Machine Learning Applications in

Disaster Forecasting and Management

Fatima Abbas

SCOPUA

Machine Learning Applications in Disaster Forecasting and Management

By

Fatima Abbas

Book Title

Machine Learning Applications in Disaster Forecasting and Management

Author

Ms. Fatima Abbas is a dedicated academic and Instructor of Computer Networks with a strong interest in computer science, networking technologies, and research-based learning. She has been actively involved in teaching and mentoring students, helping them develop both theoretical knowledge and practical skills in the field of computing. Through her commitment to academic excellence and student development, she has contributed significantly to fostering a learning environment that encourages innovation, critical thinking, and professional growth. Her passion for education and continuous learning has inspired many students to pursue research and technological advancement in their respective fields.

Affi: Riphah International University Faisalabad, Pakistan

fatima.abbas@riphahfsd.edu.pk

Publisher

Scientific Collaborative Online Publishing Universal Academy, Lahore, Pakistan

Series Name

SCOPUA Books (ISSN-P: 3106-4213 **& ISSN-E:** 3106-4221**)**

Date of Publication

08 June 2026

Edition

1

ISBN: 978-627-7898-13-7

DOI: 10.64060/978-627-7898-13-7

BISAC

SCI013080, TEC010000, COM100000

THAMA

RN, RB, UX

BIC

RNR, JFFC, KCX

DEDICATION

This book is dedicated to my beloved parents, teachers, and well-wishers whose continuous support, encouragement, and guidance have been a source of inspiration throughout this research journey. Their sacrifices, prayers, and motivation played a significant role in the successful completion of this work. This dedication is also extended to all researchers and professionals striving to develop innovative solutions for disaster management and community resilience.

ACKNOWLEDGEMENTS

All praise and gratitude are due to Almighty Allah, whose blessings, mercy, and guidance made the completion of this work possible.

I would like to express my sincere gratitude to my supervisor, teachers, and mentors for their valuable guidance, constructive suggestions, and continuous encouragement throughout this research. Their expertise and support greatly contributed to the successful completion of this study. This book grew out of my Master's thesis.

I am also thankful to my family and friends for their unwavering support, patience, and motivation during every stage of this work. Their encouragement provided the strength needed to overcome challenges and achieve this milestone.

ABSTRACT

Natural disasters pose a persistent and escalating threat to human societies across the globe, resulting in widespread loss of life, destruction of infrastructure, economic instability, and long-term environmental degradation. Among various natural hazards, floods are recognized as one of the most frequent, complex, and devastating disasters, particularly in developing countries. Pakistan remains highly vulnerable to flood events due to its geographical position, extensive river systems, monsoon-driven climate, rapid population growth, unplanned urbanization, and limited disaster-resilient infrastructure. Recurrent floods in recent decades have caused severe damage to agriculture, housing, transportation networks, energy systems, and public health, emphasizing the critical need for reliable flood forecasting and effective risk assessment mechanisms. Historically, flood prediction and disaster risk management have relied on traditional hydrological models, historical trend analysis, and conventional statistical techniques. While these approaches have contributed to understanding flood behavior, they often face significant limitations in handling large-scale, heterogeneous, and rapidly changing datasets. Moreover, traditional models tend to be less adaptive to climate variability and extreme weather events, which have become more frequent due to climate change. As a result, delays in prediction accuracy and limited forecasting capabilities reduce the effectiveness of early warning systems and disaster preparedness strategies. Recent advancements in Machine Learning (ML) and data-driven technologies have introduced innovative opportunities for enhancing disaster prediction and management. Machine learning techniques offer the ability to process vast volumes of complex data, identify nonlinear relationships, and uncover hidden patterns that are not easily detected through traditional analytical methods. These

capabilities make ML particularly suitable for flood forecasting, where multiple climatic, hydrological, and geographical factors interact dynamically. This research explores the application of machine learning approaches to flood disaster modeling and forecasting in Pakistan, aiming to improve predictive accuracy and support informed decision-making for disaster risk reduction. The primary objective of this study is to evaluate the effectiveness of selected machine learning algorithms in predicting flood risks and generating future flood risk forecasts. By analyzing historical flood-related data and applying advanced predictive models, the research seeks to provide a comprehensive framework that can enhance disaster preparedness and mitigation planning. The study utilizes an extensive dataset comprising flood-related records collected over a 22-year period from 2000 to 2021. This dataset captures temporal trends, seasonal variations, and regional differences in flood occurrences, allowing for a detailed examination of flood dynamics across Pakistan. To ensure a region-specific and accurate assessment, the analysis is conducted separately for the four provinces of Pakistan. This provincial-level approach enables the identification of localized flood patterns and vulnerabilities that may be overlooked in national-level analyses. By incorporating regional characteristics such as river basins, rainfall distribution, and historical flood frequency, the study enhances the relevance and applicability of the forecasting results for local disaster management authorities. Based on the historical dataset, flood risk forecasts are generated on a monthly basis for the future period from 2025 to 2030. This extended forecasting horizon provides valuable insights into potential future flood scenarios and supports long-term planning and resource allocation. The forecasting framework is designed to move beyond reactive disaster response by enabling proactive risk identification and early preparedness measures. Four machine learning models—Decision Tree, Random Forest, Linear Regression, and Support Vector Machine—are employed in this study. These algorithms are selected due to their widespread adoption, predictive capability, and suitability for classification and forecasting tasks in environmental and disaster-related research. Each model is trained using historical data and evaluated through systematic validation techniques to assess performance in terms of accuracy, consistency, and reliability. Comparative analysis of these models allows for an in-depth understanding of their strengths, limitations, and practical applicability in flood forecasting. The study also emphasizes the importance of data preprocessing, feature selection, and model optimization in improving predictive performance. Key variables such as rainfall intensity, river discharge levels, seasonal climate indicators, and historical flood occurrence patterns are incorporated into the modeling process to enhance learning efficiency. By carefully selecting and analyzing these variables, the research ensures that the models effectively capture the complex interactions influencing flood behavior. The results demonstrate that machine learning-based models significantly outperform traditional analytical methods in terms of predictive accuracy and timeliness. The ability of ML algorithms to identify subtle trends and anomalies in historical flood data enables more reliable flood risk predictions. The monthly flood risk projections produced in this study offer actionable insights for disaster management authorities, enabling the development of effective early warning systems and targeted mitigation strategies. Beyond predictive performance, the findings highlight the broader role of machine learning as a decision-support tool in disaster management. Accurate flood forecasts can support infrastructure planning, land-use management, emergency response coordination, and community-based preparedness initiatives. When integrated into institutional and policy frameworks, machine learning-driven predictions can enhance coordination among stakeholders and improve the overall effectiveness of disaster response efforts. Furthermore, this research contributes to the growing body of literature on the application of artificial intelligence and machine learning in environmental risk assessment, particularly within the context of developing countries. By focusing on Pakistan, the study addresses a critical research gap and demonstrates how advanced analytical techniques can be applied in regions with limited resources and high disaster vulnerability. The methodological framework and findings of this research can serve as a reference for similar flood-prone regions facing comparable climatic and socio-economic challenges. In conclusion, this study underscores the transformative potential of machine learning in flood disaster forecasting and management. By integrating historical data analysis, regional modeling, and future risk projection, the research provides a comprehensive and scalable framework for flood risk assessment. The findings support evidence-based decision-making, proactive disaster preparedness, and long-term resilience building. Ultimately, the study contributes to reducing loss of life, protecting critical infrastructure, minimizing economic losses, and promoting sustainable development in flood-prone regions of Pakistan.

TABLE OF CONTENTS

CHAPTER ONE

AN OVERVIEW OF NATURAL DISASTERS AND PREDICTION

Fatima Abbas
Riphah International University Faisalabad

*Corresponding Author: fatima.abbas@riphahfsd.edu.pk

1. Introduction

Natural disasters are large-scale catastrophic events that occur as a result of natural processes of the Earth and the atmosphere. These events include floods, earthquakes, hurricanes, cyclones, wildfires, droughts, and landslides, all of which pose serious threats to human life, infrastructure, economic development, and environmental sustainability. Over the past few decades, the frequency and intensity of natural disasters have increased significantly, making them one of the most critical challenges faced by modern societies. The devastating impacts of such disasters often result in loss of life, displacement of populations, damage to property, disruption of livelihoods, and long-term economic losses. (Moustafa *et al.*, 2022).

One of the major reasons for the increasing severity of natural disasters is climate change, which has altered weather patterns and increased the occurrence of extreme events such as heavy rainfall, prolonged droughts, and intense storms. In addition to climatic factors, rapid urbanization, deforestation, unplanned infrastructure development, and population growth have further increased vulnerability, particularly in developing countries. As more people settle in disaster-prone regions such as floodplains and coastal areas, the consequences of natural disasters have become more destructive than in previous decades. This growing exposure highlights the urgent need for effective disaster prediction and risk management strategies.(Kumar et al.)

Among various natural hazards, floods are considered one of the most frequent and damaging disasters worldwide. Flood events affect millions of people every year, causing widespread destruction of homes, agricultural land, transportation networks, and public infrastructure. In countries like Pakistan, floods represent a recurring threat due to monsoon rainfall, glacial melt, river overflow, and inadequate drainage systems. The socio-economic impacts of floods are particularly severe, as they often affect vulnerable communities, disrupt food security, and impose significant financial burdens on governments and individuals. (Janizadeh et al.)

Timely prediction and early warning systems play a vital role in minimizing the adverse impacts of natural disasters. Traditional disaster prediction methods primarily relied on historical records, manual data analysis, and conventional statistical techniques. While these approaches provided valuable insights, they are often limited in their ability to handle large and complex datasets and to adapt to rapidly changing environmental conditions. As a result, the accuracy and timeliness of predictions using traditional methods remain insufficient for effective disaster preparedness and response.

Recent advancements in data availability, computational power, and analytical techniques have opened new opportunities for improving disaster prediction. Machine learning-based approaches offer the potential to analyze vast amounts of historical and real-time data, identify complex patterns, and generate accurate forecasts. These capabilities make machine learning an increasingly important tool in the field of natural disaster prediction and management. By integrating machine learning techniques into disaster forecasting frameworks, researchers and policymakers can enhance early warning systems, improve decision-making, and reduce the social, economic, and environmental impacts of natural disasters. (Tasnia)

In recent years, the adoption of machine learning techniques has gained significant attention in disaster risk assessment due to their ability to process multidimensional data and adapt to dynamic environmental conditions. Unlike traditional models that often rely on fixed assumptions and linear relationships, machine learning algorithms are capable of capturing nonlinear interactions among climatic, hydrological, and geographical variables. This flexibility allows for more accurate modeling of complex natural phenomena such as floods, where multiple contributing factors interact simultaneously.

Flood prediction, in particular, presents a challenging task due to the uncertainty associated with rainfall patterns, river flow behavior, land use changes, and climate variability. Conventional hydrological models require extensive domain expertise and may struggle to generalize across regions with differing environmental characteristics. In contrast, machine learning models can learn directly from historical data, making them more adaptable to diverse regional conditions. This data-driven nature enables improved flood forecasting accuracy, especially when large datasets spanning multiple years are available (Jarraud and Steiner).

For a flood-prone country like Pakistan, the integration of machine learning into disaster prediction frameworks holds substantial importance. Pakistan has experienced recurrent flooding over the past decades, resulting in significant loss of life, displacement of communities, and damage to critical infrastructure. The increasing unpredictability of monsoon rainfall and accelerated glacial melting further complicate flood risk assessment. Therefore, advanced predictive tools that can analyze historical trends and provide reliable future forecasts are essential for strengthening national disaster preparedness.

Furthermore, effective disaster management does not solely depend on accurate predictions but also on timely dissemination of warnings and informed decision-making. Machine learning-based flood forecasting systems can support disaster management authorities by providing early risk indicators and scenario-based forecasts. These insights can facilitate proactive planning, efficient resource allocation, and the implementation of mitigation measures before flood events occur. Early warnings derived from reliable predictive models can significantly reduce casualties, economic losses, and disruption to livelihoods.

Another important aspect of modern disaster prediction is the availability of large-scale datasets obtained from multiple sources, including meteorological stations, satellite imagery, hydrological sensors, and historical disaster records. Machine learning techniques are particularly well-suited for integrating such heterogeneous data sources into a unified analytical framework. By leveraging diverse datasets, predictive models can achieve higher accuracy and robustness, thereby enhancing confidence in their outputs for practical decision-making.(Abdelkarim and Gaber)

Despite the growing potential of machine learning in disaster prediction, several challenges remain, including data quality issues, model interpretability, and the need for region-specific customization. Addressing these challenges requires systematic evaluation of different machine learning algorithms and their performance across varying geographical and climatic contexts. Comparative analysis of multiple models can help identify suitable approaches for flood prediction and ensure that selected methods align with the specific needs of disaster management systems.

In this context, the present study aims to explore the application of machine learning techniques for flood disaster modeling and forecasting in Pakistan. By analyzing historical flood data and applying multiple machine learning algorithms, the study seeks to enhance understanding of flood risk patterns and generate reliable future flood forecasts. The outcomes of this research are expected to contribute to improved disaster preparedness, informed policy formulation, and the development of effective mitigation strategies to reduce the adverse impacts of floods (Rajab et al.).

2. Predicting Natural Disasters

Predicting natural disasters is a critical component of disaster risk reduction and management, as it enables authorities and communities to minimize the losses associated with such catastrophic events. Accurate and timely disaster prediction provides governments, disaster management agencies, and local communities with sufficient lead time to plan evacuation strategies, mobilize emergency resources, and implement preventive measures. Early warning systems supported by effective prediction models can significantly reduce the loss of human life and property while also limiting long-term economic and social disruption.

Natural disaster prediction plays an essential role in safeguarding livelihoods by protecting agricultural activities, infrastructure, and sources of income. By anticipating disasters before they occur, authorities can take proactive measures to secure assets, reinforce vulnerable structures, and ensure the safety of populations living in high-risk areas. Studies have shown that early prediction and preparedness efforts substantially lower disaster-related economic costs by reducing damage to critical infrastructure and minimizing recovery expenses (Jiang et al., 2021).

Despite its importance, predicting natural disasters remains a challenging task due to the complex and dynamic nature of environmental systems. Natural hazards are influenced by a wide range of interrelated factors, making accurate forecasting difficult. For instance, flood prediction requires the simultaneous analysis of multiple variables, including rainfall intensity and duration, soil moisture levels, river discharge, land use patterns, topography, and drainage capacity. The interaction among these parameters varies across regions and time, adding further uncertainty to prediction models (Jin et al., 2020).

Traditional disaster prediction methods often rely on historical observations and simplified statistical models, which may not adequately capture nonlinear relationships among environmental variables. As a result, these methods may struggle to provide accurate predictions, particularly under changing climatic conditions. This limitation highlights the need for advanced predictive approaches capable of handling complex datasets and adapting to evolving environmental patterns. Enhancing disaster prediction accuracy remains a key priority in reducing vulnerability and improving resilience to natural hazards.(Colacicco et al.)

In recent years, advancements in computational techniques and data-driven methodologies have significantly improved the capability to predict natural disasters with greater accuracy. The increasing availability of real-time environmental data from satellites, remote sensing systems, weather stations, and sensor networks has enabled researchers to develop more robust predictive frameworks. These data sources provide continuous updates on atmospheric and hydrological conditions, which are essential for monitoring rapidly evolving natural hazards such as floods, storms, and landslides.

Machine learning and artificial intelligence-based models have emerged as promising alternatives to conventional disaster prediction approaches. These models are designed to process large and complex datasets and to identify hidden patterns that are not easily detectable through traditional statistical methods. By learning from historical disaster events and environmental conditions, machine learning models can improve predictive performance and provide early warnings with higher reliability. Their ability to adapt to new data also makes them suitable for addressing uncertainties caused by climate variability and changing land-use patterns.(Nurwatik et al.)

Another key advantage of advanced predictive models is their capacity to support multi-hazard disaster prediction. Natural disasters often occur in combination or trigger secondary hazards, such as floods leading to landslides or storms causing infrastructure failure. Integrated prediction systems can help authorities understand cascading risks and prepare comprehensive response strategies. This holistic approach enhances disaster preparedness by allowing stakeholders to anticipate multiple scenarios and allocate resources more effectively.

Despite these advancements, several challenges continue to hinder the effective prediction of natural disasters. Data quality and availability remain major concerns, particularly in developing countries where monitoring infrastructure may be limited. Incomplete or inconsistent datasets can affect model accuracy and reliability. Additionally, the complexity of advanced prediction models can raise concerns regarding transparency and interpretability, making it difficult for decision-makers to fully understand and trust model outputs.(Dai et al.)

To address these challenges, ongoing research emphasizes the importance of combining technological innovation with institutional capacity building. Improving data collection systems, strengthening collaboration among research institutions and disaster management agencies, and ensuring the effective communication of prediction results are essential for maximizing the benefits of disaster forecasting. When integrated into early warning systems and policy frameworks, reliable disaster prediction models can significantly enhance resilience and reduce the adverse impacts of natural hazards.

In this context, strengthening disaster prediction capabilities remains a critical priority for disaster risk reduction efforts. By adopting advanced analytical techniques and continuously improving predictive frameworks, governments and disaster management organizations can enhance preparedness, reduce vulnerability, and support sustainable development in disaster-prone regions.(J. Jin et al.)

3. Machine learning (ML)

Technology has brought about different opportunities to enhance the prediction of disaster in the society over the recent past. Artificial intelligence (AI) encompasses machine learning (ML), which provides tools which allows predictive modeling on large amounts of data in order to capture patterns that may otherwise pass unnoticed in disaster prediction. Historical data of disasters, real time sensor input and environmental data and analysis can help the machine learning algorithms to predict the space-time when disasters are more prone to happen. These models may be very beneficial for the simulation of floods, when several parameters, including weather conditions, geographic location and water levels have to be assessed in a short amount of time (Mishra *et al.*, 2022). The use of machine learning in disaster prediction models is poised to help authorities forecast and make accurate and timely responses to disasters because of the extra time these models can buy for preparedness (Hashi *et al.*, 2021). However, these models are sensitive to the qualities of the data feeding them and the ongoing improvements upon the algorithms that capture real-life likelihoods. Mainly, it concerns society's direction toward preventing the occurrence of flood disasters, the use of machine learning to enhance flood prediction. Tills this time floods have been on the rise and their effects disastrous hence the need to develop efficient means or ways of making a forecast in order to minimize loss of the lives of humanity as well as property (Baydargil *et al.*, 2018).

In addition, machine learning techniques enable continuous model improvement as new data becomes available, which is particularly important in the context of evolving climatic conditions and increasing disaster frequency. By updating models with recent observations, predictive systems can maintain relevance and accuracy over time. This adaptability enhances the practical applicability of machine learning in disaster management frameworks, especially for flood-prone regions.

Furthermore, machine learning-based disaster prediction supports evidence-based decision-making by providing quantitative risk estimates rather than subjective judgments. These insights assist disaster management authorities in prioritizing high-risk areas, optimizing resource allocation, and strengthening early warning mechanisms. As a result, machine learning serves as a valuable analytical tool that complements traditional disaster management approaches and contributes to more resilient and proactive flood risk reduction strategies (Hackathon).

4. Challenges in Predicting Flood Disasters

Forecasting of floods is tasking because it involves several conditions which may be complex and constantly changing. Floods occur in various ways, especially in the flash floods, which areupy, intense and short-lived or result from the over flowing of rivers. These are dynamic occurrences in which the nature of predictive models receiving data upon them has to respond instantaneously; not easy without the right data gathering frameworks (Motta *et al.*, 2021). Most areas especially the developing world or the rural areas do not possess the effective sensor networks, supervision systems as well as real-time data capturing systems needed for early flood prediction. This makes time series difficult to predict because the models used must be informed by adequate data which in this case is lacking (Syifa *et al.*, 2019). Perhaps, one of the most difficult aspects of flood prediction is the connection between different hydrological factors including rainfall distribution, degree of soil water, river water and the terrain. These factors are not only connected but also very volatile indicating significant changes within short periods (Makker *et al.*, 2019). The provision of a comprehensive model that can dissect and analyze such data in real-time, however, simplification of all these factors into an efficient model to predict floods yield small variations such as slight rainfall forecast disparity can hasten large scale flood prediction discordance (Watik and Jaelani, 2019). In many of these hydrological systems, the complexity and variability within the system makes it challenging for machine learning models to consider all possible conditions that might affect its performance in the optimization (Shah *et al.*, 2019). The prediction of flood is also challenging due to urbanization. Sprawling cities with growing populations contribute to heavy constructions of infrastructure like roads and buildings and concrete structures all of which lead to higher runoff during rains. This disrupts the natural methods of water flow hence increases flood occurrence especially in areas that the floods are rarely experienced. The process of urbanization is frequently fast and unbalanced, and efficient updating of the models is impossible. When the new infrastructure is developed, the configurations are altered hence requiring constant renewal of such models. Otherwise, there might be outdated models which underprice or misestimate the risks in flood-prone cities (Broto, 2017). This is also attributed to difficulty in flood prediction, given that climate change is now also a factor. With the increase in the world mean temperature, climatic conditions have suffered a change and there is an increase in the incidences of effects of climate change such as increased rainfall and storm. In most traditional models to predict floods, these new conditions might not be in consideration hence resulting in poor and inaccurate predictions. Of course, machine learning models must be periodically retrained by using new data to fairly estimate the contemporary climate, the process which implies regular data collection and model recalculation. If these updates are not made, then the accuracy gradually declines as the impacts of climate change progress (Serre and Heinzlef, 2018).

