing flood resilience in the region.

Keywords: Support Vector Machine (SVM); artificial neural network (ANN); machine learning; flood; mapping; meteorology; topography



Citation: Shehzadi, M.; Ali, R.H.; Abideen, Z.u.; Ijaz, A.Z.; Khan, T.A. Enhancing Flood Resilience: Streamflow Forecasting and Inundation Modeling in Pakistan. *Eng. Proc.* **2023**, *56*, 315. <https://doi.org/10.3390/ASEC2023-16612>

Academic Editor: Nunzio Cennamo

Published: 7 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Flood events are a pervasive global occurrence, yet their severity exhibits regional variations. Particularly in developing nations, the recurrence of floods exacts a significant toll in terms of human lives and generates severe economic turmoil, amplifying financial strains. The amplification of global temperatures, attributable to the overarching phenomenon of climate change, has led to an accelerated rate of snowmelt and heightened precipitation. Consequently, the incidence of floods has surged both in frequency and intensity. This trend is evident in Figure 1, showcasing Pakistan's elevated flood occurrence rate compared to other natural disasters [1]. The year 2021 saw floods surpassing all other calamities in South Asian countries [1], underlining the prominence of flooding as a critical issue in the region as shown in Figure 1. Annually, a staggering average of USD 300 billion worth of damages and the consequential societal repercussions have spurred researchers around the globe to earnestly address the pressing issue of flooding [2].

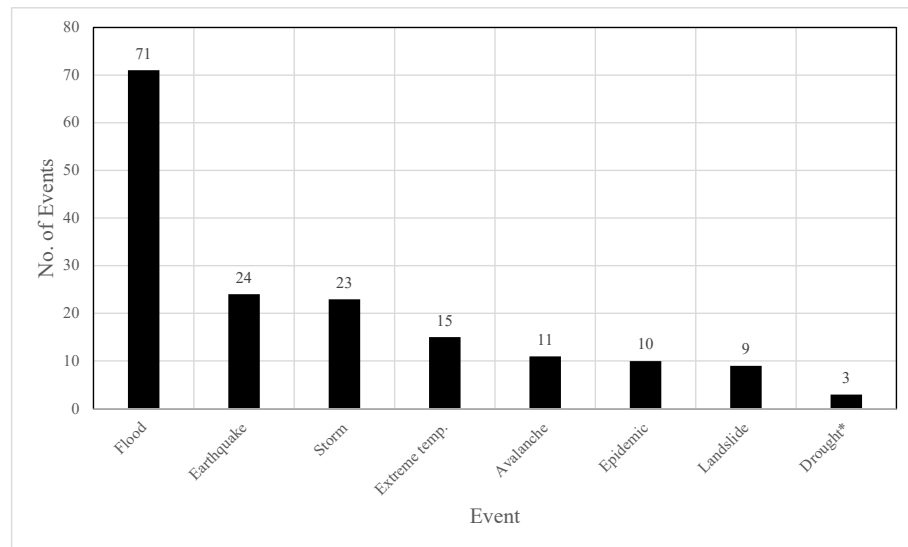


Figure 1. Number of natural disaster events in Pakistan since 1900. Data source: EM-DAT: The OFDA/CRED International Disaster Data base and United Nation Office for the Coordination of Humanitarian Affairs (OCHA) (*except data of drought).

A multitude of methodologies, models, and strategies with potential regional or global relevance have been proposed; however, effectively modeling and predicting this natural disaster has proven to be a formidable challenge [3]. For Pakistan, which has experienced a higher frequency of floods in recent years, it is evident that there is a critical need for an advanced system to detect floods in advance and accurately map flood-prone areas. Artificial Intelligence [4,5], in particular machine learning (clustering [6,7] and classification) and deep learning [8,9], have played an important role in identification of diseases [10–12], flood detection and warning systems, and in many other allied fields [13–15]. Such a system enables authorities to proactively relocate people to safer locations, reducing the impact of floods on vulnerable communities.

1.1. Literature Review

Numerous statistical approaches and methodologies, such as the climatology average method (CLIM) [16], flood frequency analysis (FFA) [17], Bayesian forecasting models (BFM) [18], and artificial neural networks (ANN) [19], among others, have been extensively employed for flood prediction within scholarly literature. In these frameworks, the intricate physical processes contributing to flooding have been encapsulated through intricate mathematical formulations. The major recent contributions in the field of flood predictions are shown in Table 1.

The significant progress in prediction systems owes much to the utilization of machine learning (ML) methods, which have offered improved performance and cost-effective solutions.