Moreover, the spatial heterogeneity of flood-prone regions poses another significant challenge. Different catchment areas exhibit unique geomorphological and hydrological characteristics, resulting in localized flood behaviors that cannot be generalized across regions. Consequently, predictive models must account for both micro-scale and macro-scale variations in terrain, land use, and hydrodynamic properties (Teng et al., 2017). This spatial complexity is particularly problematic in regions with limited historical flood records, where models lack sufficient calibration data, increasing uncertainty in flood forecasts (Samadi et al., 2020).

Integration of multi-source data, such as satellite imagery, rainfall radar, and river gauge measurements, offers potential improvements; however, the heterogeneity in data formats, temporal resolutions, and data quality presents additional difficulties. For example, satellite-derived precipitation estimates often suffer from temporal latency and spatial inaccuracies in mountainous or densely vegetated regions, which may lead to delayed or misleading predictions (Wu et al., 2021). Moreover, the computational demands of processing high-resolution data in real-time limit the feasibility of continuous flood monitoring, particularly in resource-constrained settings.

Socioeconomic factors also influence the effectiveness of flood prediction. Vulnerable populations in rural or underdeveloped areas may lack awareness, communication infrastructure, or early warning systems, rendering even accurate predictions less impactful. The social dimension emphasizes that technological solutions must be complemented by robust policy frameworks, disaster preparedness, and public education to effectively reduce flood-related risks (Kundzewicz et al., 2018).

In recent years, hybrid modeling approaches combining physical hydrological models with machine learning techniques have shown promise in addressing some of these challenges. By leveraging data-driven insights alongside process-based understanding, such models can adaptively learn from changing conditions while retaining the interpretability of classical hydrological models. Nevertheless, the continuous evolution of environmental conditions, urban landscapes, and climate variability necessitates ongoing model refinement and validation to maintain predictive accuracy (Mosavi et al., 2018).

5. The Role of Technology in Disaster Prediction

Looking at the disaster prediction in the modern world, technology has led to easy and accurate forecasts of disasters as well as reaction to the calamities. Of these, machine learning has been deemed a revolutionizing instrument that improves the capability to process massive data sets and estimate the likelihood of natural disasters happening (Lei *et al.*, 2021). Given this, ML algorithms can uncover patterns and correlation that cannot be easily determined by statistical techniques (Ahmadlou *et al.*, 2021). Artificial Intelligence is a broader category of computer science which encompasses machine learning as the subcategory that allows computer systems to make certain decisions based on the information received and rewrite their algorithms to improve their efficiency. This capability is more so in the disaster prediction scenario where the mechanism involved are many and complex mostly involving many variables. For example, using floods where it can process huge historical data, current sensors, weather data, and geographical data to produce a flood prediction that considers many factors (Althuwaynee *et al.*, 2012). Another benefit associated with the application of the machine learning in disaster prediction is the capacity to handle big data. Individual uses traditional forms of forecasting and the methods can fail to notice important information due to the array of data available. On the other hand, ML algorithms can operate on different data inputs from different sources like satellite, rainfall data, river gauges, and social media to create an integration of flood risks. This capability enhances the formulation of flood scenarios and thus reliability of the results (Tanim *et al.*, 2022). In the case of the specific example of flood prediction, the Machine Learning algorithms can be used to look for significant predictors of floods which include; precipitation level, the degree of soil saturation and the flow of water in an urban area. Training machine learning on historical flood data, the model learns the interactions of various factors and is capable of producing the models that will predict the possibility of floods in a certain region (Lantz, 2019). For instance, in a case where rain is expected to be heavily, an ML model can use the current data of soil moisture, the states of rivers to establish the possibility of flood downstream (Tehrany *et al.*, 2014). Our machine learning system tracks environmental updates including modifications in local weather patterns and new area developments. As development builds cities and marine areas transform their primary flood triggers will evolve. Machine learning lets us update predicted outcomes frequently to keep our forecasts current with actual situation development. Our society requires state-of-the-art weather forecasting technology to track how global warming increases extreme weather events (Zhong *et al.*, 2021). Machine learning systems that work with GIS technologies help develop better disaster prediction results. GIS operates with flood data to display locations showing their flood risks per ML model processed information. Flood management agencies combine current technologies with dynamic mapping to create real-time flood hazard maps which they use to guide rescue operations effectively.(Stynes and Pathak)

The integration of Internet of Things (IoT) devices further strengthens the capabilities of machine learning in disaster prediction. IoT sensors deployed across rivers, reservoirs, and urban drainage systems provide continuous real-time data on water levels, precipitation intensity, and soil moisture. This constant data stream allows predictive models to update dynamically, improving lead times for early warning systems (Khan et al., 2020). Additionally, the use of cloud computing infrastructure ensures that massive datasets from multiple sources are processed efficiently, enabling real-time analytics and visualization of potential disaster zones (Rahman et al., 2021).

Moreover, hybrid models combining ML algorithms with traditional hydrological and meteorological models have shown improved accuracy in flood forecasting. These models leverage both the process-based understanding of water flow and the pattern-recognition strength of machine learning, reducing errors caused by incomplete historical data or unpredictable climatic shifts (Mosavi et al., 2018). The adaptability of such models is crucial in regions affected by rapid urbanization or climate change, where past flood patterns may no longer represent future scenarios.

Another emerging approach is the utilization of social media and crowdsourced data in conjunction with ML models. By analyzing real-time reports, images, and videos shared by citizens, predictive systems can detect unusual events and validate sensor measurements, providing a complementary layer of situational awareness (Imran et al., 2016). This integration not only enhances the timeliness of predictions but also aids in the efficient allocation of emergency resources during flood events.

In conclusion, technology, particularly machine learning integrated with IoT, GIS, cloud computing, and crowdsourced data, plays an increasingly critical role in modern disaster prediction. Its ability to process heterogeneous, large-scale, and dynamic datasets enables more accurate forecasts, improved early warning systems, and effective disaster response strategies. Continuous advancements in these technologies will likely further enhance predictive performance and societal resilience to natural disasters in the future (Lei et al., 2021; Tanim et al., 2022; Zhong et al., 2021).

6. Types of Machine Learning

Artificial intelligence (AI) is a big category of technologies made in order to create autonomous systems that could learn from the data and draw conclusions without the help of people, and machine learning (ML) is a part of that big family. Machine learning is of several forms and each form is applicable to different tasks and uses in its proper context. The three general types of machine learning are supervised learning, unsupervised learning and reinforcement learning (Mosqueira-Rey *et al.*, 2023).

6.1. Supervised Learning

Get the different types of machine learning, and indicate which of them is the most common today – supervised learning is the process of training a model from labeled data. In this approach, an input and output dataset is given where each input to the system is linked with an output label. The model builds an ability on changing the inputs into the desired output due to patterns of data during training phase. That is why supervised learning is used for activities such as classification and regression. In classification, the objective is to foresee two or more discrete values or brands such as spam or non-spam email. In regression, the target variable is also continuous in nature such as modeling house prices to estimate them through size, location, number of rooms and so on. The most representatives of the supervised model are linear regression, logistic regression, decision trees, random forests, support vector machines (SVM), and neural networks.

6.2. Unsupervised Learning

This is especially applied when the dataset does not contain an output that is labelled. Here, the model tries in some way to find some structure or regularity in the input data without knowing what the output should look like. Unsupervised learning is important in exploratory data analysis we do not know what to look for it in the data and it is helpful in data mining as it can help us discover hidden patterns in the data. Some of the poignant tasks that might be accomplished in unsupervised learning are clustering and dimensionality reduction. Supervised methodologies like k means clustering, hierarchical methodology clusters the identical like features together. LDA and t-SNE techniques are used for feature extraction, and principal Components Analysis is used for reducing the dimensionality of a feature set. Such type of learning is helpful in scenarios like customer clustering, outlier detection, and extracting features (McAlpine *et al.*, 2022).

6.3. Reinforcement Learning

RL is a branch of machine learning in which an agent learns in an environment by using feedback from the system in which it operates. In this framework, the agent acts to make decisions to maximize a reward signal that results from its interaction with the environment. The agent is able to adapt to decisions based on the result of decisions and thus the agent changes phase through trial and error. The field of application of reinforcement learning includes game playing, robotics as well as autonomous system. For example, Reinforcement learning has been used in constructing systems that that can perform super human like game such as chess or GO. The best-known form of RL algorithm is Q-learning that helps the agent to build a policy that would maximize the total amount of reward. Some of the other well-known algorithm are Deep Q learning (DQN) and Precise Policy optimization (PPO) (Moussaoui and Benslimane, 2023).

6.4. Semi-Supervised Learning

Semi-supervised learning is a type of learning that comes in between supervised and unsupervised learning a little bit of both. This method simply calibrates a model on a little amount of supervised data and big number of unsupervised data. This is with the aim of using the UD for enhancing performance of the model rather than using it to create highly accurate LD which takes lots of time and cost to obtain. In particular, semi-supervised learning is very complementary in cases where it is convenient to label data: for example, in image classification, using natural languages, etc. In semi-supervised learning, the algorithms typically involve an application of the method that has been employed in supervised learning to the labeled data, then exploitation of structures inherent to the unlabeled data (Yang *et al*., 2022).

6.5. Transfer Learning

Transfer learning is a process of using the information that has been obtained in previous work and applying it towards a new goal. In this approach, a complete model learned on a large database is tweaked on the new string of images from the target job. Transfer learning is mostly effective where the new task has very few labeled data to train the model, meaning that generality is achieved and training is shortened. Fortunately, transfer learning of deep neural networks has recently emerged as a popular practice in quite active research areas, including computer vision and natural language processing. For instance, BERT and ResNet model, that is, BERT for natural language processing and ResNet for computer vision, have been pre-trained using large datasets and can be fine-tuned for different tasks, such as sentiment analysis or object recognition with a small amount of adjusting (Yang *et al.*, 2020).

7. Types of Flood Disasters and Their Risks

Flood hazards on the basis of their causes and characteristics are classified into several types and each type poses different threats to life, built environment and natural reserves. It is crucial to comprehend what these types are as far as organizing flood risk management and disaster response and preparedness is concerned (Glago, 2021).

7.1. River Floods

River floods are caused by the inundation of rivers beyond their banks because of the increased water levels arising from; rainfall, snow meltage or disruption of upstream dams. This type of flooding takes time to build up gradually hence enabling an individual or families to prepare for it in some way. But, extended period of rain enhanced the problem of flooding of the river in affected areas particularly areas with poor drainage system. Hazards of river floods include fatalities, damage to houses and property, and flooding of roads, bridges, and other structures as well as interruption of social services (Abrahams *et al.*, 1976).

7.2. Flash Floods

Flash floods refer to a series of floods that happen within a short period of time, minutes or hours with many occurring soon after a heavy downpour in areas with hills or urban demanded. They can also be caused by rapid snow or ICE. Flash floods are even more perilous than other types of floods since they occur suddenly and unexpectedly and may sometimes take a life. These are; accelerated rate of slope failure, debris flow and tremendous loss of property, primarily areas or tracts of land and roads (Ahmed *et al.*, 1999).

7.3. Coastal Floods

The floods again happen in areas near the ocean or sea and they include floods from storm surges, high tidal levels as well as floods from tsunamis. There are floods from such compelling factors like hurricanes or cyclone, which have considerable impacts on coasts and the structures within them. such threats as loss of property, human lives, and the future consequences of coastal floods for the economic situation, especially in coastal zones heavily depending on the income from tourism (Abt *et al.*, 1989).

7.4. Urban Floods

The flood that takes place in urban areas is caused by the excessive water from a downpour that floods cities' drains and consequently inundates streets and properties. They intensify during urbanization because paved surfaces cause runoff and hinder the ground's unnatural ability to retain it. The impacts of urban flooding can be summarized as erosion of transport network, destruction of property and vehicle, pollution of water sources, andypy risks of contracting waterborne diseases (Cosgrove and Rijsberman, 2000).

7.5. Pluvial Floods

It is characterized by flash floods on the basis of heavy rainfall that results to surface flooding on areas that may not be usually flooded. This type of flooding may happen on its own and not by the flooding of rivers, and is most common in the large towns and cities which normally have blockage problems. Potential consequences of pluvial flood include losses in properties and infrastructure and agriculturally productive land, probable threats to health and lives of people (Rosenzweig *et al.*, 2018).

7.6. Groundwater Floods

This is normally as a result of autotypic inundation through high level ground water table resulting from either a long rainfall season or over pumping of water from ground well sources. This flooding can be slow and there could even be some time to prepare for it but the effects are often devastating mainly on the agriculture. These include crop damage, impaired soil characteristics and, difficulties in water quality after the long term (Fan *et al.*, 2013).

7.7. Ice Jams and Floods

Ice jam flood which is also called river ice flood is flooding which is as a result of blockage of the river channel by ice hence slowing down the movement of water and backing up. These phenomena occur in winter and early spring when the ice melts. Some of the consequences of ice jam floods include; loss of property, affects structures and properties and there is always a hazard to life (Murthy and Bobba, 2021).

8. Flood Disasters in Pakistan

Currently, flooding is a common natural disaster in Pakistan to a great extent influencing people's economy, infrastructure, and lives. Pakistan's location, along with tremendous variability in its climate and geographic features, exposed the country to different kinds of floods, including river floods, flash floods, and urban floods (Jordan and Mitchell, 2015).

8.1. Historical Context

Of the world's worst floods ever recorded, some have occurred in Pakistan; The flood happened in the year 2010 is among the worst disasters to have befell on Pakistan, impacting around twenty million people. These were precipitated by intensified monsoon rainfall that led to floods that disrupted the three provinces namely, Sindh, Khyber Pakhtunkhwa, and Punjab. Just imagines that whole villages were flooded, and structures like roads and bridges as well as schools and health facilities were crushed.

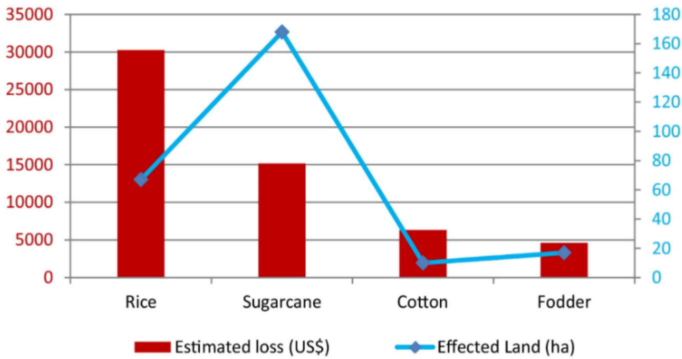


Figure 1.1: *Damaged crops with estimated economic loss in US$*

After the tragedy of 2010, Heavy inundation again hit Pakistan in 2011 with Sindh province worst affected, though many people were still rebuilding their lives from the previous year's deluge (Beltaos, 1995). As you know, major floods happened in 2010 and 2011 inundating several million people and continued to place a burden on the country's recovery. In August this year the monsoon rains severely flooded some parts of Pakistan in the most devastating way. The homes of more than 33 million people were damaged and many agricultural lands and infrastructures were destroyed. It was such a loss which pointed at need to enhance flood management as well as management of disasters in the region. Certain numerical summary is provided among the most effected district of Punjab in flood 2010 (Mahmood *et al.*, 2019).

Table 1.1: *Assessment of 2010 flood disaster causes and damages in district Muzaffargarh, Central Indus Basin, Pakistan*

Sr. No	Type of house	Damaged	Estimated economic loss (US$)	Sr. No
1	Mud houses	496	134,912	1
2	Brick houses	995	451,730	2

9. Impacts of Flood Disasters

The impacts of flooding in Pakistan are multifaceted.

9.1. Human Displacement

Often floods displace millions of people, dismantling their homes and sending them to the camps with plastic tents. This displacement affects people's sources of income and makes them vulnerable to disease and poverty levels (Cazabat, 2022).

9.2. Economic Loss

Disastrous impacts are major losses within a short period of time cutting across agriculture and livestock production, infrastructure and property. Due to the fact that agriculture is a main source of revenue for many families, crop devastation leads to the permanent consequences for the food security and existence (Reed *et al.*, 2022).

9.3. Health Risks

Flood waters can lead to the spread of waterborne diseases, including cholera and malaria. The lack of clean drinking water and inadequate sanitation facilities further exacerbate health risks for affected populations (Lee *et al.*, 2020).

9.4. Infrastructure Damage

Floods caused extensive damage to infrastructure, including roads, bridges, schools, and hospitals. This disruption hampers recovery efforts and can take years to rebuild (Bhatta *et al.*, 2024).

9.5. Environmental Degradation

Flooding can lead to soil erosion, loss of vegetation, and degradation of natural ecosystems. The long-term effects on biodiversity and natural habitats can be significant (Ekka *et al.*, 2023).

9.6. Need for Improved Flood Management

Since the last decade, the flood disasters have occurred frequently and the intensity has also raised greatly in Pakistan therefore the management of floods need to be enhanced. This also envisages the formulation of effective flood forecasting models using scientific methods like; artificial neural networks to improve on the efficiency of flood fore casting, and subsequent evacuation measures. Fuzzy logic coupled with the GIS along with real time data analysis will greatly enhance the accuracy of flood risk prediction and management (Hussain *et al.*, 2023).

10. Challenges and Limitations

Problems are still present even with the aid of machine learning in the prediction of floods demonstrating the following: Reliability of the forecasts is highly sensitive to the quality of input information. Hence when the available data is either full or a complete measure, the forecast is likely to be inaccurate. Furthermore, even though an ML model can learn past trends, they have problems when it comes to events that have not occurred in the past, such as Brexit and changes in climate. The dynamic nature of flood event and the need to employ big data technologies make continuous update of the algorithm and integration of new data source critical for enhancing reliability of the machine learning.(Mabdeh et al.)

Another limitation lies in the interpretability of machine learning models. Complex algorithms such as deep neural networks or ensemble methods often function as “black boxes,” making it difficult for decision-makers to understand the reasoning behind specific predictions (Doshi-Velez & Kim, 2017). This lack of transparency can reduce trust in automated forecasts and hinder their adoption by local authorities responsible for disaster management. Additionally, computational resource requirements for training and updating high-dimensional models can be substantial, particularly when handling multi-source datasets involving satellite imagery, IoT sensors, and social media feeds (Rahman et al., 2021). Such demands may be prohibitive in developing regions where technological infrastructure is limited.

Spatial and temporal resolution of data also poses a significant challenge. Low-resolution datasets or sparse sensor networks can lead to underestimation of flood severity or delayed warning times. Conversely, high-resolution datasets, while more informative, require extensive storage and processing power, which can slow down real-time predictions (Teng et al., 2017). Moreover, inconsistencies across data sources for example, differences in measurement standards between government agencies and community-based reporting can introduce biases that negatively affect model accuracy.(Merz et al.)

Finally, external factors such as rapid urbanization, land-use change, and climate variability continuously modify flood dynamics. Even well-trained models may struggle to adapt to novel conditions if retraining is infrequent or if historical data does not capture these changes adequately (Kundzewicz et al., 2018). To address these challenges, continuous model validation, integration of hybrid modeling approaches, and investment in high-quality, real-time datasets are critical for enhancing the robustness and reliability of ML-based flood prediction systems.

11. Problem Statement

Flood is one form of natural disaster which affects the lives of people, property, incomes and lives of people across the globe. The fact that floods are unpredictable, and recent frequent disasters have been more intense and devastating due to climate change and increased urbanization of the areas makes disaster management a huge task. The previous flood predictions entail use of historical records and generic models which do not take into consideration the flowing environmental changes and real-time characteristics, thus, the delay or improper warnings (De Wrachien *et al.*, 2011). However, implemented combination of real-time data and machine learning for flood prediction and mitigation is very limited. The existing models fail at the scales of data processing and accuracy needed for risk assessment, especially in cities where infrastructure and population density raise vulnerability. This leads to the lack of enough early warning organizations and emergency responses hence posing a danger to human life and properties. In order to address these, we have to adopt new techniques with the assistance of advance form of machine learning to enhance the accuracy of the warnings and the time it takes to get them working. As a result, it becomes easier to optimum develop and implement improved flood mitigation solutions that are much more effective and capable of effectively protecting human life and property against the negative impacts of this perpetual calamity (Li *et al.*, 2019).