Table 1. Literature review table showing the contributions of various authors for Flood Forecasting.

Papers	Datasets	Year Published	Deep Learning Used	GANS	SVM	ANN	Machine Learning Used	Topographic Factors	Hydrological Factors	Land Cover Factor	Geological Factor	Flood Extent Mapping
Weng et al. [2]	.	July, 2023	yes	yes	yes	.	.	.
do Lago et al. [20]	.	March, 2023	yes	yes
Saikh et al. [21]	Geospatial data and Flood Inventory Map	May, 2023	yes
Saber et al. [22]		May, 2023	yes
Farooq et al. [1]	.	March, 2023	yes
Proposed Solution	Data of Pakistan	July, 2023	.	.	yes	yes	yes	yes	yes	yes	yes	yes

1.2. Our Objectives

The core objective of this research article centers on introducing a novel flood forecasting and river inundation mapping system in Pakistan, filling a void in the existing landscape. This endeavor involves the integration of diverse elements, encompassing rainfall and temperature data, topographic influences, and geological considerations. Through the amalgamation of these components, the research endeavors to construct an adept flood forecasting and inundation mapping mechanism tailored to Pakistan's context. The overarching aim is to elevate the accuracy of flood predictions within the region and fortify readiness strategies for potential flood occurrences.

1.3. Research Questions

This study focuses on enhancing flood resilience in Pakistan through streamflow forecasting and inundation model. Following are the few main research questions discussed in this study.

1. How can streamflow forecasting and inundation modeling be effectively integrated to enhance flood resilience in Pakistan?
2. What are the key factors influencing accurate streamflow forecasting in Pakistan's unique hydrological conditions?
3. How can advanced inundation modeling techniques be utilized to develop efficient flood risk mapping and response strategies in Pakistan?

1.4. Novelty of This Study

This research introduces a novel approach to flood prediction and extent mapping in Pakistan. It combines Support Vector Machine (SVM) and Artificial Neural Network (ANN) models, supported by comprehensive analysis of influencing factors. The study adopts a multidimensional method considering meteorological, topographic, hydrological, land cover, and geological elements, in contrast to conventional techniques.

1.5. Executive Summary

This study aims to enhance flood forecasting and mapping in Pakistan using machine learning, particularly Support Vector Machine (SVM) and Artificial Neural Network (ANN). SVM achieved 94% accuracy and ANN reached 95%, highlighting their reliability. By employing these techniques, the research seeks to improve flood prediction precision, bolstering flood resilience and disaster readiness in the area.

2. Materials and Methods

2.1. Dataset

Datasets of different features such as rainfall, temperature, landsat images, historical flood, water level, weather data, land cover of Pakistan is collected from different sources. An overview of the datasets used in this study is shown in Table 2.

Table 2. Datasets used in this study and their sources

Rainfall (1901–2016)	Kaggle
Temperature (1901–2016)	Kaggle
Water Level (2022–2023)	WAPDA river flow
Historical flood data	Google, Different sources
Land cover data	Land use atlas of Pakistan, Env. Ministry Gov. of Pakistan
Weather data	https://weatherspark.com/ (accessed on 24 June 2023)

2.2. Overall Workflow

By using machine learning approaches, particularly Support Vector Machine (SVM) and Artificial Neural Networks (ANN), to predict floods based on temperature and rainfall data, the current methodology aims to tackle this challenging job. The overall workflow deployed in this study is shown in Figure 2.