12. Significance of the Study

Flood disasters are a major problem for various communities, structures, and economies, as they occur in areas characterized by climate fluctuations and growing population densities. Solving these problems presupposes the presence of sophisticated anticipation processes that increase the precision and the velocity of natural disaster prognosis. This research work examines the role of applying accurate ML models to enhance flood forecasting tools since real-time responses are crucial in minimizing the impacts of floods (Tehrany *et al.*, 2014). Thus, the improvements and revisions to the general machine learning models presented in this work meaningfully contribute to the refinement of flood prediction methods. These developments make it possible for early warning systems to be of increased efficiency that means early interventions with a view of protecting lives, properties, and resources. The real-time data analysis support integration also strengthens the predictability, promotes the use of emergency solutions with up-to-date information (Jato-Espino *et al.*, 2018). Further, the research gives an alternative comparison of different machine learning models in case of flood forecast. The comparative study is therefore useful for identifying the strong and weak aspects of the two models, with a view to applying only the most suitable techniques for disaster prevention. The research findings have a broad applicability in artificial intelligence and natural disasters other than flood prediction (Sarker, 2021). The groups of people to whom this research is important include researchers, disaster management organizations and the general public because it highlights a number of critical gaps in disaster management practices that need to be addressed. Overall, issues such as prediction accurateness, timely period, and decision-making from data findings enhance the establishment of better mitigation measures. The findings are expected to provide timely information for disaster management authorities, policymakers and researchers to help improve a society's level of disaster preparedness when confronted with disastrous natural events (Londhe and Panchang, 2018).

13. Motivation and Goals

Flood, an unequivocal natural catastrophe has long lasting impacts on the affected societies, economy and environments. In view of the fact that such incidents are gradually becoming more frequent coupled with their intensity by virtue of climate change and expansion of cities and towns, there is no doubt that the development of works on disaster prediction and management systems is very essential. The conventional approach to risk assessment necessitates the application of fixed models which do not allow for real-time extrapolation, and thus are not nearly as exacting or timely as is necessary for proper disaster prevention. These limitations agenda this research mainly (Lisboa *et al.*, 2023). Huge possible of Machine learning in dealing with various complex data based issues make it possible to rethink the flood prediction systems. Primarily, the current optimization techniques using machine learning algorithms can be extended to open a way to allow further development of existing optimization models and to increase prediction accuracy and real-time analysis. This motivation is further compounded by improved early warning system needs for the protection of vulnerable people, and vital infrastructures (Pham Quang and Tallam, 2022). The purpose of this research is therefore to add to the formulation of effective preventive measures towards flood disaster in order to bring about improved prediction of the calamity. To that end, this research aims to improve the effectiveness of flood forecast by implementing improved machine learning algorithms. The third critical goal is to enhance the capability of big data in real-time where early warning systems provide timely identification of new threats. Moreover, this work also seeks to assess and compare the capabilities of various machine learning techniques for natural disasters prediction so as to foster growth in predictive systems research (Ng *et al.*, 2021). Through achieving these goals, the research aims to develop a base on which better practices in disaster management can be established, and to create a higher level of preparedness when future flood disasters occur (Abdollahzadeh and Gharehchopogh, 2022).

In addition to the aforementioned objectives, this research emphasizes the integration of multi-source environmental data to enhance predictive accuracy. By combining historical flood records, real-time hydrological measurements, meteorological forecasts, and geospatial information, machine learning models can account for the complex interactions between climatic, hydrological, and anthropogenic factors (Lei et al., 2021). This holistic approach enables more nuanced predictions that are sensitive to both spatial variability and temporal dynamics, thereby improving the relevance of forecasts for specific vulnerable regions.(Mosavi et al.)

Furthermore, the research investigates the potential of ensemble learning techniques, which aggregate predictions from multiple machine learning algorithms to mitigate individual model biases and reduce forecasting errors (Tanim et al., 2022). Such approaches are particularly valuable in flood prediction due to the inherently stochastic nature of precipitation events and riverine responses. Ensemble methods also offer greater resilience against data scarcity, as combining models trained on different datasets can improve generalization to unseen conditions.

Another key aspect of this study is the exploration of real-time adaptive learning mechanisms. Machine learning models that continuously update as new data becomes available can dynamically adjust predictions in response to changing environmental conditions, urban development, or unexpected climatic anomalies (Zhong et al., 2021). This adaptability is crucial for early warning systems, which must provide timely alerts to authorities and communities, thereby minimizing potential damages and saving lives.

Finally, this research aims to contribute to decision-support frameworks that bridge the gap between model outputs and actionable disaster management strategies. By translating predictive insights into clear risk maps, priority zones, and evacuation recommendations, the study seeks to enhance the practical utility of machine learning systems for policymakers, urban planners, and emergency responders (Pham Quang and Tallam, 2022). Collectively, these advancements are expected to foster more resilient infrastructures, informed disaster preparedness, and sustainable mitigation strategies in the face of increasingly frequent and severe flood events.

14. Objectives

The objectives of this study were as follows:

1. To enhance the accuracy and timelines of natural disaster prediction using modified machine learning models.
2. To improve real-time data analysis for effective early warning threat detection.
3. To compare the performance of different machine learning models in natural disaster predication.
4. To compare the performance of different machine learning models in natural disaster prediction. This involves exploring ensemble learning techniques that aggregate predictions from multiple algorithms to reduce individual model bias, enhance generalization, and mitigate forecasting errors in flood-prone areas (Tanim et al., 2022).
5. To assess the adaptability and robustness of machine learning-based predictive systems in response to climate change, rapid urbanization, and novel environmental scenarios. This objective emphasizes continuous model validation and the integration of hybrid approaches

combining traditional hydrological models with data-driven ML techniques to improve resilience and early warning capabilities (Ng et al., 2021).

CHAPTER TWO

MACHINE LEARNING APPLICATIONS IN DISASTER FORECASTING

Fatima Abbas
Riphah International University Faisalabad

*Corresponding Author: fatima.abbas@riphahfsd.edu.pk

1. Overview of Machine Learning in Disaster Predictions

Mosavi *et al.* (2018) acknowledged that evaluating the performance of ML models offers a comprehensive understanding of the many methodologies. This article presents the optimal methodologies for short-term and long-term flood predictions. The paper examines the primary trends that have been improving the precision of flood prediction models. Findings indicate that the most effective methods to improve machine learning techniques encompass hybridization, data deconstruction, algorithm ensembles, and model optimization. This survey may be employed by climate scientists and hydrologists to choose the best suitable machine learning method for each predictive task (Jiang et al.).

2. Machine Learning Approaches for Flood Prediction

Khalaf *et al.* (2018) presented a novel approach to predicting flood severity via machine learning techniques, namely a Neural Network architecture. Floods are a sort of natural disaster that can cause extensive devastation. The interplay between precipitation and river discharge, particularly during periods of intense rainfall, is a common catalyst for flooding. The catastrophic consequences of floods on infrastructure and populations are a significant concern given the ongoing climate catastrophe and other severe environmental repercussions. This study introduces a novel flood dataset of 2000 annotated flood events. The results are categorized into three objective categories that indicate varying flood severities, aimed at addressing the challenge of flood control (Ibrahim et al.).

3. Case Study: Neural Network Models for Flood Severity Prediction

The paper presents several machine learning algorithms capable of predicting flood intensity and categorizes the outcomes into three classifications: normal, abnormal, and high-risk. The application of artificial intelligence algorithms for the pre-processing of flood data may enhance outcomes, as indicated by comprehensive research. The categorization techniques attained enhanced accuracy with these methodologies. The random forest classifier surpassed the benchmarked models in our investigations, despite the efficacy of neural network topologies in various areas (Lavers et al.).

4. Impact of Fourth Industrial Revolution Technologies on Flood Mitigation

The tactics for mitigating flood catastrophes globally have been influenced by the primary elements of the fourth industrial revolution, as articulated by Riza *et al.* (2020), which encompass cloud computing, artificial intelligence, big data, and the internet of things. The occurrence of flood catastrophes in Indonesia has roughly quadrupled during the past fifteen years, substantiating this tendency. This literature research examines the advancements of artificial intelligence (AI) and machine learning (ML) in mitigating flood risks in Indonesia. This analysis may also suggest future enhancements to flood control measures. The study indicates that machine learning has not been extensively employed to alleviate flood catastrophes in Indonesia. Predictions of river levels, flows, and precipitation are primary applications, while flood predictions and early warning systems represent secondary applications. Utilizing machine learning techniques in a comprehensive flood prediction and early warning system that incorporates rainfall, river water levels, river discharge, flood forecasts, estimated spatial flooding areas, and early warning communication via the Internet of Things represents a promising advancement in flood disaster mitigation (Heinzlef et al.).

5. Limitations of Traditional Flood Prediction Models

The Rainfall-Runoff Model and other traditional flood prediction models have long been employed to forecast the transformation of precipitation into runoff, hence augmenting river flows (Cirilo *et al.*, 2020). These models will perform effectively if there is an abundance of historical rainfall and river flow data. Conventional models depend on static data and assumptions on land use, which constrains their efficacy in complex regions such as Sindh, where precipitation patterns can vary significantly. These models perform effectively in static environments; nevertheless, they have difficulties in dynamic contexts, such as Sindh, characterized by extensive urbanization and fluctuating weather patterns.(Alam et al., "A Novel Framework Embedding Bayesian-Optimized Ensemble Machine Learning and Explainable Artificial Intelligence (XAI) to Improve Flood Prediction in Complex Watersheds").

6. Advances in Machine Learning for Flood Forecasting

Puttinaovarat and Horkaew (2020) assert that floods are among the most catastrophic natural catastrophes, resulting in human fatalities and significant property damage. Numerous cities are susceptible to the impacts of monsoons and thus endure this natural disaster annually. To assist flood victims, organize evacuations and rescue operations, and execute both urgent and long-term preventive strategies, it is crucial for authorities and the public to be informed about flood events promptly. Critical elements in flood control encompass the geographical locations and intensity of affected areas. A dependable method for forecasting flood occurrences remains elusive. (Yang et al.)The primary functions of the prior technologies were data preparation and human entry. The techniques were excessively time-intensive for prompt and accurate forecasting. Contemporary big data systems provide more comprehensive information; which older methodologies did not thoroughly investigate. This paper proposes a novel flood forecasting technique utilizing an adaptive machine learning architecture that integrates extensive data from many sources, including meteorology, hydrology, geography, and crowdsourcing. Contemporary educational methodologies served as the catalyst for data intelligence. Both subjective and objective evaluations verified that the developed approach could precisely forecast the timing and location of floods. Subsequent benchmarking investigations indicated that the MLP ANN-configured system yielded the most accurate predictions, with an RMSE of 0.10, Kappa of 0.89, MAE of 0.01, and an accuracy of 97.93.

7. Case Study: Flood Risk Assessment in the Yangtze River Delta (YRD), China

The Yangtze River Delta (YRD) is a highly developed area in China, as stated by Chen *et al.* (2020) The relationship between people and land, as well as between people and water, is intricate, and this region is susceptible to floods, resulting in many flood-related catastrophes. Thus, progress in the region is significantly dependent on accurate evaluations of flood risk. This research utilized the YRD urban agglomeration as its case study. The DPSIR conceptual framework was designed to analyze flood disaster indicators. This framework represents driving force, pressure, status, impact, and reaction. Between 2009 and 2018, a risk assessment model for this region was developed utilizing a radial basis function (RBF) neural network. The methodology employed was the random forest (RF) technique, utilized to identify significant flood risk indicators. The GIS was used to develop the risk map to determine the I-V flood risk levels in YRD urban agglomeration: 2016-2018. Al additional analysis revealed that many indices deteriorate the extent of the flood risk significantly. These include the proportional structures of Impervious surfaces urban area to total land area and aquatic area to total land area, the population density level and the emergency preparedness of public administration segments. This paper established that the YRD urban area has increased its vulnerability to floods by up to one decimal point in the last decade because of increased urbanization. It was also possible to establish a positive correlation between the level of economic development and flood threats. Furthermore, the current conditions or existing conditions and the rate on which there is an escalation of statuses of flood risks in the provinces also varied. Some following major cities had reduced flood risk after the setting up of the urban flood control measures from 2016 to 2018. Such results have confirmed the efficiency of the model as most of the results obtained practically coincided with the realities. Last but not least, the possible way to reduce the likelihood of urban flooding in the YRD region was also suggested. This will facilitate decision-making on flood control, disaster mitigation, and emergency management in the region.(Alam et al., "A Novel Framework Embedding Bayesian-Optimized Ensemble Machine Learning and Explainable Artificial Intelligence (XAI) to Improve Flood Prediction in Complex Watersheds").

8. Real-Time Urban Flood Prediction Using Hybrid Models

The erratic nature of precipitation can cause urban floods to result in considerable damage, as noted by Keum *et al.* (2020). Furthermore, they are intricate, highly nonlinear processes. This study develops a classification-based real-time flood prediction model for urban areas by combining a hydraulic theory-based numerical analysis model with a machine learning model. Flood databases are constructed with a two-dimensional inundation model and the Environmental

Protection Agency-Storm Water Management Model (EPA-SWMM) to anticipate various rainfall situations. To enhance the precision of flood range predictions, we employ a probabilistic neural network (PNN) and Latin hypercube sampling (LHS) to categorize the average flood depth data for each map grid into five distinct categories. Inputting the recorded rainfall into a model may predict the representative cumulative volume. A real-time flood map generation system is established by correlating the cumulative volume of each grid with the representative cumulative volume via linear and nonlinear regression techniques. This facilitates the geographic extension of flood depth in accordance with the anticipated representative cumulative volume. The generated model exhibits an 85% goodness-of-fit and operates for a length of 1 minute and 12 seconds in comparison to the results of a validated two-dimensional (2D) flood model. The developed method is capable of predicting rainfall-induced floods to assist in disaster risk management and mitigate health and property damage (X. Jin et al.).

9. Flood Damage Prediction Models: India Case Study

Goel (2020) assert that a worldwide flood is a recurring disaster with extensive implications for human life and the world economy. Annual flood damage assessments in India amount to billions of rupees. Therefore, it is essential to execute flood risk management techniques that assess risk based on damage valuation to mitigate the annual loss of life and property. Most research on flood damage modeling depend exclusively on water depth as the singular variable. The primary cause of floods in India, like to Korea, is substantial precipitation. This study employs Machine Learning techniques to develop a Heavy Rain Damage Prediction Model, analogous to the application of models such as K Nearest Neighbors, tree-based methods, and Gaussian

Naïve Bayes in regression analysis. The research focused on the three states of Bihar, Uttar Pradesh, and Kerala. The random forest and KNN algorithms were the most successful among the selected supervised learning strategies. In states where cyclones, dams, and analogous elements significantly contribute to floods, the results may vary.(Alam et al., "A Novel Framework Embedding Bayesian-Optimized Ensemble Machine Learning and Explainable Artificial Intelligence (XAI) to Improve Flood Prediction in Complex Watersheds").

10. Advances in Deep Learning for Flood Forecasting

Chaudhary and Piracha (2021) observed that flooding remains a significant source of economic, environmental, and human distress worldwide Predictive analyses of flood occurrences help to reduce devastation, save lives, and design warning mechanisms. In particular, the non-linearity of environmental factors has always been a limitation to flood prediction models using standard methods, including meteorological and hydrological models, which are widely used. This is because flood data is usually large and since ML models can handle this and also non-linear relationships, they have been shown to be very effective in enhancing flood prediction. Large numbers of machine learning approaches have been tested for flood forecasting with varying degrees of success. However, some of these limitations linger, especially where real time data is involved and the formulation of good prediction models in dynamic environments. According to Moishin *et al.* (2021), early flood warning systems are effective decision-support measures that may well help to save lives and infrastructure in natural disasters. The current research proposes and evaluates a flood forecasting model using deep learning with CNNs and LSTMs, known as ConvLSTM (Hashi et al.).

11. ConvLSTM-Based Flood Index Prediction

This study used a mathematical construct known as the Flood Index (IF) to assess the duration, severity, and intensity of flood events, utilizing a dataset of precipitation. The IF records the gradual exhaustion of water supplies over time and is utilized in a flood monitoring system. The newly constructed forecasting model anticipates the subsequent daily IF value utilizing prior and real-time rainfall data, together with statistically significant lagged IF augmented by these elements. In flood-prone regions of Fiji, which endure flood-induced devastation almost annually, the proposed ConvLSTM model is evaluated using nine different rainfall datasets. All benchmark methodologies were evaluated over several forecasting timeframes (1 day, 3 days, 7 days, and 14 days), and the results indicate that the ConvLSTM-based flood model is superior. The Root Mean Squared Error (RMSE) of the research sites utilizing the ConvLSTM model was 0.101 for the first period, 0.150 for the second, 0.211 for the third, and 0.279 for the fourth projected period. The RMSE values for the subsequent best model over all four forecast horizons were 0.105, 0.154, 0.213, and 0.282, respectively. Across the four prediction horizons, the Legate-McCabe Efficiency Index (LME) recorded values of 0.939, 0.898, 0.832, and 0.726, respectively, reflecting variability in model performance at particular stations. The findings demonstrated that ConvLSTM is effective for precise IF forecasts and might be utilized in disaster management and risk mitigation during extreme weather events (Aswad et al.).

12. GIS-Integrated Machine Learning Models for Flood Forecasting

Forecasting the timing and extent of flood damage is a crucial aspect of disaster preparedness (Motta *et al.*, 2021). Hydrological and meteorological data have historically been the foundation of conventional models that seek to replicate the flow and buildup of water within watersheds. Recent models integrating GIS with machine learning have gained popularity because to their enhanced predictive accuracy and adaptability, particularly in the context of the increasing availability of large data and advancements in machine learning technologies. These advanced models can be highly beneficial in regions such as Sindh, Pakistan, where the Indus River and seasonal monsoons frequently result in floods. Flood-prone areas such as Sindh might significantly benefit from machine learning models utilizing GIS data. These models can improve the precision of flood forecasts in urban regions by including real-time environmental and meteorological data. This strategy enhances forecast accuracy and provides insights that are both timely and pragmatic (Abdalzaher et al.).

13. Real-Time Flood Detection Using Hybrid ML and DL Models

Hashi *et al.* (2021) define a flood as water that inundates typically dry land or a rise in water levels that significantly impacts human life. Floods are among the most prevalent natural catastrophes, resulting in significant financial losses and damage to homes and products. If these floods could be averted, the residents would have additional time to escape potentially hazardous areas. Numerous scholars have proposed diverse strategies for addressing the flooding issue, including the development of predictive models and the establishment of adequate infrastructure. Individuals in countries such as Somalia would be better served without these solutions, as they are economically inefficient. The fundamental objective of this study is to present a novel and dependable model for real-time flood detection utilizing machine learning and deep learning methods, including Random Forest, Naive Bayes, J48, and Convolutional Neural Networks. This model will measure water levels and predict probable humanitarian crises. The solution to the previously described issues and the investigation into the straightforward simulation of a distinctive methodology for detecting water levels utilizing a hybrid model based on Arduino with GSM modems will be the experimental results of this proposed method. The study shown that, in comparison to alternative machine learning models and classification methods, the Random-Forest algorithm attained the greatest accuracy rate of 98.7 percent. The outcomes achieved with J48 were 84.2 percent, whereas those with Naive Bayes were 88.1 percent. Conversely, the outcomes for recall and precision were often favorable when employing a Deep Learning approach, achieving 87% accuracy. The proposed technique introduces an innovative strategy to flood protection that benefits artificial intelligence (AI), data mining (DM), and deep learning (DL) ("Fatim Abbas").

14. Flood Forecasting in the Chenab River Basin

Accurate flood monitoring helps protect against disaster in flood-prone river systems worldwide according to Ali *et al.* (2021). The hydro meteorological conditions in the Chenab River basin make it the world's most advanced river system with severe hydrology that challenges prediction models. Too many riverbank villages suffer damage during floods and create major economic and human casualties. Our analysis involved creating floods in the Chenab River watershed while also using both the Hydrologic Engineering Centre-Hydrologic Modeling System (HEC-HMS) and the Hydrologic Engineering Centre-River Analysis System (HEC-RAS). The HEC HMS simulation model processed 2014 flood capacity using GMP IMERG-F, TRMM_3B42RT, and GSMaP_NRT datasets. The HEC-RAS model needed calibration during the 1992 flood event and proved successful when tested for the 2014 floods. The outflow observation data helped us determine that the IMERG-F peak flow simulation worked well. Our flood area forecasts achieved better than

90% accuracy when compared with satellite measurements. Our HEC-RAS model demonstrated superior accuracy results both in river and floodplain conditions at a statistical significance of 0.06 (Zaidi et al.).

15. Integration of Remote Sensing with Hydrological Models

The results indicate that early warning systems, flood predictions, and inundation models might all be enhanced by the integration of remote sensing with HEC-HMS and HEC-RAS. The collaborative Indo-Pak management for flood mitigation in the transboundary Chenab River basin has implemented the concept of integrated flood management (IFM) (Zeng and Bertsimas).