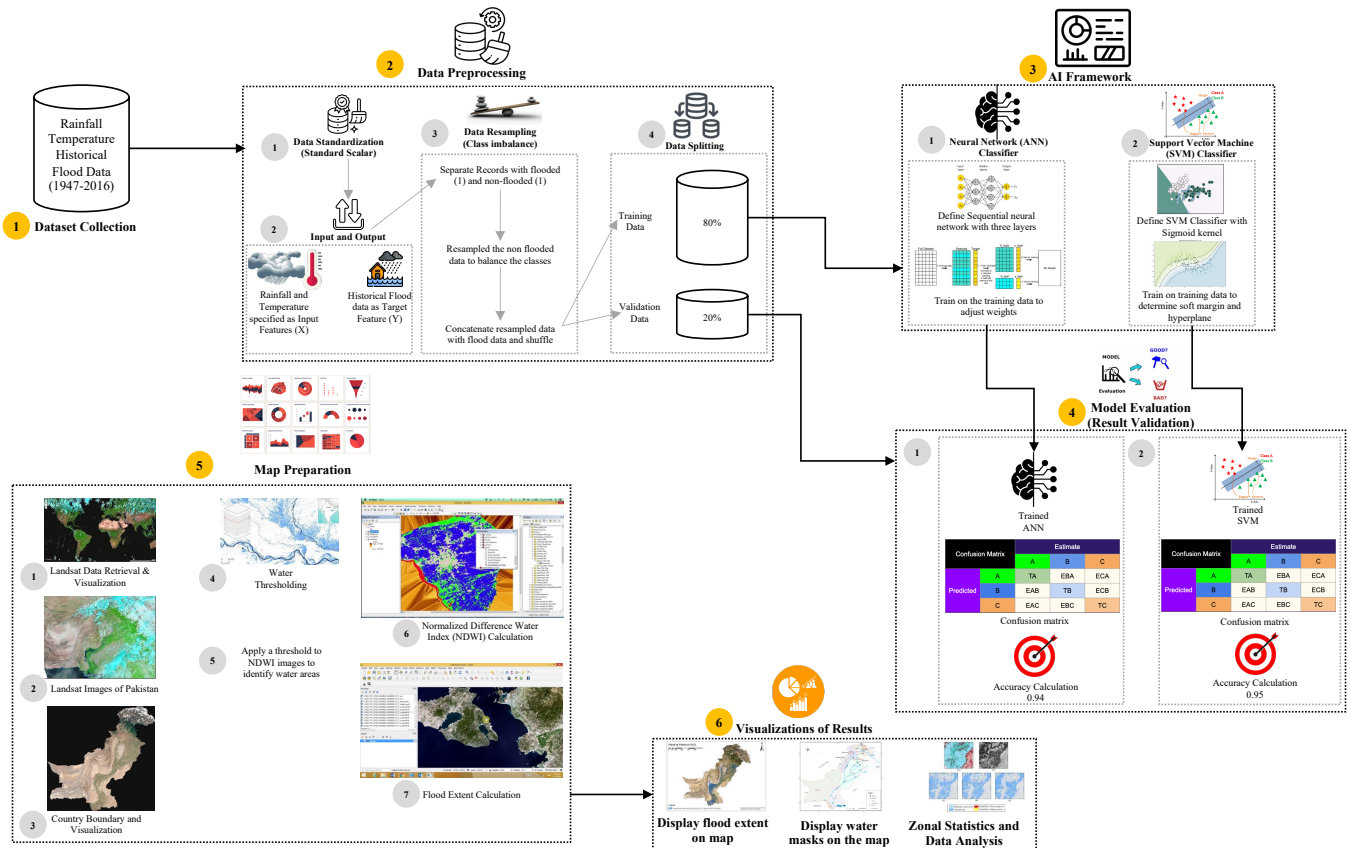


Figure 2. Workflow diagram deployed in the current study.

2.2.1. Data Preparation

The methodology's first step entails gathering and preparing pertinent meteorological data. Using the Pandas library, historical flood records, rainfall, and temperature data are combined into a comprehensive dataset. To help predictive modelling, further data preparation converts categorical flood labels into binary values.

2.2.2. Data Division into Training and Validation Data Sets

The dataset is divided into training and testing sets using scikit-learn's train-test-split function to enable efficient model training and evaluation. Additionally, feature values are standardized using the StandardScaler during model training to guarantee consistent scaling and convergence.

2.2.3. Data Balancing Techniques

Given that floods are very infrequent, the class imbalance in the dataset presents a key issue in flood prediction. The Synthetic Minority Over-sampling Technique (SMOTE), which artificially augments the minority class to provide a more balanced dataset, is used to remedy problem.

2.2.4. Modeling Techniques Deployed in this Study

For this study, SVM and ANN models were deployed. A SVM model with a Sigmoid kernel was trained on the balanced dataset using the scikit-learn module. An ANN

model is constructed using TensorFlow and Keras. The architecture consists of sequential dense layers, with ReLU activation functions, culminating in Sigmoid output, with binary cross-entropy loss and the Adam optimizer. The Neural Network's predictive efficacy is scrutinized in a manner mirroring SVM evaluation.

For the comprehensive evaluation of performance of both models, a suite of metrics is employed, including confusion matrix analysis, classification reports, and accuracy scores.

Upon completion of flood forecasting, geospatial analysis techniques utilizing Google Earth Engine are implemented, involving visualization of satellite imagery, calculation of the Normalized Difference Water Index (NDWI), thresholding for water identification, and determination of flood extent. The methodology culminates in zonal statistics for flood-affected areas within country boundaries.