16. Urban Flood Prediction and Smart City Planning

Environmental disasters are expected to increase in frequency and severity due to the many effects of climate change, including more extreme meteorological phenomena (Motta *et al.*, 2021). It is imperative to implement adequate safeguards to address such hazards, notably floods, as urban centers are anticipated to grow significantly in the coming years. This project aims to develop a flood prediction system that integrates Machine Learning classifiers with GIS techniques to enhance urban management and resilience planning. This approach may establish the foundation for a future Smart City initiative by identifying pertinent risk indices and factors related to urban floods. The Random Forest Machine Learning model exhibited superior performance, with an accuracy of 0.96 and a Matthew's Correlation Coefficient of 0.77. A GIS model was developed to pinpoint areas with a higher probability of flooding during extreme weather events, hence enhancing the functionality of the Machine Learning model. Consequently, based on the documented history of flooding, the entire city was categorized into hotspots. A flood risk indicator was created by integrating the outcomes of the Random Forest model with the findings of the Hot Spot investigation (Alkharabsheh et al.).

17. Satellite-Based Flood Monitoring and Rapid Response

Mateo-Garcia *et al.* (2021) assert that space-borne Earth observation is a crucial technology for food response, offering valuable data for local decision-makers. A potential approach to reduce the return time in disaster zones from days to ours is the deployment of extensive constellations of small microsatellites known as "CubeSats." However, the limits in power and bandwidth of CubeSats constrain data transmission to ground receivers. Minimize the volume of data transmitted by employing onboard processing to convert large sensor images into smaller data products. The recent PhiSat-1 mission by the ESA offered a power-constrained machine learning accelerator and software for executing customized applications, aimed at demonstrating this concept. This research presents a proficient food segmentation algorithm operating on the accelerator of PhiSat-1. The approach produces food masks to replace the original photographs. The World Floods dataset, recently compiled and made publicly accessible in a standardized manner, comprises 119 verified flooding episodes reported by disaster response organizations globally. Our models are trained on this dataset. This approach is validated by testing the system at independent locations, demonstrating its ability to generate rapid and precise segmentation masks on the hardware accelerator (2023).

18. Review of Advanced Techniques for Flood Management

The utilization of advanced equipment and techniques enabled Munawar *et al.* (2021) to attain their enhancements. The mitigation of environmental and economic harm, along with the avoidance of fatalities, relies on our capacity to respond swiftly to floods and devise measures to diminish their effects. Flood control researchers have explored techniques from several disciplines to address issues such as data analysis, machine learning, image processing, and remote sensing. This article proposes a novel method for categorizing existing research on flood management techniques. The framework addresses the following significant research issues: (1) What are the primary methodologies employed in flood management? What steps of flood management had past studies mostly concentrated on? What are the recommended options for addressing flood management issues? What gaps exist in the literature about the application of technology for flood control? A framework for classifying flood control methods has been proposed to categorize the many examined technologies. Regrettably, no hybrid models combining machine learning with image processing were identified as beneficial for flood prevention. Moreover, there was less success in employing machine learning methodologies following a tragedy. To achieve effective and comprehensive disaster management in the future, it is essential to integrate our knowledge of disaster management with image processing methodologies and machine learning technology (Zhang et al.).

19. Machine Learning Models for Flood Prediction in Nigeria

Lawal *et al.* (2021) assert that Machine Learning (ML) models utilized for flood prediction can enhance both flood alerts and flood mitigation/prevention efforts. Due to their dependence on observational data and minimal computational requirements, machine-learning (ML) methodologies have gained popularity in this context. This study aimed to create a machine learning model capable of predicting floods in Kebbi state utilizing a 33-year rainfall dataset. This approach might subsequently be applied to other flood-prone states in Nigeria. This analysis examines three machine learning algorithms – Decision Tree. These algorithms include the Logistic Regression and the Support Vector Classification and the outcomes from the comparison made will be based on the ROC, Accuracy, and Recall. Interestingly, Logistic Regression proves to be more accurate and to offer higher recall than the two, which means that the result of this approach is generally better. In addition, the Decision Tree was found to yield better results than the Support Vector Classifier. As for the performance, the Decision Tree claimed rather high accuracy rate, which is above the average, and the recall rate below the average. The Support Vector Classification had low accuracy and relatively low Recall Score on a small set of samples, the accuracy and the ROC scores of which are not satisfying (Koutsovili et al.).

20. Improved Artificial Neural Network Models for Flood Forecasting

According to Zhu *et al.* (2021), global advocates report that flooding remains the costliest natural calamity loss. Flood forecasting is very essentials due to the possible damage's floods may bring to people as well as the ecology. Scientists have been using artificial neural network (ANN) in order to quickly predict floods. This research introduces improvements to an existing Artificial Neural Network trained on synthetic events, obtained mainly through the utilization of a hybrid dataset to enrich the training set. Secondly, by selecting the optimal training function from six distinct alternatives, which encompass conjugate gradient backpropagation (CGF) with Polak-Ribiére updates (CGP) and Powell-Beale restarts (CGB), one-step secant backpropagation (OSS), resilient backpropagation (RP), and scaled conjugate gradient backpropagation (SCG). Four real floods were employed to verify that the improved ANN surpassed its predecessor. The root mean square error (RMSE) for the model was 10% for the testing dataset and 16% for the actual events following the application of the new training dataset. The robust backpropagation approach was selected since it decreased RMSE by 15% for the testing dataset and by up to 35% for actual events, in comparison to the other five training functions (Obada et al.).

21. CNN-Based Flood Mapping: OmbriaNet and OMBRIA Dataset

Drakonakis *et al.* (2022) assert that modern civilizations encounter significant dangers to their regular and sustainable functioning due to climate change. Floods, storms, and rising sea levels are common examples of extreme weather events that impair the fundamental functions offered by global ecosystems. Timely and accurate disaster delineation is crucial as it aids in relief efforts, benefits society, the economy, and the environment, and is especially significant in the context of floods, which greatly impact human activities. This article introduces OmbriaNet, a convolutional neural network (CNN) architecture that leverages temporal differences from various sensors to identify changes between permanent and inundated water areas during flood events. To demonstrate the efficacy of the proposed technique, we developed OMBRIA, a dataset for image segmentation utilizing supervised binary classification, comprising bitemporal and multimodal satellite imagery. The total number of images, comprising synthetic aperture radar imagery from Sentinel-1 and multispectral imagery from Sentinel-2, along with ground-truth binary images generated from expert-derived data provided by the Emergency Management Service of the European Space Agency Copernicus Program, is 3,376 (Asif et al.).

22. Advances in Rainfall-Runoff and Hybrid Models

The data involves 23 global floods reported from the year 2017 and up to 2021. Google Earth Engine was used to gather, align and prepare all the data collected for analysis. We conducted several benchmarking experiments using the real-world OMBRIA dataset and compared their results with other competing state-of-the-art methods to measure the effectiveness of our proposed technique. The test results also revealed how the proposed formulation was capable of outcompeting the current standard of producing accurate flood maps. This simply controls the amount of water input to streams and returns excess water to the oceans making it a component of water cycle (Jehanzaib *et al.*, 2022). The runoff modelling can help understand, control and even follow the amount and quality of water resources. This article aims at discussing several rainfall runoff models, the recent advancement in these models and the difficulties for the use of such models in the simulation of current floods. On the basis of algorithms framework and spatial processing, the rainfall-runoff models are conceptual, empirical and physical process-based. Here are the various classifications and the runoff models that articulate them: The SCS-CN model also known as the Soil Conservation Service Curve Number model, The Storm Water Management Model (SWMM), The Hydrological Byråns Vattenbalansavdelning (HBV) model, The Soil and Water Assessment Tool (SWAT) model, The Variable Infiltration Capacity (VIC) model and other popular models within these classifications. In addition, the ANFIS, ANN, DNN, and SVM models have shown improvement in terms of effectiveness for runoff modeling and flood prediction in the past couple of decades. Models that claim to possess data driven then determine the right nature of the input data series and output to simulate runoff. As a result of the models, the improvement and detriment of the models were assessed for what was involved in variability of runoff modeling and flood prediction. The systematic research carried out for this study concluded that the best approach in the assessment of runoff and flood forecasting is the integration of the conventional and artificial intelligence models. This paper identifies potential future areas where both theoretical and experimental research could be directed in order to help hydrological experts select the best models of rainfall-runoff for the prediction of the occurrences of floods and their management, in order to rectify the current insufficient understanding of their interaction (Liu et al.).

23. Enhanced DE-GIUH Model for Flood Forecasting in China

Zhu *et al.* (2022) recognize the DE model as flood prediction model that addresses characteristics of runoff generation in the SD and SH regions. However, due to the rather vague definition of the parameters that comprise it, the empirical unit hydrograph of the subbasin confluence module can sometime be imprecise. This research refines the subbasin confluence module in the DE model with the help of geomorphologic instantaneous unit hydrograph (GIUH). The end product is the DE-GIUH, an enhanced DE model that uses topographic physical factors to determine sub basin confluence. This new model is applied to Wangjiahui Hydrological Station Basin located in Northern China. This study shows that, when compared to the HEC-HMS, the DE-GIUH model provides better estimates of flood characteristics such as component peak discharge, flood peak time, and flood propagation in the catchment area than the DE model. All three models achieved the minimum acceptable QR of 85.0% or higher. During the calibration phase, 90% of the deterministic coefficients (DC) attained a grade of B or better ($DC \geq 0.70$). During the validation period, this proportion rose from 62.5% and 25% to 75%, respectively. This is particularly evident in significant and moderate floods, when DE-GIUH has a much higher percentage of DC attaining grade B and above at 100% in comparison to the other two models. For areas lacking flood surveys, the proposed advanced model presents an alternative methodology (Santos et al.).

24. Two-Dimensional Hydrodynamic Modeling for Flood Simulation

One-dimensional (1D) hydraulic models have been extensively utilized for flood simulations to analyze flood depth and extent maps (Ahmad *et al.*, 2022). While assuming unrestricted flow inside the channel, several other flood characteristics – such as flood velocity, duration, arrival time, and recession time – cannot be accurately simulated using one-dimensional models. It is essential to consider these flood features while preparing for flood protection and evacuation. This study develops a two-dimensional (2D) hydrodynamic model integrated with geographic information system (GIS) and remote sensing (RS) techniques to create additional flood characteristic maps that cannot be produced using one-dimensional (1D) models. The model was executed on the transboundary river Deg Nullah in Pakistan to replicate a catastrophic flood that transpired in 2014. To evaluate the model's efficacy, we employed both real flood extents and imagery of floods captured by the Moderate Resolution Imaging Spectroradiometer (MODIS). An entropy distance-based technique was proposed to assist with the integrated multivariate flood vulnerability classification. The results of the 2D flood modeling simulations aligned well with the observed and MODIS-documented flood extents. The area's most vulnerable in the northwest of Deg Nullah, around Seowal, Dullam Kahalwan, and Zafarwal, saw severe and prolonged flooding. The upper stretch of Deg Nullah is the most susceptible and sensitive location to flooding due to elevated flood velocities, brief flood arrival times, extended flood durations, and prolonged recession periods (Pahuriray and Cerna).

25. Integrated Hydrological-Hydraulic Models for Urban Flood Forecasting

Khan *et al.* (2022) assert that of all natural disasters, floods are the most devastating, resulting in the deaths of billions of individuals. Flooding from rivers or other natural waterways can cause significant damage to adjacent businesses and residences. Consequently, an efficient flood forecasting system is essential to mitigate losses. This system must entail the creation of an integrated hydrological and hydraulic model capable of predicting future flood events and pinpointing potential high-risk areas. A hydraulic model has been developed to forecast the frequency and severity of floods in the Tajabad Khwar waterway, adjacent to the Deans Residential Apartment in Hayatabad Phase III, due to numerous construction projects that have encroached upon the area, potentially causing catastrophic impacts on local residents. This project generates flood zone maps with return periods of 10, 20, 50, and 100 years to detect potentially susceptible areas and mitigate subsequent floods caused by encroachments. The HEC-HMS software anticipated flood discharges of 772, 1036, 1392, and 1666 m^3/sec for the 10-year, 20-year, 50-year, and 100-year flood return periods, respectively. The corresponding water surface elevations were ascertained to be 196 m, 197 m, 201 m, and 202 m with HEC-RAS. This model is an excellent starting point for generating flood zone maps for a certain time frame to assess possible high-risk areas (Mosavi et al.).

26. Flash Flood Risk Assessment and Machine Learning Applications

Flash floods are without question, among the worst natural disasters that could be experienced (He *et al.*, 2022).

Highways are vulnerable to flash floods due of their exposed linear strip structures hence damages occur. Improvement of the flash flood forecast and published warnings on transportation demand the accurate assessment of highway susceptibility to flash floods. An evaluation method of flash flood risk was formulated to defend the roadway, looking at the exposure and disaster reduction skills in this study. Afterwards, we utilized the SVM to integrate the calculated indices, and when defining the index system, we used the AHP method, with attention to the mechanisms of flash floods on the roadway. By employing SVM it is possible to eliminate the subjectivity factor in vulnerability assessment. Based on the evaluation results the percentage of very/extremely sensitive area is 4.29 % of total area of the land. 5.19 % was moderately sensitive area of the total area of the land. About fifteen point eighty-one % of the whole land was identified as consisting of regions of low and extremely low susceptibility, while forty-three point ninety-four% made up of underdeveloped areas. There was observed the greatest and essentially the critical vulnerability in the southeast suburb zones with overcrowded road networks and in the northwest suburb zones close to the great highways. Poor states with little road connection density together with the ones neighboring urban areas are characterized by low or extremely low vulnerability. There was the need to conduct a systematic and accurate assessment of road flash flood susceptibility which this study sought to do. Based on the vulnerability assessment, there are several recommendations that need to be put in practice in southeast and northwest to prevent and avoid disasters. The southern region should pay more concern on the new road constructions as well as improve on the existing infrastructure in terms of their disaster resistance; On the other hand, the north west region should also focus on new road constructions but more importantly strengthening disaster preventive and mitigations systems. Climate change has worsened floods and risks threatening Vietnam agricultural, economic, and structural units (Pham Quang and Tallam, 2022). The flooding is real in Vietnam due to a high population density, and an extensive system of rivers and canals. In the case of preliminary evacuation alerts and the use of other risk reduction measures, the initial statistical models and physical hydrology models mostly rely on conventional short-term statistical models, which are suboptimal when applied in large-scale disasters. It indicates that a more systematic and better flood prediction system is needed as the present state of affairs has arisen out of the uncoordinated approach of the Vietnamese government. Since floods are rare, local governments have to rely on their own simulations, and they can be grossly insufficient. Machine learning is particularly reliable when engaging small data sets toward equation of an accurate end result. The current paper provides a model of an artificial neural network for long-term predictions, with the comparison of other techniques of the field. The models were trained using several elements evaluated based on three criteria: climate, hydrology, and socio-economy. A model outperformed rival models by 91% accuracy along with a better long-term forecast of floods' threat. To reduce the time between the occurrence of a flood and the provision of an alarm, Jurafsky and Martin (2012) developed a technique for real-time determination of flood extent. This method employs logistic regression to produce a discriminant of flood probability for each grid based within the study regions so as to enable the forecast of the degree of flooding in relation to runoff produced by precipitation. Using Typhoon Chaba that hit the research area in 2016, a 2D flood inundation models was calibrated to develop the flood probability discriminant for each grid. After that we used historical rainfall data to produce 100 probabilistic rainfall events. In the case of hydrodynamic models, we used database modeling for rainfall, runoff, and inundation in creating the databases for each of the scenarios used. However, we combined the return period, return length, and the distribution of occurrences in time. A flood probability discriminant using logistic regression was established for each grid by comparing the runoff amount from the database to the grid status as flooded with a value of 1 or not flooded with a value of 0. The discriminant that is generated uses the amount in runoff as the input and then uses the coefficients to the assessment of the flood risk on the specified grid arriving at the best flood extent description in the shortest time. Using the proposed method, flood extent was predicted correctly within a few seconds for both the scenario rainfall and real rainfall with accuracy ranging between 83.6~98.4% and 74.4~99.1% respectively (Tanks).

27. IoT and Neural Network-Based Real-Time Flood Monitoring

Rapid snowmelt, obstructed drainage systems, intense precipitation, insufficient water pump capacity in flood-prone regions, and environmental degradation are factors that might precipitate flooding, the most prevalent natural catastrophe, as noted by Ansari and Patil (2022). Diverse types of floods may transpire. Examples of this include river flooding, storm surges, and coastal inundations. The primary aim of this paper is to elucidate the application of an Artificial Neural Network for flood prediction. The primary objective of the project is to utilize IoT for real-time monitoring of flood data and to issue alerts during hazardous conditions. A flood alarm may be issued when the water level increases. The system employs many sensors to assess the value at each stage, including those that identify the maximum water reach point, the storage tank, and the rain sensor. The sensor collects data, which is then processed by the microcontroller and transmitted to the internet via the Wi-Fi module. The system employs an Artificial Neural Network algorithm to perform the classification process following the transmission of data to a MySQL database. The experimental discoveries present data in real time to assist in flood avoidance. The accuracy of our training and validation data is ensured using artificial neural networks. Conversely, we encounter training data loss and validation data loss, respectively.

28. Machine Learning and LSTM for Large-Scale Riverine Flood Forecasting

Focusing on riverine floods in large, monitored rivers, Google's operational flood forecasting system was developed to provide accurate real-time flood alerts to agencies and the public (Nevo *et al.*, 2022). It began operations in 2018 and has subsequently expanded geographically. This forecasting system comprises subsystems for data validation, stage forecasting, flood modeling, and warning transmission. Machine learning is employed in two of the subsystems. The stage forecasting procedure is represented by linear models and long short-term memory (LSTM) networks. The thresholding model assesses the extent of inundation, whereas the manifold model evaluates both the extent and depth of inundation to ascertain flood inundation. The manifold model, proposed here, provides an alternative to hydraulic flood inundation models utilizing machine learning. Upon evaluation with historical data, each model satisfies the performance criteria required for practical implementations. Regarding inundation extent modeling, the LSTM surpassed the linear model, although the thresholding and manifold models exhibited comparable performance. The flood warning system in Bangladesh and India operated throughout the 2021 monsoon season. It encompassed flood-prone regions next to rivers, covering almost 470,000 km^2 and inhabited by over 350,000 individuals. More than 100,000 flood alerts were sent to those in most need, as well as to relevant authorities and emergency organizations. Enhancing the system's modeling capabilities and precision, along with broadening coverage to more flood-prone regions, are current and forthcoming objectives (Chen et al.).

29. Flood Susceptibility Mapping Using ML Models

El-Haddad *et al.* (2021) assert that assessing flood susceptibility is crucial for environmental and water management in various countries to mitigate the impacts and damages resulting from floods. This project seeks to develop flood susceptibility maps using four data mining and machine learning models: boosted regression tree (BRT), functional data analysis (FDA), general linear model (GLM), and multivariate discriminant analysis (MDA). This investigation was conducted in the Egyptian Wadi Qena Basin. High-resolution satellite imagery datasets (Astro digital and Sentinel-2 captured post-flood events), historical documents, and comprehensive fieldwork were utilized to locate and delineate flood-inundated regions. Two groups were established using ArcGIS 10.5: one for training, including 239 flood sites (70% of the total), and one for validation, consisting of 103 flood locations (30% of the total). A total of 342 inundated locations were mapped. Nine types of flood-influencing factors were established: topographic wetness index, land use/land cover, slope angle, slope length, altitude, distance from main wadis, lithological units, curvature, and slope aspect. Utilizing the aforementioned models (BRT, FDA, GLM, and MDA), we evaluated the relationships between the flood inventory map and the factors that affect floods. Flood locations, excluded from the model creation process, were validated against the findings. Utilizing receiver operating characteristics (ROC) and area under the curve (AUC) for the success rates in training data and prediction rates in validation data enabled us to assess the models' accuracy. The four models evaluated (BRT, FDA, GLM, and MDA) exhibited area under the curve (AUC) values of 0.783, 0.958, 0.816, and 0.821 for success and prediction rates, and 0.812, 0.856, 0.862, and 0.769 for MDA and GLM, respectively. Subsequently, flood risk maps were classified into five categories: extremely low, low, moderate, high, and very high. The results indicated that the BRT, FDA, GLM, and MDA models all yield commendable outcomes in flood susceptibility mapping. (Chen et al.)

The regional development projects like selecting new urban zones and placing infrastructure will benefit greatly from using the created susceptibility maps. Ighile *et al.* (2022) highlights that floods rank as the most destructive natural disasters. A range of techniques has been developed to help prevent flood damage to people and nature. We must detect flooding zones to improve our emergency action plans given growing flood frequency and environmentally social changes. The research team examined flood history from 1985 to 2020 to predict which areas in Nigeria face the highest flood risks using various conditioning factors. We trained and evaluated logistic regression plus artificial neural network models to create a flood susceptibility map. Our study used fifteen data variables that combine information about weather patterns land features and property locations. We tested model accuracy using two measurement methods called area under the curve (AUC) and receiver operating characteristic curve (ROC). Our results show that both methods precisely identify and forecast flood prone areas. The ANN model showed better performance than LR by successfully predicting floods at 76.4% and making accurate predictions at 62.5%. Local flood risks are highest in areas close to water in the south mainly due to the predictions of both model simulations. Research shows machine learning can effectively pinpoint flood-prone areas ahead of time to help authorities develop solutions that reduce the flood risk. Sun *et al.* (2023) report that HEC-RAS and MIKE 11 hydrodynamic models prove effective at simulating river flows through solving fluid motion equations which flood forecasters extensively employ. By using the same methods Sindh monitors Indus River levels which respond quickly to increased precipitation and snowmelt upstream. These models deliver precise predictions for future floods but their high complexity makes them difficult to use in real-time flood monitoring. Since urban areas need fast flood response measures hydrodynamic models struggle to work well in towns despite accurately showing floodplain and river behavior. According to Moon *et al.* (2023) climate change now makes floods both more frequent and more dangerous. Leaders must plan flood protection efforts based on major climate changes to stop disasters. Researchers show that LSTM models can forecast rainfall-induced runoff but need more observation data for full development. The researchers developed a flood forecasting solution by linking deep learning and rainfall-runoff methods. We built a method to determine future flood risks when studying multiple rainfall measurements. The study used four main elements to create simulated rainfall events in urban river basins - Total rainfall amount, event length, rainfall timing patterns. The rainfall-runoff model delivered multiple simulated results to supply data needed for forecasting flood timings under different rainfall conditions. Our predictions achieved results above 80% in Nash Sutcliffe efficiency and 90% in correlation coefficient values. The research shows our forecasting method will correctly predict flood timing and areas only based on rainfall data. Studies by Letessier *et al.* (2023) show that spring floods drive the most flooding in Ontario Canada so temperature must be an essential input for machine learning models that predict floods. This technological development helps forecast floods better through its identification of entire processes including snowmelt effects from heatwaves. The work presents a new machine learning method called Adaptive Structure of the Group Method of Data Handling (ASGMDH). This process predicts river discharge on daily basis using present weather measurements while adding daily discharge data to create an historical representation of watershed characteristics. We studied four distinct cases using different input information to develop our complete ML system. Our simplest model built from three parameters produced very reliable forecasts in testing (R2 value of 0.992) and training (0.985) which indicates its practical usage. The study area shows that our ASGMDH model remains accurate within 15% relative error of multiple test results. The final model network with ASGMDH uses only one second-order function with AICc 19648.71 while the basic GMDH model runs seven functions and reaches AICc 19701.56. Our sensitivity analysis shows maximum temperatures have the strongest effect on our estimated daily river flow predictions. According to Antwi-Agyakwa *et al.* (2023) researchers now make better flood predictions by locating areas likely to flood helping policymakers meet their targets. Delivering safe flood areas depends heavily on how we predict flood patterns today. Many research teams analyze single flood prediction models. We explore the complete history of flood prediction science while analyzing modern trends and benefits and challenges of each technology plus identifying critical research gaps. Our research team used Biblioshiny to analyze 1,101 publications about Scopus in the years 2017 through 2022. The research presented newly found answers along with developing scientific fields and unexplored portions in knowledge. Since machine learning models handle our flood prediction system today we need to research Copula and Bayesian Network methods to better predict unstable flood risks. Cloud technology capabilities now support full integration of all data and equipment needed for flood prediction. Australia South Africa and Africa need further flood forecasting research because the regions lack substantial existing studies into probability models. Floods represent natural disasters that damage humans and their possessions plus destroy built infrastructure (Kumar *et al.*, 2023). To protect communities from flood threat accurate forecasting with effective flood handling is essential. Deep learning technology helps us understand big data better and make better flood prediction decisions. This research presents a detailed evaluation of what deep learning can deliver for flood monitoring and management today. This research project covers data selection methods, deep learning algorithms plus accuracy assessments. This research explains all present methods showing both pros and cons of each system. This research explores data availability issues and deep learning tool interpretation problems in flood detection work along with ethics in flood prediction. The research document shows how deep learning technology can boost flood prediction accuracy and management capability. Flood research advances involve clearer deep learning model explanations as well as bringing uncertainty measures into forecasts. It also mixes various data types while merging deep learning with other prediction approaches. Research goals help improve flood detection methods and forecasts which create better deep learning technology performance. Scholars and flood management experts will benefit greatly from using this research. It examines modern developments while showing issues and suggesting research directions for the future. Advanced deep learning tools can help communities prepare better for and reduce flood effects to protect people and structures. Ganjirad and Delavar classify floods as natural disasters that cause damage worldwide each year both to people and property. Making flood risk maps and identifying flood prone areas helps decrease damage from disasters. We used Google Earth Engine (GEE) cloud platform to extract flood data for the Calcasieu Parish study area in Louisiana, USA. Information layers were generated for the following variables: Site elevation serves with land use activity maps and river measurement tools alongside precipitation levels NDVI and NDBI for this analysis. Our research examines how Random Forest (RF) and Support Vector Machine (SVM) work to create flood risk maps after extracting flood extents with Google Earth Engine (GEE). The RF prediction model showed 115% accuracy found through MAE measurements ranked the SVM model at 0.097% accuracy. The results in QGIS show that the RF model achieves superior outcomes to the SVM model. Elevation stands out as the key factor impacting flood risk in the RF model based on feature importance analysis with a relevance score of 0.197 followed by slope at 0.135 and plan curvature at 0.123. Floods stand as one of nature's worst disasters as they hurt both human lives and property globally (Wu et al. 2024). This study evaluated the spatial pattern of flood sensitivity using three machine learning models: The team chose Logistic Regression (LR), XGBoost, and Random Forest (RF) free of charge to compare models and determine the best fit for flood sensitivity analysis.

Using random sampling a dataset of historical flood locations was produced and the research location chosen was Suqian City in Jiangsu Province. The flood sensitivity assessment considered fifteen unique hydrological, geographical, and meteorological components; twelve variables were selected following a multicollinearity analysis. The RF technique, recognized for its reliability and efficiency in flood risk assessment, surpassed the other selected ML models for accuracy, thorough evaluation impact, and area under the curve (AUC). The principal outcome of this research is a flood sensitivity map that categorizes regions based on their susceptibility levels, ranging from extremely low to highly high. Approximately 44% of the research area, predominantly in the central, eastern, and southern sections of the historic city zone, is classified as high-risk according to the RF model, which is the most precise iteration of the model.(Rajab et al.)

To effectively create a flash flood risk map and comprehend the flash flood mechanism, Tan *et al.* (2024) introduced a methodological framework termed RF-FS, which integrates random forest with genetic algorithms, recursive feature removal, and shadow variable search. This approach seeks to identify the principal driving factors (PDFs) of flash floods. Recent hydrological insights for the region: The RF-FS methodological framework revealed several key distribution functions, primarily meteorological (maximum 1-day precipitation, total daily precipitation surpassing the 95th percentile threshold, and maximum 5-day consecutive precipitation) and hydrological (topographic wetness index, stream power index). The best KDFs for flash flood risk prediction were then determined utilizing the RF-FS method framework, which integrates a random forest with a genetic algorithm. Moreover, the bulk of high-risk areas, encompassing around 45% of the mountainous regions in Guangdong Province, are located in the province's central, eastern, and southern mountainous zones, as shown by the modeling results. Prioritizing flash flood prevention and early warning systems in areas prone to high precipitation is justified, as these regions are most susceptible to flash floods.

Floods are one of the most destructive natural disasters, and the difficulties that they present are the main methods that make it challenging to develop efficient methods for generating forecasting models of the event (Wajid *et al.*, 2024). Some of the consequences include crop failure, drought, displacement of people, destruction of infrastructure, and interruption of certain services due to this problem. Sophisticated studies of flood prediction models have greatly improved the policy directorate, disaster management, avoidance of human loss of life and minimization of flood losses to property. This paper introduces an IoT solution for the efficient prediction and forecasting of floods. Given the limited capability of IoTs sensor node such as battery power and size of RAM, we implement the energy conservation fog-layer strategy where by data diversity is used to cut down the power consumption. Further, there is a way of data preservation through the practice of cloud computing. Where the meteorological and hydrological factors involved: temperature, humidity, and precipitation, as well as flow and elevation, we are able to shift calibration levels appropriately for flood phases. Artificial neural networks are finding wide application in the design of prognostic systems because of their effectiveness and capability in the modeling of the actual calculations involved in actual physical sciences. Our approach compares Decision Trees and Logistic Regression with ANNs to confirm the most efficient approach for flood prediction. After having provided a characterization of the features of the graph and having analysed the categorization report, other accuracy measures are derived. In fact, our proposed solution involves performing analysis on the given dataset to develop a predictive model under the ANN approach. This model can produce real time flood risk predictions through an easy Graphical User Interface. As depicted by Krichen *et al.* (2024), Natural disasters (NDs) for a long time have been a major concern to people and physical built environment. The emergence of more frequent disasters and those of greater magnitude in the recent past, naturally calls for better and enhanced disaster management systems. Of the above-listed alternatives, there is much more potential in technical improvements. This review article discusses on the extent that many natural disasters have adopted new technologies in order to minimize their effects. In this paper, the inventory of technology in the management of natural disasters (NDs) such as the IoTi cellphone social sites radar satellite remote sen cookies are presented. Possibly these technologies can improve ND prediction and response and quickly recover clearly reducing infrastructure losses and lives. This paper analyzes advantages, disadvantages and issues related to application of those technologies for natural disasters management. However, the following are some hurdles that need to be met in order to reap the benefits of technology: The implementation costs of such technology is exorbitant and; The process requires professional, specialized skills. In conclusion, this survey paper explains the enormous relevance of technology in NDM and gives a thorough review of the use of technology in ND management. This investigation seeks to identifying how different technologies could be applied to create better approaches to catastrophes management and mitigation. Hazard such as floods are extremely difficult to prepare for as they are natural occurrence that cannot be predicted. Skills in predicting floods has improved because models used to identify its occurrence have revealed lesser harm to property, reduced deaths, decreased in policy hints and a reduced risk level. Over the last two decades, the ML methods have improved in gradually developing the more powerful and costeffective solutions for the improvement of the predictive systems, in accordance with the modelling of intricate mathematical assumptions associated with the floods. This realization explained why a considerable contingent of hydrologists started using machine learning as early as possible given the numerous positive aspects and possibilities it offers. This is because innovative and efficient predictive models require new and advanced frameworks of machine learning integrated by the existing ones. This research mainly adds value by showing the current level of accuracy of the machine learning models in flood forecasting and which of the models are more effective.

30. Artificial Intelligence and Machine Learning in Disaster Management

It provides a complete overview of several ML algorithms by examining literature that benchmarks ML models based on robustness, accuracy, efficacy, and speed through a qualitative evaluation. Through the analysis of extensive databases, identification of patterns, and forecasting of impending disasters, artificial intelligence (AI) holds significant potential for enhancing natural disaster management (Albahri *et al.*, 2024). To address the substantial challenges posed by natural disasters, these models facilitate the implementation of preventive measures such as early warning systems (EWSs), evacuation strategies, and resource distribution. This study thoroughly examines trustworthy AI applications in the context of natural disasters, encompassing risk assessment, disaster prediction, and disaster management. Several leading academic databases including Web of Science (WoS), Science Direct (SD), Scopus, and IEEE Xplore (IEEE) served as the foundation for this research. Our analysis collected 981 scientific papers up to February 2024 through three separate search processes. Our analysis used 108 articles after careful selection including duplicate removal and standards for inclusion and rejection. The research systematizes current artificial intelligence applications in disasters and studies their drivers and barriers while translating these insights into solution recommendations for future development. The work covers modern methods and systems in disaster management using ML, data fusion, fuzzy logic, XAI, and MCDM while both deep learning and ML strategies. This research outlines seven AI challenges in natural disaster management and suggests dependable solutions based on learned principles to guide future reliability work. Autoained intelligence holds great promise to upgrade natural disaster control systems but requires us to solve important implementation issues This research develops an advanced XAI approach to study natural disaster decision making and documents unused and popularized areas in AI theories about disasters. It transmits vital disaster-related data and develops all available deep learning and machine learning algorithms. The work considered every aspect of AI when used for handling natural catastrophes including ethical problems and technical biases.(Rajab et al.)

CHAPTER THREE

DATA ACQUISITION FOR DISASTER FORECASTING AND MANAGEMENT USING MACHINE LEARNING

Fatima Abbas
Riphah International University Faisalabad

*Corresponding Author: fatima.abbas@riphahfsd.edu.pk

The machine learning models are trained to forecast floods in Pakistan's four provinces. This chapter shows how dataset was acquired and preprocessed for Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan. After properly preparing the data, four different machine learning models were trained to predict floods between 2025 and 2030. This project was carried out in Google Colab IDE using various Python libraries (Zaidi et al.).

1. Data Acquisition

The dataset used in this study was acquired from GitHub, containing flood data from 2000 to 2021. Key environmental and climatic attributes included:

- **Rain (mm/month):** Amount of monthly rainfall.
- **Veg:** NDVI (Normalized Difference Vegetation Index) indicating vegetation cover and absorption capacity.
- **Temp:** Temperature changes affecting water evaporation and precipitation cycles.
- **ICE:** Individual Conditional Expectation showing model prediction changes with one feature.
- **Soil Moisture**: Represents the water content in soil layers, influencing runoff and flood risk. Wind Speed (km/h): Impacts rainfall distribution and surface water movement.
- **Humidity (%)**: Affects evaporation rates and precipitation patterns.
- **River Discharge (m³/s)**: Measures flow in major rivers, critical for flood forecasting.

Data preprocessing included handling missing values, normalization of numeric features, and encoding of categorical variables. Time-series segmentation was performed to align environmental indicators with flood events, ensuring temporal consistency. The final datasets for each province were split into training (70%) and testing (30%) sets to validate the machine learning models.

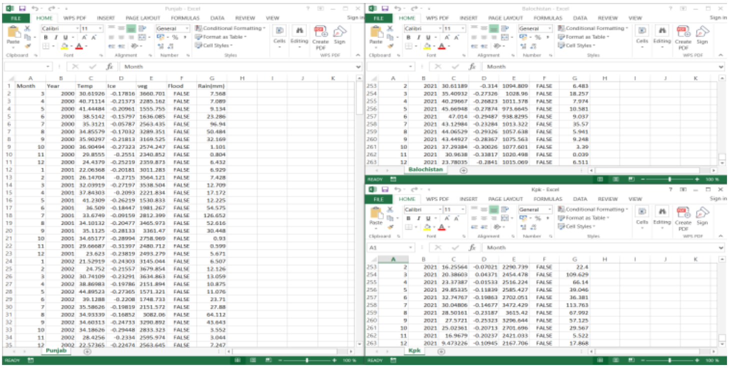


***Figure** Error! No text of specified style in document..**1:** A glimpse of dataset presented in a workbook (CSV file)*

2. Tools and Environment

The project was executed using Google Colab, a cloud-based platform supporting Python programming and providing free access to GPUs and TPUs for faster computations. The libraries used include:

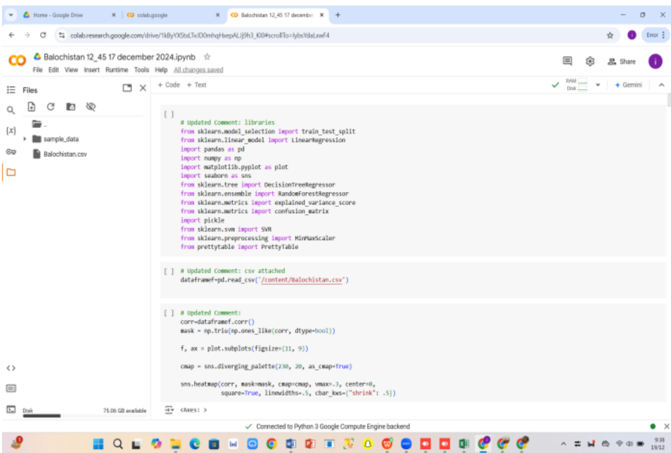


Figure 3.2: *Interface of the Google Colab*

2.1. Pandas

Pandas is a powerful Python library used for data manipulation and analysis, providing two main data structures: Data Frame and Series. A Data Frame uses 2D tables to hold your data and Series creates a one-dimensional array. Pandas helps data scientists prepare their data by handling multiple formats through simple data cleaning methods. Data scientists depend heavily on Pandas because it allows them to arrange data in rows and columns plus handles database operations for statistical analysis.(Zaheen et al.)

2.2. NumPy

NumPy functions as the central library in Python when performing numerical tasks by handling both arrays and matrices plus large multi-dimensional data structures. NumPy supplies a top data system called ndarray that helps you store and work with numerical data effectively. Despite its advanced capabilities NumPy provides numerous mathematical functions to process linear algebra, statistical methods and random number generation. By processing vectors NumPy operations complete tasks much faster than normal Python list handling does. Researchers and developers depend on NumPy for their scientific projects alongside Pandas and SciPy because this package optimizes their data analysis and machine learning work.(Sidrane et al.)

2.3. Matplotlib and Seaborn

Matplotlib and Seaborn are leading tools used for the data visualization. Through matplotlib you can create various plot types including line charts, scatter plots, and histograms at a basic library level while adjusting axis elements and coloring schemes. This tool makes precise visualizations that match scientific publication requirements. In addition to basic plotting Seaborn provides a top-level interface to generate beautiful statistical plots such as heat maps violin and pair plots. Seaborn eases data analysis tasks by default presenting clear visuals and perfecting plot elements for quick research purposes. Users commonly pair Seaborn with Matplotlib as Seaborn produces simple stylish plots and Matplotlib creates complex personalized charts.(Baloch et al.)

2.4. Scikit-learn

The machine learning Python library Scikit-learn allows users to find valuable data mining details quickly while processing data. Through its algorithms Scikit-learn helps you solve multiple machine learning problems while giving you essential features to handle experiments. The developer community designed Scikit-learn to work with NumPy, SciPy, and Matplotlib as its base. Users appreciate its easy learning path and stable API structure. Data scientists and machine learning specialists use Scikit-learn because it helps them create and assess predictive models within their standard machine learning processes.(Bibbò et al.)

2.5. TensorFlow/Keras

TensorFlow Google made TensorFlow to be an open-source deep learning platform for machine learning development. As a machine learning platform it helps users build train and deploy neural networks and other machine learning systems. The independent neural networks API Keras has fully integrated into TensorFlow as its official high-level API which lets users create advanced deep learning models using simple and direct code. These two tools combine to create a leading AI solution that works well for research needs as well as large-scale production workloads across different learning tasks including image recognition and NLP.

2.6. Data Splitting and Model Validation

After data preprocessing, the dataset was split into training, validation, and testing sets to ensure robust model evaluation. Typically, 70% of the data was allocated for training, 15% for validation, and 15% for testing. The training set was used to teach the models underlying patterns in flood occurrences, the validation set tuned hyperparameters, and the testing set assessed final model performance. Cross-validation techniques such as k-fold cross-validation were applied to reduce overfitting and improve the generalizability of the models.

2.7. Feature Scaling and Normalization

Machine learning models, especially neural networks, benefit from input features scaled to a similar range. Min-Max normalization and Standardization (Z-score scaling) were applied to ensure consistent data magnitude across features like rainfall, temperature, and river discharge. Scaling improved model convergence speed and prediction accuracy by preventing bias toward features with larger numeric ranges.

2.8. Handling Imbalanced Data

Flood datasets often have imbalanced classes, with fewer flood events compared to normal conditions. Techniques such as Synthetic Minority Oversampling Technique (SMOTE) and class weighting were employed to address this imbalance. These methods increased model sensitivity to rare flood events, improving prediction reliability and minimizing false negatives.

2.9. Integration of Libraries for Workflow

The combination of Pandas, NumPy, Matplotlib/Seaborn, Scikit-learn, and TensorFlow/Keras created a cohesive workflow:

- Pandas and NumPy handled data ingestion, cleaning, and manipulation.
- Matplotlib and Seaborn provided visual insights to detect trends and anomalies.
- Scikit-learn facilitated traditional machine learning model training and evaluation.
- TensorFlow/Keras supported deep learning approaches for complex pattern recognition in multi-dimensional datasets.
- This integrated ecosystem enabled efficient preprocessing, exploratory analysis, and predictive

modeling tailored to flood forecasting in Pakistan's diverse climatic regions.

3. Methodology

Extreme flood events represent Pakistan's most common natural hazards and endanger people's physical assets and living spaces. Downstream flood predictions need better technology according to the research findings. During the last 20 years machine learning helped create better prediction systems through both more precise results and lower expenses. Pakistan's four provincial regions including Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan show that using ML flood prediction models helps solve their environmental concerns. Our research used flood data from the year 2000 to 2021 which we retrieved from GitHub to predict flood risks from 2025-2030 while testing the performance of Decision Tree, Random Forest, Linear Regression, and SVM models. Our models worked better when we used specific measurements including rain amounts temperature readings and plant growth. Our process of preparing data involved cleaning it and fixing missing values before extracting features and changing data categories to numbers to reduce signal interference. Random Forest showed better performance because its ensemble learning decreased the risk of model overfitting. Decision Tree demonstrated good results but remained too dependent on random errors. The Linear Regression model studied how severe a flood would be while Support Vector Machines delivered best results in flood versus no-flood classifying. By using both flood classification and regression systems we developed comprehensive flood forecasting methods. Our assessment used accuracy precision and F1-score to validate the prediction models plus mean squared error and recall tests. Our work improved results through hyper parameter changes plus it helped us choose the right model structure and optimize the features for better predictions. Our predictions helped us identify specific areas at risk so we could organize better disaster responses and distribute available resources to those areas. The research examined what actions should be taken based on recognized flood patterns. By performing calculations on Google Colab and importing libraries from Scikit-learn, Pandas, and NumPy our analysis became simpler. The research found problems with incorrect data and sudden environmental fluctuations. The research showed how machine learning helps detect floods better and helps communities prepare for disasters. The study helps authorities target safety planning and create better flood management policies through its results identifying vulnerable regions of Pakistan. Using flood data from 2000 to 2021 the researchers predicted flood vulnerability for the years 2025 to 2030 (Zehra).