3. Results and Discussion

In this study, we forecasted floods and mapped river inundation in Pakistan using a variety of variables, including rainfall, temperature, topography, and geology. The outcomes show how well the algorithm was able to map potential inundation zones and anticipate flood events. Results show that on the basis of different factors, we obtained an accuracy of 94% for SVM and 95% for ANN. Figure 3 shows the flood extent in 2021 and in 2022 in Pakistan.



Figure 3. Flood Extent of 2021 and 2022.

The results of river inundation mapping were similarly encouraging. The inundation mapping model accurately identified possible regions at danger of flooding by integrating topographic features and geology. SVM and ANN displayed 94% and 95% overall accuracy rates, respectively, for identifying flood-prone areas. The model successfully identified and mapped inundation zones during flood events, with a high precision of 95% and a recall rate of 94%.

Seasonal precipitation variability and land use changes were found to be two important variables that affect Pakistani streamflow forecasting accuracy. The AI algorithms (SVM, ANN) employed in flood prediction and classification operate on established assumptions and are trained using accessible data, potentially lacking a comprehensive grasp of the intricate elements shaping flash floods [23].

The integration of streamflow forecasting and inundation modeling in our study achieved a high level of accuracy, enhancing flood resilience in Pakistan by providing rapid warnings and precise flood extent mapping through real-time streamflow data and advanced modeling techniques.

Seasonal precipitation variability and land use changes were identified as pivotal factors affecting the accuracy of streamflow forecasting in Pakistan. Understanding and incorporating these variables are critical for improving the precision of flood predictions.

4. Conclusions

In conclusion, this conference paper shows significant progress in improving Pakistan's flood resilience through the integration of streamflow forecasting and flood modeling. Using advanced machine learning algorithms such as Support Vector Machine (SVM) and Artificial Neural Network (ANN) results in impressive 94% and 95% accuracy rates. The study addresses the key factors influencing accurate flow forecasting and explores the application of advanced flood modeling to flood risk mapping and response strategies. The results promise to improve flood preparedness and disaster management in the region and provide valuable information for future research to further improve flood resistance and effectively combat the challenges of a changing climate.

Author Contributions: Conceptualization, R.H.A.; Methodology, M.S.; Software, M.S.; validation, Z.U.A.; formal analysis, T.A.K.; writing—original draft preparation, M.S., A.Z.I. and R.H.A.; writing—review and editing, R.H.A. and T.A.K.; supervision, R.H.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data repositories used in the project are linked.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Farooq, M.S.; Tehseen, R.; Qureshi, J.N.; Omer, U.; Yaqoob, R.; Tanweer, H.A.; Atal, Z. FFM: Flood forecasting model using federated learning. *IEEE Access* **2023**, *11*, 24472–24483. [\[CrossRef\]](#)
2. Weng, P.; Tian, Y.; Liu, Y.; Zheng, Y. Time-series generative adversarial networks for flood forecasting. *J. Hydrol.* **2023**, *622*, 129702. [\[CrossRef\]](#)
3. Niknam, S.; Dhillon, H.S.; Reed, J.H. Federated learning for wireless communications: Motivation, opportunities, and challenges. *IEEE Commun. Mag.* **2020**, *58*, 46–51. [\[CrossRef\]](#)
4. Ijaz, A.Z.; Ali, R.H.; Sarwar, A.; Khan, T.A.; Baig, M.M. Importance of Synteny in Homology Inference. In Proceedings of the 2022 17th International Conference on Emerging Technologies (ICET), Swabi, Pakistan, 29–30 November 2022 ; pp. 234–239.
5. Ali, R.H.; Khan, A.A. Tracing the evolution of FERM domain of Kindlins. *Mol. Phylogenetics Evol.* **2014**, *80*, 193–204. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Ali, R.H.; Bogusz, M.; Whelan, S. Identifying clusters of high confidence homologies in multiple sequence alignments. *Mol. Biol. Evol.* **2019**, *36*, 2340–2351. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Ali, R.H.; Muhammad, S.A.; Khan, M.A.; Arvestad, L. Quantitative synteny scoring improves homology inference and partitioning of gene families. In Proceedings of the Eleventh Annual Research in Computational Molecular Biology (RECOMB) Satellite Workshop on Comparative Genomics, Lyon, France, 17–19 October 2013 ; Volume 14, pp. 1–9.
8. Ali, N.; Ijaz, A.Z.; Ali, R.H.; Ul Abideen, Z.; Bais, A. Scene Parsing Using Fully Convolutional Network for Semantic Segmentation. In Proceedings of the 2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Regina, Canada, 24–27 September 2023 ; pp. 180–185.
9. Aslam, M.H.; Hussain, S.F.; Ali, R.H. Predictive analysis on severity of Non-Alcoholic Fatty Liver Disease (NAFLD) using Machine Learning Algorithms. In Proceedings of the 2022 17th International Conference on Emerging Technologies (ICET), Swabi, Pakistan, 29–30 November 2022 ; pp. 95–100.
10. ul Abideen, Z.; Khan, T.A.; Ali, R.H.; Ali, N.; Baig, M.M.; Ali, M.S. DocOnTap: AI-based disease diagnostic system and recommendation system. In Proceedings of the 2022 17th International Conference on Emerging Technologies (ICET), Swabi, Pakistan, 29–30 November 2022 ; pp. 148–153.
11. Ali, N.; Ansari, S.; Halim, Z.; Ali, R.H.; Khan, M.F.; Khan, M. Breast cancer classification and proof of key artificial neural network terminologies. In Proceedings of the 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), Karachi, Pakistan, 14–15 December 2019; pp. 1–6.