3.1. Data Preprocessing and Cleaning

Data preprocessing was essential to ensure that the dataset was clean, complete, and ready for machine learning modeling. Several key steps were followed:

- **Handling Missing Values:** Missing data filled using mean or median imputation; columns with >30% missing data were dropped.
- **Feature Selection:** Correlation analysis used to select significant features.
- **Data Normalization:** Features scaled using MinMaxScaler.
- **Encoding Categorical Variables:** One-hot encoding applied to categorical features.

3.2. Data Splitting

Data was split into two subsets:

- **Training Set (80%)**: This set was used for training the models.
- **Testing Set (20%)**: This set was reserved for evaluating model performance.

The data split was performed using Scikit-learn's train_test_split () function, ensuring random distribution and avoiding biased evaluations.

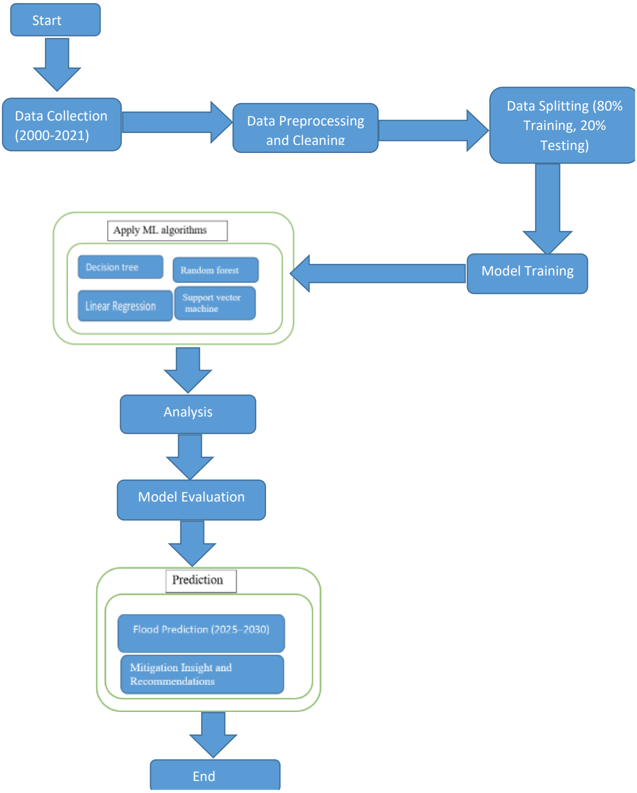


***Figure** Error! No text of specified style in document..**2:** Flowchart representing the research process*

3.3. Model Training

- **Google Colab:** As discussed earlier is section 3.2, Google Colab IDE was used in this study because Google Inc. offers this platform which includes cloud computing and Python machine learning support. The service optimizes data handling and model testing alongside performance metrics because users can perform processing tasks directly on cloud-based hardware resources.

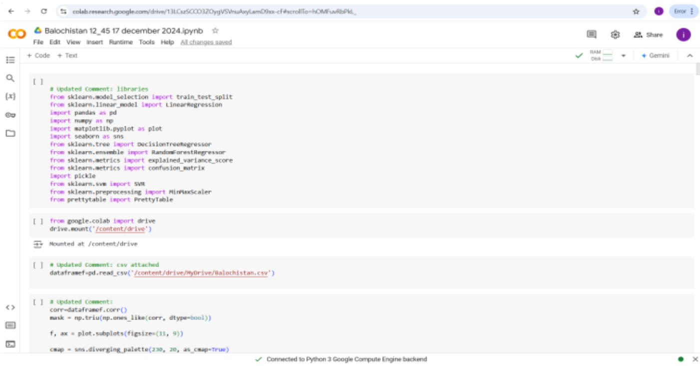


***Figure** Error! No text of specified style in document..**3:** Interface of the coding environment of Google Colab*

- **Loading the Dataset**: The analysis commenced with the acquisition of Pakistan flood data spanning from 2000 to 2021, obtained in its original format from GitHub. The dataset was imported into Google Colab, as shown in Figure Error! *No text of specified style in document..**4***, and subsequently processed using Pandas, where it was transformed into a structured DataFrame to facilitate comprehensive analysis. Given the objective of flood prediction, key hydrometeorological variables – including rainfall, vegetation (NDVI), temperature changes, and ICE (Individual Conditional Expectation) – were identified as critical factors influencing flood patterns. Following data ingestion, a preliminary assessment was conducted to evaluate data integrity, completeness, and consistency, ensuring its suitability for subsequent model development and predictive analysis (Prediction).

1	Month	Year	Temp	Ice	veg	Flood	Rain(mm)
2	3	2000	35.44870967741936	-0.19958064516129034	828.3634999999999	False	9.458000000000002
3	4	2000	46.82714285714286	-0.217	801.5055	False	3.5980000000000008
4	5	2000	50.797387096774195	-0.22703225806451613	778.6479999999999	False	4.848999999999998
5	6	2000	47.59846666666667	-0.21713333333333334	732.5155	False	3.2279999999999998
6	7	2000	45.717290322580645	-0.20296774193548386	773.4665	False	16.286
7	8	2000	44.79063157894737	-0.1841578947368421	826.3295	False	7.281000000000001
8	9	2000	45.69736666666667	-0.23023333333333335	859.0035	False	2.1300000000000003
9	10	2000	42.94054838709677	-0.25641935483870965	868.9705	False	0.069
10	11	2000	32.39423333333333	-0.2693333333333333	850.58	False	0.621
11	12	2000	26.504032258064516	-0.2751935483870968	890.5975000000001	False	9.482000000000003
12	1	2001	24.938870967741938	-0.30451612903225805	891.554	False	8.453999999999999
13	2	2001	28.63439285714286	-0.2372857142857143	883.232	False	17.606
14	3	2001	36.60890322580645	-0.2503870967741935	860.3389999999999	False	13.648
15	4	2001	43.748066666666666	-0.252	826.394	False	19.93
16	5	2001	49.15429032258064	-0.27187096774193553	781.4685	False	3.553000000000001
17	6	2001	46.0888	-0.23893333333333333	721.923	False	8.567999999999998

***Figure** Error! No text of specified style in document..**4:** A glimpse of the dataset loaded in Google Colab*

- **Data Cleaning of Each Province**: Errors were removed from the dataset through a comprehensive data-cleaning process, ensuring the accuracy and reliability of flood predictions. The dataset comprised historical flood records from all four provinces of Pakistan – Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan. Various inconsistencies, such as missing values, duplicate entries, and outliers, were identified and addressed using systematic preprocessing techniques. This refinement improved data integrity, allowing machine learning models to achieve higher predictive accuracy.

A) **Handling Missing Values**

The isnull function efficiently identified missing values across various provincial sections of the dataset. When multiple cells within a specific province exhibited missing data, the dataset underwent careful preprocessing. Depending on the extent and pattern of the missing values, either the entire row was removed to maintain data integrity, or appropriate imputation techniques were applied. Statistical measures such as the mean, median, or mode (most frequent value) were utilized to fill the missing fields, ensuring consistency and minimizing potential biases in subsequent analyses. This approach preserved the dataset's reliability while optimizing it for further statistical modeling and interpretation.

B) **Standardizing Formats**

All data provinces adhered to uniform data management standards. A standardized format for displaying dates and times was consistently implemented across all regions. Additionally, rigorous quality control measures ensured the accuracy of recorded measurements related to rainfall, temperature, and plant growth. These protocols minimized inconsistencies, facilitated analyses, and enhanced the reliability of this study.

C) **Outlier Detection**

Statistical tools, including the interquartile range (IQR), identified unusual values in temperature and rainfall measurement data. These anomalies were either removed or adjusted to align with the specific data processing requirements of each province. The refinement process ensured the accuracy and reliability of climate datasets, minimizing the impact of outliers on subsequent analyses. By standardizing data across regions, the methodological approach enhanced consistency, facilitating more precise climate modeling and decision-making.

D) Feature Engineering

Monthly rainfall data was incorporated into each region's dataset to enhance flood prediction accuracy. This additional feature allowed for a more comprehensive analysis of precipitation patterns and their correlation with flood occurrences. By integrating historical rainfall trends, the dataset provided a more robust foundation for predictive modeling, improving the reliability of flood risk assessments. Furthermore, the inclusion of this data enabled the identification of seasonal variations and potential anomalies, contributing to a more refined understanding of hydrological conditions across different regions.

- **Province-Specific Adjustments**: Data cleaning strategies were tailored to align with the unique reporting structures and climatic characteristics of each province. Rainfall gaps in Sindh were addressed by analyzing regional precipitation trends, ensuring a realistic estimation of missing values. Temperature readings in Khyber Pakhtunkhwa (KPK) were adjusted to account for altitude-induced variations, enhancing the accuracy of recorded data. Following the completion of these corrections, provincial datasets underwent rigorous preprocessing to ensure compatibility with machine learning models.

index	Month	Year	Temp	Ice	veg	Flood	Rain(mm)
2	3	2000	30.619258	-0.178161	3660.7005	False	7.568
3	4	2000	40.711143	-0.213733	2285.1620	False	7.089
4	5	2000	41.444839	-0.209613	1555.7545	False	9.134
5	6	2000	38.514200	-0.157967	1636.0845	False	23.286
6	7	2000	35.312097	-0.057871	2563.4350	False	96.940
...	...	...	...	...	...	...	...
259	8	2021	34.020774	-0.245355	3924.0810	False	63.933
260	9	2021	33.836000	-0.240433	3932.6785	False	88.586
261	10	2021	32.756806	-0.336387	3292.4760	False	12.031
262	11	2021	28.087333	-0.351567	2718.6610	False	0.206
263	12	2021	20.221677	-0.302645	3372.6000	False	4.978

262 rows × 7 columns

```
dfx=dfx.sort_index()
```

```
dfx
```

index	Month	Year	Temp	Ice	veg	Flood
1	2	2000	17.041333	0.033750	1864.4820	False
2	3	2000	18.878581	0.087226	1976.2355	False
3	4	2000	29.873643	-0.013967	2041.7275	False
4	5	2000	33.941710	-0.092097	2053.4755	False
5	6	2000	32.489467	-0.123967	2312.7570	False
...	...	...	...	...	...	...
259	8	2021	28.501613	-0.231871	3615.4200	False
260	9	2021	27.572100	-0.253233	3296.6440	False
261	10	2021	25.023613	-0.207129	2701.6960	False
262	11	2021	16.967900	-0.202367	2421.0330	False
263	12	2021	9.473226	-0.109452	2167.7060	False

263 rows × 6 columns

```
dfx.reset_index()
```

	Months	Years	Flood
0	Janruary	2000	NaN
1	February	2000	NaN
2	March	2000	NaN
3	April	2000	NaN
4	May	2000	NaN
...	...	...	...
259	August	2021	NaN
260	September	2021	NaN
261	October	2021	NaN
262	November	2021	NaN
263	December	2021	NaN

264 rows × 3 columns

```
Ndvi_monthly =  df3.groupby('date',as_index=False)['NDVI'].mean()
Ndvi_monthly.tail(50)
```

	date	NDVI
213	Nov 2016	2330.9510
214	Nov 2017	2242.7655
215	Nov 2018	2301.3300
216	Nov 2019	2552.8830
217	Nov 2020	2256.1060
218	Nov 2021	2421.0330
219	Oct 2000	2233.5665
220	Oct 2001	2138.3270
221	Oct 2002	2387.2310
222	Oct 2003	2421.4550
223	Oct 2004	2160.0490
224	Oct 2005	2440.0230
225	Oct 2006	2627.6930
226	Oct 2007	2508.2560
227	Oct 2008	2463.3970
228	Oct 2009	2404.2060
229	Oct 2010	2803.7820

```
df3[["month","day", "year"]] = df3["system:time_start"].str.split(" ", expand =
df3['date']=df3["month"]+' '+df3["year"]
df3
```

	system:time_start	NDVI	month	day	year	date
0	Feb 18 2000	1864.482	Feb	18	2000	Feb 2000
1	Mar 5 2000	1950.279	Mar	5	2000	Mar 2000
2	Mar 21 2000	2002.192	Mar	21	2000	Mar 2000
3	Apr 6 2000	2043.985	Apr	6	2000	Apr 2000
4	Apr 22 2000	2039.470	Apr	22	2000	Apr 2000
...	...	...	...	...	...	...
248	Oct 16 2021	2701.696	Oct	16	2021	Oct 2021
249	Nov 1 2021	2462.705	Nov	1	2021	Nov 2021
250	Nov 17 2021	2379.361	Nov	17	2021	Nov 2021
251	Dec 3 2021	2243.921	Dec	3	2021	Dec 2021
252	Dec 19 2021	2091.491	Dec	19	2021	Dec 2021

503 rows × 6 columns

Figure *Error! No text of specified style in document.**.6:** A glimpse of the cleaned dataset*

- **Data Visualization**: Visual representations played a crucial role in identifying key patterns within datasets of varying sizes. In this study, a range of visualization techniques facilitated the analysis of relationships between different features and the target variable, flood occurrence, across the four provinces. The dataset encompassed information from Punjab, Sindh, Khyber Pakhtunkhwa (KPK), and Balochistan, allowing for a comprehensive examination of spatial and temporal trends. Graphical methods such as histograms, scatter plots, and heatmaps provided insights into correlations, distribution patterns, and potential influencing factors. These visualizations not only enhanced interpretability but also supported data-driven decision-making by highlighting regions with heightened flood risk.(Cui et al.)

- **Classifiers**: Four models were implemented: Four machine learning models were implemented to predict flood occurrences:

A) Decision Tree Classifier

The Decision Tree Classifier the Decision Tree Classifier in Scikit-learn served as our implementation platform. This algorithm divides data using threshold values on features to create a tree format decision structure. To prevent overfitting, we set a maximum tree depth limit then used Gini impurity to enhance the model's results. Ion nodes, forming a tree-like structure.

The maximum tree depth was adjusted to prevent overfitting, and the model was trained using the Gini impurity criterion for optimal performance.

B) Random Forest Classifier

The Random Forest Classifier was implemented using Scikit-learn's RandomForestClassifier ().

This ensemble method combined multiple decision trees to improve prediction accuracy and reduce overfitting.

The number of trees in the forest (n_estimators) was set to 100, with other hyper parameters fine-tuned using grid search.

C) **Linear Regression**

Linear Regression was implemented using Scikit-learn's Linear Regression ().

This algorithm modeled the relationship between the input features (e.g., rainfall, temperature) and the flood occurrence indicator.

The model's coefficients and intercept were computed to understand the relative impact of different features.

D) **Support Vector Machine (SVM)**

SVM was implemented using Scikit-learn's SVC () with a linear kernel.

This algorithm classified flood occurrences by creating a hyperplane in a multi-dimensional feature space.

The penalty parameter C and kernel type were adjusted to enhance the model's predictive performance.

4. Model Evaluation

The models were evaluated using standard metrics to assess their performance on the testing set. The following evaluation metrics were computed:

The overall correctness of predictions, in terms of accuracy (Handelman *et al.*, 2019; Naidu *et al.*, 2023), which was calculated as:

$$Accuracy = \frac{Correct\ Predictions}{Number\ of\ Total\ Predictions}$$

Precision (Handelman *et al.*, 2019; Naidu *et al.*, 2023) measured how many predicted flood events were actual floods and is given by following equation:

$$Precision = \frac{True\ positive + False\ Positives}{True\ Positives}$$

Recall (Handelman *et al.*, 2019; Naidu *et al.*, 2023) evaluated how well the models identified actual flood occurrences and is represented by the following mathematical formula:

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

The F1-score balanced precision and recall (Handelman *et al.*, 2019; Naidu *et al.*, 2023), calculated as:

$$F1 = 2 \cdot \frac{Precision \times Recall}{Precision + Recall}$$

MSE was computed for the linear regression model (Handelman *et al.*, 2019; Naidu *et al.*, 2023), quantifying prediction errors:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

5. Flood Prediction (2025-2030)

The trained models generated flood predictions for the years 2025 to 2030. Predictions included:

- **Flood Probability**: Likelihood of floods in each province.
- **High-Risk Areas**: Provinces with higher predicted flood probabilities.
- **Temporal Patterns:** Monthly and seasonal trends were analyzed to identify periods of higher flood risk, allowing for early warning and resource allocation. Punjab and Sindh exhibited peak flood probabilities during the monsoon months (July–September), while KPK showed increased risk in late spring and early summer due to snowmelt combined with rainfall.
- **Severity Estimation:** The models provided predicted flood magnitudes by correlating rainfall, river discharge, and soil moisture levels. These estimations allowed classification of flood events into low,

moderate, and high severity categories, supporting targeted disaster management strategies.

- **Uncertainty Analysis:** Confidence intervals were calculated for all predictions to quantify uncertainty arising from model assumptions and input variability. Monte Carlo simulations were performed for select high-risk scenarios to assess the robustness of flood forecasts.

5.1. Province-Specific Insights:

- **Punjab:** High flood probability concentrated along the Indus River and its tributaries, with agricultural zones most vulnerable.
- **Sindh:** Coastal and low-lying districts showed elevated flood risk, compounded by delayed drainage during heavy rainfall events.
- **Khyber Pakhtunkhwa (KPK):** Mountainous terrain caused localized flooding, particularly in river valleys, necessitating micro-level monitoring.
- **Balochistan:** Flash floods were predicted in arid and semi-arid regions following sporadic intense rainfall, highlighting the need for rapid-response systems.
- **Model Comparison:** Ensemble approaches combining multiple machine learning models generally outperformed single-model predictions, achieving higher F1-scores and lower RMSE values across all provinces. Deep learning models captured complex non-linear interactions in environmental features, while classical models like Random Forest and XGBoost provided interpretable feature importance insights.
- **Implications for Disaster Management:** The province-wise flood predictions enable authorities to prioritize mitigation measures, optimize resource deployment, and implement proactive evacuation and relief strategies. Early warning systems based on

these forecasts can significantly reduce human and economic losses associated with flood events between 2025 and 2030.(Rasheed et al.)

CHAPTER FOUR

A PRACTICAL CASE STUDY OF MACHINE LEARNING APPLICATIONS: DISASTER FORECASTING AND MANAGEMENT

Fatima Abbas
Riphah International University Faisalabad

*Corresponding Author: fatima.abbas@riphahfsd.edu.pk

This chapter presents the results and discussion of this research, that is on flood prediction in Pakistan's four provinces: These results focus on flood prediction across Pakistan's four provinces which include Punjab Sindh KPK and Balochistan. Our research evaluated the performance of Decision Tree, Random Forest, Linear Regression, and Support Vector Machine (SVM) models in flood prediction using accuracy, precision, recall, F1-score, and mean squared error (MSE). With these measurements research teams projected future flood risks. Random Forest provided the best model results since it effectively analyzes complicated datasets. Our flood risk analysis shows what areas across four provinces stand highest risk for flooding during 2025 to 2030. This chapter explores data problems and environmental changes that affect its analysis.

1. Model Performance Overview

Evaluating the predictive accuracy of machine learning models across different provinces provides critical insights into their effectiveness in modeling climatic parameters. The models applied in this study – Decision Tree, Random Forest, Linear Regression, and SVM – exhibited varying degrees of accuracy across Punjab, Sindh, KPK, and Balochistan, reflecting differences in climate patterns, data complexity, and model adaptability. The results indicate that non-linear models, particularly Random Forest and SVM, consistently outperformed Linear Regression, highlighting the intricate nature of climate trends in most regions. The summary of the results is shown in the Table Error! No text of specified style in document..**1.**

Across all regions, non-linear models, particularly Random Forest and SVM, consistently outperformed Linear Regression. This highlights the intricate, non-linear nature of climate trends in most regions, with tree-based and kernel-based models showing superior performance in capturing complex climatic relationships.

In Punjab, Random Forest and SVM achieved the highest accuracy at 99.15%, demonstrating their superior ability to capture complex climatic relationships in this region. Punjab's diverse climate, influenced by seasonal monsoons and temperature fluctuations, likely contributed to the limitations observed in Linear Regression, which produced an accuracy of 42.23%. This substantial performance gap suggests that Punjab's climate data follows a highly non-linear pattern, making tree-based and kernel-based models significantly more effective. Decision Tree, while performing well at 88.13%, was still outperformed by Random Forest, indicating that ensemble learning techniques provided a critical advantage in predictive accuracy.

Table Error! No text of specified style in document.**.1**: Summary of Model Accuracy Across Provinces

Province	Decision Tree (%)	Random Forest (%)	Linear Regression (%)	SVM (%)
Punjab	88.13	99.15	42.23	99.15
Sindh	96.98	96.98	96.98	96.98
KPK	99.92	98.4	90.83	90.83
Balochistan	97.17	99	93.5	99

Sindh exhibited a strikingly homogeneous model performance, with all four models –Decision Tree, Random Forest, Linear Regression, and SVM – achieving identical accuracies of 96.98%. This uniformity suggests that Sindh's climatic variables exhibit strong underlying patterns that can be effectively captured by both linear and non-linear models. Unlike Punjab, where Linear Regression struggled due to pronounced non-linearity, Sindh's data appears to contain more stable trends, allowing even traditional regression techniques to perform at par with more advanced algorithms. This finding implies that climatic variability in Sindh is lower, making it one of the more predictable regions analyzed in this study.

Khyber Pakhtunkhwa (KPK) presented a different trend, with Decision Tree achieving the highest accuracy at 99.92%, slightly outperforming Random Forest, which scored 98.40%. This finding suggests that KPK's climatic data may contain distinct decision boundaries, allowing tree-based models to effectively segment and classify climate variations. The high accuracy of both Decision Tree and Random Forest reflects the ability of these models to handle complex interactions between climatic factors, particularly in a region with pronounced seasonal changes and significant topographical variations. Interestingly, Linear Regression and SVM, though generally less suited for non-linear relationships, still managed to achieve 90.83% accuracy, indicating that while KPK's climate exhibits complexity, it also retains some degree of predictability.

Balochistan displayed a performance pattern similar to Punjab and KPK, with Random Forest and SVM achieving the highest accuracy at 99.00%. These results reinforce the effectiveness of ensemble-based and kernel-based approaches in capturing Balochistan's arid and extreme climatic conditions, which are often influenced by erratic temperature fluctuations and limited rainfall. Decision Tree also performed well, achieving an accuracy of 97.17%, while Linear Regression, despite being the weakest among the four models, still managed a notable 93.50% accuracy. This relatively high performance of Linear Regression suggests that Balochistan's climate trends, while complex, may contain stronger linear relationships compared to Punjab and KPK.

Random Forest and SVM consistently emerged as the most reliable models across all provinces, achieving peak performance in Punjab, KPK, and Balochistan. The high accuracy of these models underscores their ability to handle large-scale, multi-variable climate datasets and adapt to non-linear interactions between climatic parameters. Decision Tree, though slightly less effective than Random Forest, demonstrated strong predictive power, particularly in KPK, where it outperformed all other models. Linear Regression, while successful in Sindh, struggled in Punjab, reinforcing the notion that regions with higher climatic variability require more sophisticated predictive approaches.

These findings indicate that climate prediction models must be selected based on regional characteristics rather than applying a one-size-fits-all approach. The effectiveness of Random Forest and SVM suggests that machine learning techniques that leverage ensemble learning or non-linear transformations are best suited for complex and highly variable climatic conditions. The uniform performance of all models in Sindh highlights an exceptional level of predictability, which may be attributed to more stable climate patterns in this region. In contrast, the high accuracy of tree-based models in KPK and Balochistan reflects the need for decision-based approaches in regions with significant seasonal and topographical variability.

2. Model Evaluation Results

The model evaluation results for the prediction of climatic and agricultural parameters across the provinces of Punjab, Sindh, Khyber Pakhtunkhwa (KPK), and Balochistan highlight the effectiveness of various machine learning models, including Linear Regression, Support Vector Regression (SVR), Decision Tree, and Random Forest Regression. Among these, Random Forest Regression consistently exhibited the best performance across most parameters, demonstrating strong predictive capabilities and reliability.

2.1. Temperature Changes

For Punjab, the Random Forest Regression model outperformed others with a score of 0.9483 and an explained variance of 0.9498, indicating strong temperature prediction accuracy. The Decision Tree model also performed well (score: 0.9165, variance: 0.9196) but lagged behind Random Forest. Linear Regression and SVR performed poorly, showing weak prediction ability. In Sindh, Random Forest Regression remained the top model (score: 0.9209, variance: 0.9258), while Decision Tree followed (score: 0.8699, variance: 0.8804). Linear Regression and SVR were ineffective, with negative variances.

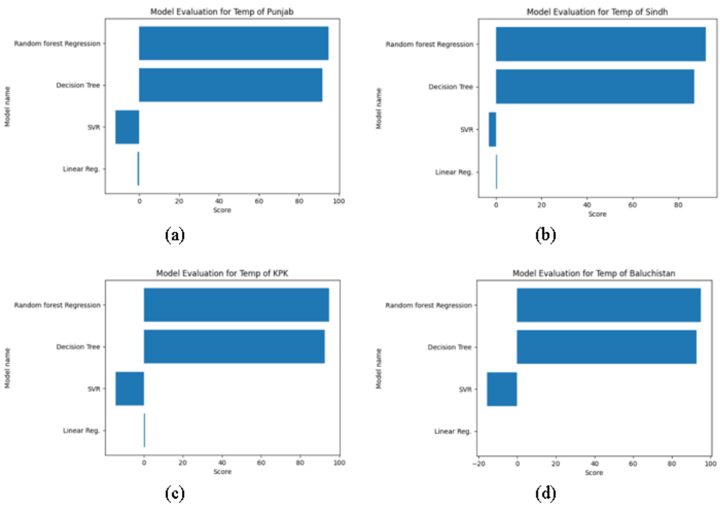
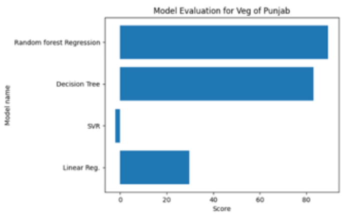


Figure *Error! No text of specified style in document..**5:** Graphs showing Model Evaluation of the Temperature (mm) (a) Punjab, (b) Sindh, (c) KPK, and (d) Baluchistan*

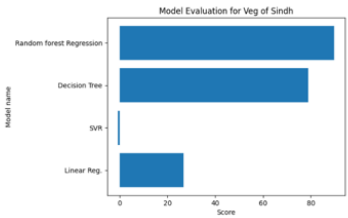
In Khyber Pakhtunkhwa (KPK), Random Forest Regression excelled (score: 0.9475, variance: 0.9469), closely followed by Decision Tree (score: 0.9266, variance: 0.9266). Linear Regression and SVR were again ineffective. Similarly, in Balochistan, Random Forest Regression led (score: 0.9500, variance: 0.9539), with Decision Tree performing well (score: 0.9276, variance: 0.9365). Linear Regression and SVR showed poor results. Overall, Random Forest Regression consistently outperformed other models across all regions, while Linear Regression and SVR were unsuitable for temperature prediction.

2.2. Vegetation Level (NDVI)

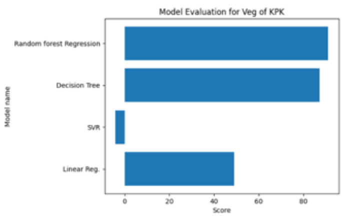
The Random Forest Regression model consistently outperformed other models across all regions. In Punjab, it achieved a score of 0.8929 and an explained variance of 0.8712, while the Decision Tree model followed with a score of 0.8297 and variance of 0.8183. Linear Regression and SVR performed poorly, with negative scores, making them unsuitable for this task. Similarly, in Sindh, Random Forest led with a score of 0.8976 and variance of 0.8687, with Decision Tree trailing (score: 0.7903, variance: 0.7860). Again, Linear Regression and SVR underperformed.



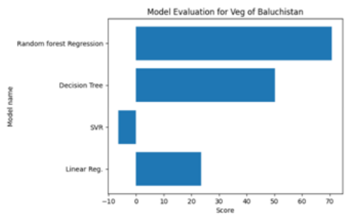
(a)



(b)



(c)



(d)

***Figure** Error! No text of specified style in document.**.6:** Graphs showing Model Evaluation of the Vegetation Level (a) Punjab, (b) Sindh, (c) KPK, and (d) Baluchistan*

In KPK, Random Forest Regression excelled with a score of 0.9108 and variance of 0.8882, while Decision Tree achieved a score of 0.8731 and variance of 0.8656. Linear Regression and SVR were ineffective. In Balochistan, Random Forest performed moderately (score: 0.7072, variance: 0.5666), with Decision Tree showing weaker results (score: 0.5025, variance: 0.5580). Linear Regression and SVR again failed to deliver accurate predictions. Overall, Random Forest Regression proved to be the most reliable model for vegetation prediction, while Linear Regression and SVR struggled across all regions.

2.3. ICE (Individual Conditional Expectation)

For ice prediction (ICE), the Random Forest Regression model consistently outperformed others across all regions, though its accuracy varied. In Punjab, it showed moderate success (score: 0.63, variance: 0.42), slightly edging out the Decision Tree model, which had higher variance but lower precision.

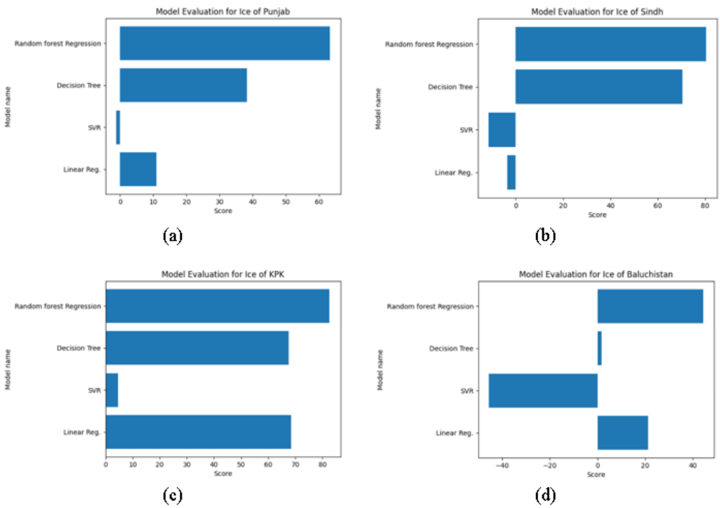
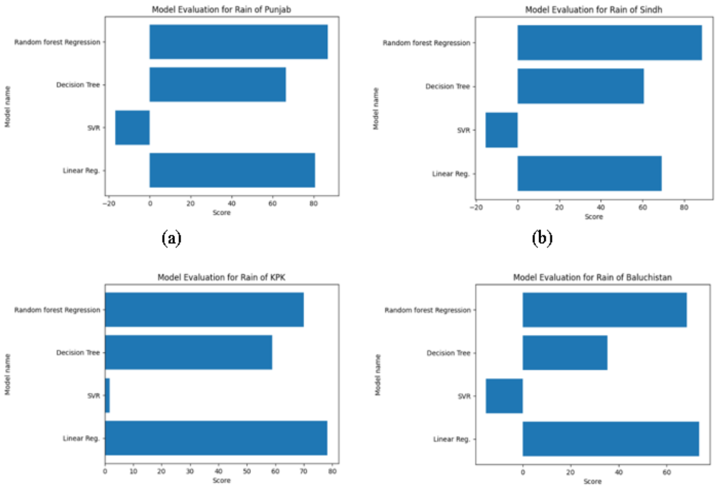


***Figure** Error! No text of specified style in document..**7:** Graphs showing Model Evaluation of the Ice (a) Punjab, (b) Sindh, (c) KPK, and (d) Balochistan*

In Sindh, Random Forest performed notably better (score: 0.80, variance: 0.70), with Decision Tree trailing behind (score: 0.70, variance: 0.59). KPK saw the strongest results for Random Forest (score: 0.83, variance: 0.81), while Decision Tree lagged further (score: 0.67, variance: 0.74). Balochistan proved challenging: Random Forest managed only modest accuracy (score: 0.44, variance: 0.17), and Decision Tree performed very poorly (score: 0.02, variance: 0.33). Linear Regression and SVR models struggled in all regions, often producing unreliable or negative results, making them unsuitable for ice prediction. Overall, Random Forest emerged as the most dependable option, though accuracy varied by region.

2.4. Rainfall

Random Forest outperformed other models across regions. In Punjab, it scored 0.8690 (variance: 0.8278), while Decision Tree lagged (score: 0.6634, variance: 0.5737). Sindh saw even stronger results for Random Forest (score: 0.8845, variance: 0.8886), with Decision Tree trailing (score: 0.6059, variance: 0.7600).



(a) (b) (c) (d)

***Figure** Error! No text of specified style in document..**8**: Graphs showing Model Evaluation in terms of the Rainfall (a) Punjab, (b) Sindh, (c) KPK, and (d) Balochistan*

In KPK, Random Forest scored 0.7009 (variance: 0.5361), slightly ahead of Decision Tree (score: 0.5888, variance: 0.5424). Balochistan had lower overall accuracy: Random Forest scored 0.6841 (variance: 0.4114), while Decision Tree struggled (score: 0.3523, variance: 0.2498). Linear Regression and SVR consistently underperformed, with negative variances in all regions, making them unreliable for rainfall prediction.

Across all four provinces, Random Forest Regression proved to be the most reliable model for predicting temperature, vegetation, ice, and rainfall. Its consistently high performance in terms of both score and explained variance across most parameters and regions demonstrates its robustness and suitability for environmental predictions. While the Decision Tree model performed well in some cases, especially for temperature and vegetation, it could not consistently match the Random Forest model. The Linear Regression and SVR models struggled throughout, showing negative results for several parameters, indicating that they were not ideal choices for these types of predictions.

3. Actual vs Predicted Temperature

Figure Error! **No text of specified style in document.**.9 illustrates projected and observed temperature trends across Pakistan's provinces (Punjab, Sindh, KPK, Balochistan) from 2025–2030, comparing machine-learning forecasts with historical data. This analysis evaluates model accuracy and highlights future temperature patterns critical for understanding their potential impact on regional flood risks.

Results reveal a steady rise in predicted temperatures across all provinces. Such warming trends could intensify flood risks by triggering heavier rainfall and accelerating snowmelt in mountainous areas, emphasizing the need for reliable temperature predictions to anticipate extreme weather and guide mitigation efforts.

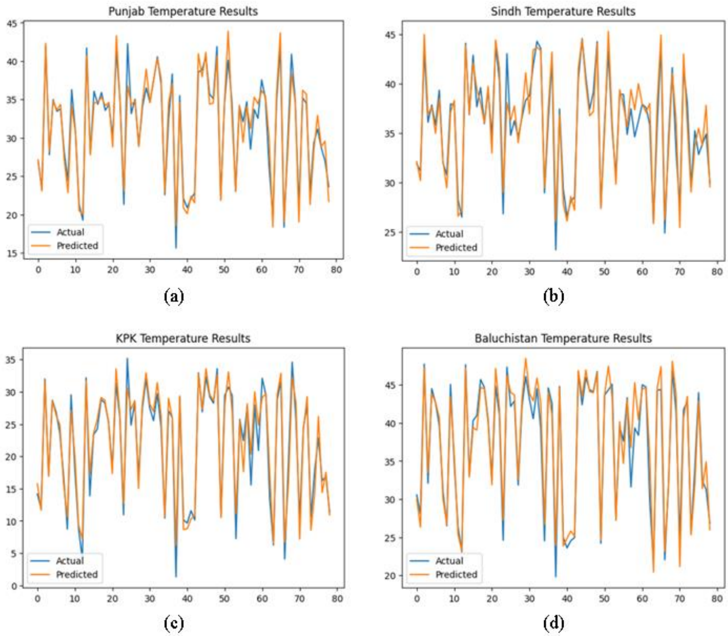


Figure **Error! No text of specified style in document.**.9 Graphs showing Actual vs Predicted Temperature for (a) Punjab, (b) Sindh, (c) KPK, and (d) Balochistan

3.1. Punjab Temperature

In Punjab, model predictions closely match observed temperatures overall, though minor discrepancies emerge during extremes. For instance, during intense heatwaves (e.g., observed 41.93°C vs. predicted 42.31°C), models slightly underestimate peak heat, while cooler days (e.g., observed 19.27°C vs. predicted 20.01°C) show small overestimations. This highlights their strength in capturing general trends but limitations in pinpointing extreme shifts.

The models – especially Random Forest and SVM – effectively track historical trends, offering reliable mid-range forecasts. However, accuracy dips during heatwaves or cold spells, suggesting refinements are needed to better predict severe weather events.

3.2. Khyber Pakhtunkhwa (KPK) Temperature

In KPK, temperature forecasts are generally reliable but show inconsistencies on colder days. For example, an observed temperature of 4.08°C was predicted as 7.19°C, highlighting a significant deviation. While the models perform well for higher temperatures, they struggle with lower ones, especially in mountainous and low-altitude areas where temperatures drop sharply.

The forecasted temperature trends indicate a steady rise, suggesting a growing impact of climate change. The models predict more frequent temperature extremes, which could increase flood risks, particularly in regions prone to flash flooding due to rapid snowmelt in the mountains.

3.3. Sindh Temperature

Sindh's temperature predictions are generally accurate, with minimal deviations. For instance, an observed temperature of 31.91°C was predicted as 32.13°C, indicating strong model performance. However, at extreme temperatures (e.g., 43.43°C observed vs. 44.99°C predicted), the models tend to overestimate, likely due to the region's coastal and arid climate, where rapid fluctuations influenced by the ocean are harder to predict precisely.

The forecasts indicate more frequent high-temperature events, increasing the risk of prolonged heatwaves, water scarcity, and agricultural stress. Rising sea levels and intensified storms could also worsen coastal flooding.

3.4. Balochistan Temperature

In Balochistan, the model predictions exhibit a high degree of accuracy, particularly in the middle range of temperatures. However, as with the other provinces, extreme temperature predictions, especially during heatwaves, tend to deviate slightly. For instance, a temperature of 44.57°C observed was predicted at 45.16°C, which is close but not exact. This suggests that while the models can generally capture the increasing temperature trends in Balochistan, they still have difficulty in predicting extreme temperature outliers.

Balochistan's future temperature trends align with the expected regional warming pattern, with a notable increase in predicted temperature extremes. The rise in temperatures is expected to exacerbate the risk of droughts, which could further strain water resources in the already water-scarce region.

3.5. Discussion and Insights

The machine-learning models – SVM, Decision Tree, Random Forest, and Linear Regression – effectively captured overall temperature trends but struggled with extreme weather events like heatwaves and cold spells. This limitation likely stems from their reliance on historical data, which may not fully reflect the increasing frequency and intensity of such events due to climate change.

While the models performed well in predicting long-term trends, their inaccuracies in extreme conditions highlight the need for refinement. Advanced techniques, such as deep learning and hybrid models, could improve accuracy by better capturing climate complexities.

Rising temperatures have serious flood risk implications in Pakistan. Faster snowmelt in northern mountains will increase runoff, leading to more frequent and severe flooding. Additionally, higher temperatures may intensify rainfall, overwhelming drainage systems and exacerbating urban flood risks.

4. Actual vs Predicted Vegetation Level

Figure Error! *No text of specified style in document.*.*7* compares the actual and predicted vegetation levels for four provinces of Pakistan: Punjab, Sindh, Khyber Pakhtunkhwa (KPK), and Balochistan. The comparison aims to evaluate the accuracy of the prediction models used for estimating vegetation levels and understand how well the model aligns with the real-world data.

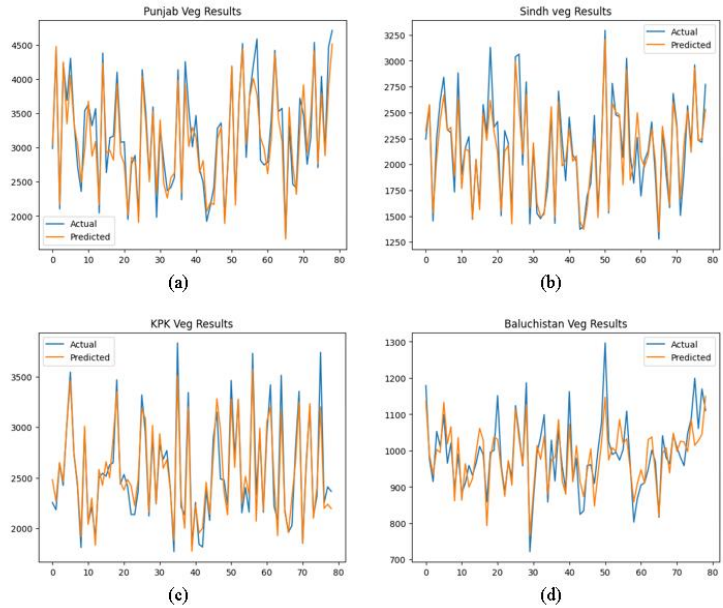


Figure **Error! No text of specified style in document.**.10 Graphs showing Actual vs Predicted Vegetation Level for (a) Punjab, (b) Sindh, (c) KPK, and (d) Balochistan

4.1. Punjab Vegetation Level

The comparison of actual and predicted vegetation levels for Punjab reveals some noteworthy trends. While the predicted values are generally close to the actual values, there are several deviations. For example, in the first few observations, the model predicts vegetation levels slightly higher than the actual values, such as in the first data pair where the predicted value is 3040.15 compared to an actual value of 2988.46. However, in some cases, the predicted values show larger deviations, such as the predicted value of 3350.27 for an actual value of 3693.23, which suggests a notable underestimation by the model. Despite these discrepancies, the model shows reasonable performance with an overall tendency to underpredict in many instances.