12. Iftikhar, M.; Ali, N.; Ali, R.H.; Bais, A. Classification of Parkinson Disease with Feature Selection using Genetic Algorithm. In Proceedings of the 2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Regina, SK, Canada, 24–27 September 2023; pp. 522–527.
13. Abdullah, A.; Ali, N.; Ali, R.H.; Abideen, Z.U.; Ijaz, A.Z.; Bais, A. American Sign Language Character Recognition using Convolutional Neural Networks. In Proceedings of the 2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Regina, SK, Canada, 24–27 September 2023; pp. 165–169.
14. Yang, Y.; Tu, S.; Ali, R.H.; Alasmay, H.; Waqas, M.; Amjad, M.N. Intrusion detection based on bidirectional long short-term memory with attention mechanism. *Comput. Mater. Contin.* **2023**, *74*, 801–815. [[CrossRef](#)]
15. Ijaz, A.Z.; Ali, R.H.; Ali, N.; Laique, T.; Khan, T.A. Solving Graph Coloring Problem via Graph Neural Network (GNN). In Proceedings of the 2022 17th International Conference on Emerging Technologies (ICET), Swabi, Pakistan, 29–30 November 2022; pp. 178–183.
16. Sankaranarayanan, S.; Prabhakar, M.; Satish, S.; Jain, P.; Ramprasad, A.; Krishnan, A. Flood prediction based on weather parameters using deep learning. *J. Water Clim. Chang.* **2020**, *11*, 1766–1783. [[CrossRef](#)]
17. Aziz, K.; Rahman, A.; Fang, G.; Shrestha, S. Application of artificial neural networks in regional flood frequency analysis: a case study for Australia. *Stoch. Environ. Res. Risk Assess.* **2014**, *28*, 541–554. [[CrossRef](#)]
18. Haddad, K.; Rahman, A. Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework—Quantile Regression vs. Parameter Regression Technique. *J. Hydrol.* **2012**, *430*, 142–161. [[CrossRef](#)]
19. Kim, S.; Matsumi, Y.; Pan, S.; Mase, H. A real-time forecast model using artificial neural network for after-runner storm surges on the Tottori coast, Japan. *Ocean. Eng.* **2016**, *122*, 44–53. [[CrossRef](#)]
20. do Lago, C.A.; Giacomoni, M.H.; Bentivoglio, R.; Taormina, R.; Junior, M.N.G.; Mendiondo, E.M. Generalizing rapid flood predictions to unseen urban catchments with conditional generative adversarial networks. *J. Hydrol.* **2023**, *618*, 129276. [[CrossRef](#)]
21. Saikh, N.I.; Mondal, P. Gis-based machine learning algorithm for flood susceptibility analysis in the Pagla river basin, Eastern India. *Nat. Hazards Res.* **2023**, *3*, 420–436. [[CrossRef](#)]
22. Saber, M.; Boulmaiz, T.; Guermoui, M.; Abdrabo, K.I.; Kantoush, S.A.; Sumi, T.; Boutaghane, H.; Hori, T.; Binh, D.V.; Nguyen, B.Q.; Bui, T.T. Enhancing flood risk assessment through integration of ensemble learning approaches and physical-based hydrological modeling. *Geomat. Nat. Hazards Risk* **2023**, *14*, 2203798. [[CrossRef](#)]
23. Nakhaei, M.; Nakhaei, P.; Gheibi, M.; Chahkandi, B.; Wacławek, S.; Behzadian, K.; Chen, A.S.; Campos, L.C. Enhancing community resilience in arid regions: A smart framework for flash flood risk assessment. *Ecol. Indic.* **2023**, *153*, 110457. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.