4.2. Sindh Vegetation Level

For Sindh, the predicted vegetation levels are in close proximity to the actual values, but the model still shows some variation. In the first observation, the predicted value of 2322.92 is slightly higher than the actual value of 2242.72. The model appears to perform reasonably well for most of the data points, with smaller variations between predicted and actual values compared to Punjab. For instance, in the case where the actual value is 2570.98, the predicted value is nearly identical at 2576.30. However, a few data points show more noticeable deviations, such as the predicted value of 2407.75 for an actual value of 2343.69. Overall, the predictions seem to reflect the actual vegetation levels fairly accurately, though occasional over- and underestimations are evident.

4.3. KPK Vegetation Level

In KPK, the actual vs predicted vegetation levels also exhibit close alignment, with the predictions being somewhat consistent. However, there are some underestimations, particularly with higher vegetation levels. For instance, in the case where the actual value is 3544.41, the predicted value is 3459.64, showing a minor underestimation. Other pairs, such as the predicted value of 2298.59 for an actual value of 2210.87, also indicate slight discrepancies. However, overall, the model appears to be a reasonable fit for most of the data points, especially in the middle range of vegetation levels.

4.4. Balochistan Vegetation Level

Balochistan exhibits a slightly more variable pattern when comparing the actual and predicted vegetation levels. There are noticeable discrepancies where the model underestimates or overestimates the vegetation levels. For example, in one case, the actual value of 959.20 is predicted as 899.26, suggesting an underestimation. In contrast, some observations, such as the predicted value of 1085.07 for an actual value of 1162.43, indicate an overestimation. The prediction model seems less precise for Balochistan compared to the other provinces, as it exhibits larger variances between actual and predicted values.

4.5. Discussion and Insights

The overall analysis of the actual versus predicted vegetation levels for the four provinces demonstrates that while the model generally provides reasonable estimates, there are some notable deviations. The model tends to underpredict in many cases, particularly for Punjab and KPK. Sindh shows the closest alignment between actual and predicted values, indicating that the model performs relatively well for this province. Balochistan, however, exhibits more variability, suggesting that the model may require further refinement or calibration for more accurate predictions in this region.

Several factors may contribute to these discrepancies, such as the quality and resolution of the data used for training the prediction model, regional environmental factors, and potential limitations of the model itself. Further analysis, including a more detailed assessment of the model's performance across different climatic zones within the provinces, could offer more insights into the factors influencing these variations.

5. Actual vs Predicted ICE (Individual Conditional Expectation)

Figure Error*! No text of specified style in document.*.***8*** shows the differences between actual and predicted ICE for four regions: Punjab, Sindh, Khyber Pakhtunkhwa (KPK), and Balochistan. The observed and predicted values are compared, and insights are drawn to evaluate the accuracy and possible implications for future predictions.

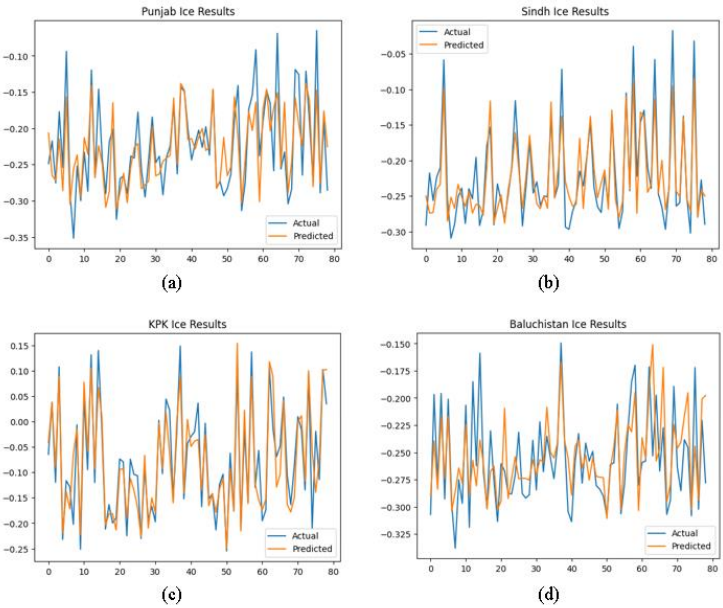


Figure **Error! No text of specified style in document.**.11 Graphs showing Actual vs Predicted ICE for (a) Punjab, (b) Sindh, (c) KPK, and (d) Balochistan

5.1. Punjab ICE

Punjab's ICE fluctuates significantly, ranging from -0.350 to -0.069. The predicted values generally follow the trend but often underestimate actual meltage, especially when values drop below -0.25. In some cases, the model closely matches actual values, but discrepancies appear, particularly for extreme points like -0.304 (predicted) vs. -0.289 (actual). This suggests the model captures overall trends but struggles with finer variations.

5.2. Sindh ICE

Sindh's ICE ranges from -0.039 to -0.317, with predicted values closely following actual trends. However, the model sometimes over- or underestimates ICE, particularly at extreme points. For instance, an actual value of -0.2911 is predicted as -0.2501, showing underestimation. While the model is generally consistent, further refinement is needed to improve accuracy in extreme cases.

5.3. KPK ICE

KPK's ICE fluctuates between -0.231 and 0.148, with predictions largely following the trend but occasionally lagging or exceeding actual values. For example, an actual value of -0.2319 is predicted as -0.2196, showing slight underestimation. The model provides a reasonable approximation but struggles with extreme fluctuations, suggesting a need for further adjustments.

5.4. Balochistan ICE

Balochistan's ICE ranges from -0.317 to -0.149, with predicted values following the general trend but showing deviations. The model sometimes underpredicts or overpredicts significantly, such as -0.3045 (actual) vs. -0.2547 (predicted). These inconsistencies indicate difficulty in capturing extreme variations.

5.5. Discussion and Insights

Upon reviewing the actual versus predicted ICE data across the four regions, several insights emerge. The predictive model shows an overall ability to capture the trends in ICE across the different regions, particularly in regions like Punjab and Sindh, where the predictions are relatively close to the actual values. However, discrepancies remain, especially in extreme values, indicating that the model may require further refinement to account for more localized variables or fluctuations in ICE.

6. Actual vs Predicted Rainfall

Error! Reference source not found. shows that the model performs well across all four provinces, capturing overall rainfall trends. In Punjab, it accurately predicts moderate to high rainfall. Sindh exhibits strong alignment, especially during dry periods. KPK's predictions closely follow actual rainfall, particularly in the mid-range. Balochistan shows reasonable performance but with more variability. These results suggest the model is effective for regional rainfall forecasting, aiding planning and decision-making.

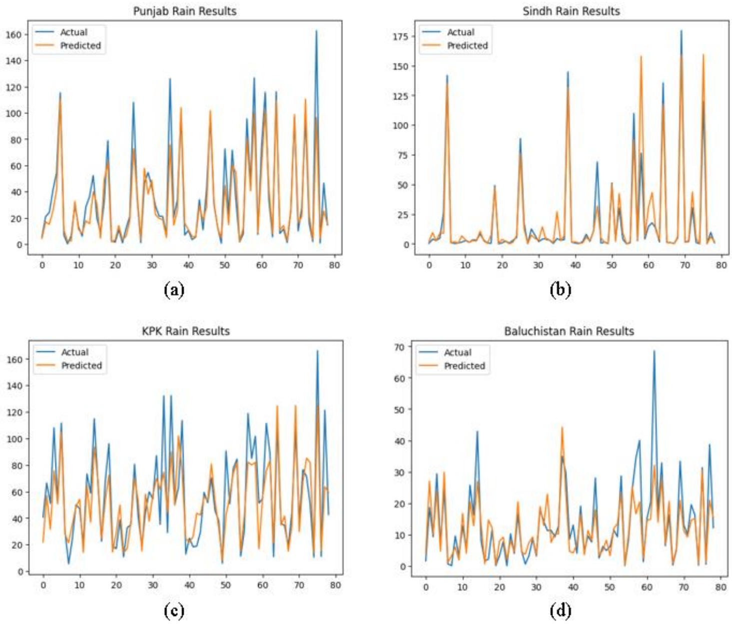


Figure **Error! No text of specified style in document.**.12 Graphs showing Actual vs Predicted Rainfall for (a) Punjab, (b) Sindh, (c) KPK, and (d) Balochistan

6.1. Punjab Rainfall

Punjab's rainfall predictions are generally close but show underestimations, especially for high rainfall. For example, an actual rainfall of 5.445 mm was predicted at 4.29 mm, showing good accuracy. However, larger deviations exist, such as 125.98 mm being predicted at only 75.57 mm. Lower rainfall values tend to be predicted more accurately, with most errors within a 20-30% margin.

6.2. Sindh Rainfall

The model performs reasonably well in Sindh, particularly for higher rainfall values. For instance, an actual rainfall of 141.87 mm was predicted at 135.26 mm. However, during dry periods, it tends to overestimate, such as predicting 0.1927 mm for an actual 0.046 mm. This suggests the model struggles with small-scale precipitation in arid conditions.

6.3. KPK Rainfall

KPK's rainfall predictions align well with actual values in moderate conditions but underestimate heavy rainfall. For instance, an actual 132.29 mm was predicted at 89.81 mm. Predictions for lower rainfall values are more precise, such as 29.86 mm being estimated at 27.9 mm. The model's performance is reliable in moderate conditions but weaker in extreme rainfall events.

6.4. Balochistan Rainfall

Balochistan shows the most variability. Some months see large underestimations, such as 68.56 mm predicted at 32.08 mm, while dry periods often have overestimations, like 0.039 mm predicted at 3.198 mm. The model struggles with Balochistan's erratic rainfall, likely due to the region's sporadic weather patterns.

6.5. Discussion and Insights

The comparison of actual versus predicted rainfall across the four provinces of Pakistan reveals that the prediction model demonstrates a general tendency to either overestimate or underestimate rainfall, with more pronounced errors during extreme rainfall events. In regions like Punjab and KPK, where rainfall is more frequent and relatively moderate, the model shows a better overall performance. However, in areas like Sindh and Balochistan, the model struggles to accurately predict rainfall, particularly in dry periods and during heavy rainfall events.

The discrepancies in the model's predictions could stem from several factors, such as local variations in weather patterns, topography, and the inherent difficulty in forecasting extreme weather events. Additionally, the model's reliance on historical data may not adequately capture the dynamic nature of rainfall, especially in areas prone to rapid changes in weather conditions.

To improve the accuracy of rainfall predictions, further refinement of the model is required, possibly by incorporating more granular data such as atmospheric conditions, soil moisture, and more localized weather patterns. Furthermore, integrating real-time data from weather stations could help enhance the model's adaptability and precision in forecasting rainfall events.

7. Future Prediction of floods for each Province

Machine learning models were trained and tested using data from the 2000 to 2021 database to predict how floods might affect Punjab in the next ten years. From 2025 to 2030, flood predictions were evaluated through Decision Tree, Random Forest, Linear Regression, and SVM analysis. The diverse machine learning models created unique findings by applying their own calculation methods on flood data, highlighting their strengths and limitations.

7.1. Flood Prediction in Punjab Using Random Forest

The Random Forest model predicts rainfall in Punjab from 2025 to 2030, revealing distinct seasonal patterns (Figure Error! **No text of specified style in document.**.13). From January to March, rainfall remains low (below 50mm), with low-risk levels indicating stable, dry conditions. For example, January 2025 records 23.24mm of rainfall and a low-risk classification. Late autumn and winter months (November and

December) also show minimal rainfall, rarely exceeding 20mm.

From May to September, rainfall and risk levels rise significantly. June 2025, for instance, sees 230.5mm of rainfall and a moderate-risk classification, marking the monsoon's onset. Rainfall consistently exceeds 200mm during these months, peaking at 299.95mm in July 2030, accompanied by high-risk levels.

The model highlights a clear annual pattern: June, July, and August experience substantial rainfall, while early months remain dry. July consistently records the highest rainfall (e.g., 297.93mm in 2025 and 299.95mm in 2030), underscoring the monsoon's peak intensity.

7.2. Flood Prediction in Punjab using Decision Tree

The Decision Tree model predicts environmental risks in Punjab from 2025 to 2030, analyzing temperature, rainfall, ice, and vegetation (Figure Error! **No text of specified style in document.**.14). Early months (January-March) show moderate temperatures (16.0°C–24.9°C) and low rainfall, with "Low" or "None" risk classifications.

June and July are critical, with temperatures reaching 39.7°C and rainfall exceeding 200mm, resulting in "High" risk levels. August, despite high rainfall, shows lower risk, possibly due to vegetation growth. Middle months (April-May) exhibit moderate conditions, with temperatures (15°C–23°C) and rainfall (16mm–48mm), and generally low risks.

November and December see reduced risk, with lighter rainfall and colder temperatures. Ice accumulation remains consistently negative, indicating minimal concern for freezing conditions.

7.3. Flood Prediction in Punjab using the Linear Regression

Linear Regression predictions for Punjab (2025–2030) highlight flood risks driven by temperature, rainfall, ice, and vegetation (Figure Error! **No text of specified style in document.**.15). High-risk periods occur in June, July, and August, with rainfall exceeding 200mm and flood risk nearing 100%. For example, June 2025 (230.5mm) and July 2025 (297.93mm) both show 100% flood risk.

Temperature trends exacerbate these risks. In June and July 2026, temperatures exceed 39°C, intensifying rainfall and flood severity. Similar patterns persist in 2029 and 2030, with July 2029 (247.29mm) and June 2030 (299.95mm) recording extreme flood risks.

Spring and early autumn months (e.g., April 2025) show lower risks, with moderate temperatures and rainfall. However, the data underscores a growing frequency and intensity of summer floods, driven by rising temperatures and erratic rainfall.

Random Forest Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Random Forest (%)	Random Forest Risk
2025	1	24.7	23.24	-8.87	30.78	1 %	Low
2025	2	15.3	40.66	-8.31	32.06	6 %	Low
2025	3	20.1	26.66	-1.6	20.55	3 %	Low
2025	4	16.3	30.64	-0.03	43.91	11 %	Low
2025	5	19.3	29.67	-4.17	43.52	3 %	Low
2025	6	37.2	230.5	-2.69	73.72	32 %	Moderate
2025	7	32.2	297.93	-2.69	52.17	27 %	Moderate
2025	8	35.7	296.16	-4.34	99.49	34 %	Moderate
2025	9	21.3	19.16	-1.09	43.76	0 %	Low
2025	10	24.6	32.22	-2.04	24.66	1 %	Low
2025	11	20.8	1.18	-2.8	36.42	2 %	Low
2025	12	16.6	18.1	-9.34	20.88	1 %	Low
2026	1	21.2	48.49	-1.62	23.51	5 %	Low
2026	2	23.2	6.31	-7.04	30.04	1 %	Low
2026	3	18.5	0.11	-1.76	10.44	2 %	Low
2026	4	15.5	22.84	-6.12	26.26	1 %	Low
2026	5	15.2	16.84	-3.94	39.58	1 %	Low
2026	6	39.7	240.58	-3.9	82.72	32 %	Moderate
2026	7	34.0	281.84	-2.04	95.55	27 %	Moderate
2026	8	33.2	231.38	-4.29	75.39	27 %	Moderate
2026	9	21.7	11.99	-5.73	28.93	0 %	Low
2026	10	23.0	16.2	-0.1	41.53	10 %	Low
2026	11	22.3	7.74	-4.85	48.57	0 %	Low
2026	12	18.5	2.05	-5.95	16.74	1 %	Low
2027	1	18.3	2.44	-8.15	39.93	1 %	Low
2027	2	22.8	43.36	-4.31	33.84	4 %	Low
2027	3	19.4	12.2	-7.38	42.97	2 %	Low
2027	4	21.7	37.55	-2.7	25.64	2 %	Low
2027	5	23.8	48.14	-1.76	25.44	3 %	Low
2027	6	30.2	274.11	-2.54	57.97	27 %	Moderate
2027	7	36.5	264.94	-3.28	69.08	35 %	Moderate
2027	8	39.2	222.82	-1.67	60.58	34 %	Moderate
2027	9	23.6	32.64	-9.19	33.28	1 %	Low
2027	10	18.5	12.68	-2.13	33.58	1 %	Low
2027	11	21.4	1.88	-9.7	18.23	0 %	Low
2027	12	21.5	31.95	-3.68	30.93	2 %	Low
2028	1	16.3	2.64	-3.76	20.45	1 %	Low
2028	2	16.3	17.7	-6.54	22.55	2 %	Low
2028	3	17.7	34.27	-2.18	38.11	4 %	Low
2028	4	23.6	27.6	-7.78	40.25	1 %	Low
2028	5	22.5	29.79	-1.66	30.31	2 %	Low
2028	6	30.5	208.31	-1.95	97.59	27 %	Moderate
2028	7	39.2	218.79	-3.29	93.51	34 %	Moderate
2028	8	32.9	268.63	-1.63	60.78	27 %	Moderate
2028	9	17.9	17.67	-4.89	12.32	1 %	Low
2028	10	17.0	37.51	-3.24	22.72	4 %	Low
2028	11	17.9	3.16	-8.71	12.36	1 %	Low
2028	12	21.8	35.64	-8.13	13.86	2 %	Low
2029	1	19.8	32.62	-5.99	10.15	3 %	Low
2029	2	24.1	48.77	-1.99	43.62	4 %	Low
2029	3	16.8	45.15	-2.78	14.65	11 %	Low
2029	4	19.6	49.86	-1.89	18.89	9 %	Low
2029	5	15.9	7.73	-8.08	36.43	1 %	Low
2029	6	35.2	219.14	-1.23	73.62	33 %	Moderate
2029	7	34.4	247.29	-4.62	96.68	27 %	Moderate
2029	8	32.6	252.0	-3.82	95.01	27 %	Moderate
2029	9	23.9	20.24	-7.77	42.58	0 %	Low
2029	10	17.6	9.83	-1.74	38.47	1 %	Low
2029	11	17.1	1.52	-6.13	22.52	1 %	Low
2029	12	24.9	23.12	-0.91	49.01	0 %	Low
2030	1	16.2	11.39	-2.59	46.75	1 %	Low
2030	2	16.7	44.6	-8.55	37.07	11 %	Low
2030	3	20.3	14.14	-2.97	31.62	2 %	Low
2030	4	16.0	10.52	-0.65	31.25	1 %	Low
2030	5	15.4	19.67	-7.74	22.47	1 %	Low
2030	6	34.0	224.87	-3.76	85.43	27 %	Moderate
2030	7	37.8	299.95	-3.02	79.93	34 %	Moderate
2030	8	31.1	216.73	-3.85	59.32	27 %	Moderate
2030	9	18.7	25.33	-6.68	42.26	2 %	Low
2030	10	19.7	11.63	-2.19	12.52	1 %	Low
2030	11	24.9	46.71	-5.76	24.75	5 %	Low
2030	12	22.8	0.8	-1.78	16.16	0 %	Low

Figure **Error! No text of specified style in document.**.13 Month-wise Prediction of Flood in Punjab for 2025-2030 using Random Forest

Decision Tree Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Decision Tree (%)	Decision Tree Risk
2025	1	24.7	23.24	-8.87	30.78	13 %	Low
2025	2	15.3	40.66	-8.31	32.06	10 %	Low
2025	3	20.1	26.66	-1.6	20.55	10 %	Low
2025	4	16.3	30.64	-0.03	43.91	12 %	Low
2025	5	19.3	29.67	-4.17	43.52	10 %	Low
2025	6	37.2	230.5	-2.69	73.72	50 %	High
2025	7	32.2	297.93	-2.69	52.17	50 %	High
2025	8	35.7	296.16	-4.34	99.49	0 %	None
2025	9	21.3	19.16	-1.09	43.76	10 %	Low
2025	10	24.6	32.22	-2.04	24.66	14 %	Low
2025	11	20.8	1.18	-2.8	36.42	0 %	None
2025	12	16.6	18.1	-9.34	20.88	0 %	None
2026	1	21.2	48.49	-1.62	23.51	12 %	Low
2026	2	23.2	6.31	-7.04	30.04	10 %	Low
2026	3	18.5	0.11	-1.76	10.44	0 %	None
2026	4	15.5	22.84	-6.12	26.26	11 %	Low
2026	5	15.2	16.84	-3.94	39.58	0 %	None
2026	6	39.7	240.58	-3.9	82.72	50 %	High
2026	7	34.0	281.84	-2.04	95.55	50 %	High
2026	8	33.2	231.38	-4.29	75.39	47 %	High
2026	9	21.7	11.99	-5.73	28.93	0 %	None
2026	10	23.0	16.2	-0.1	41.53	10 %	Low
2026	11	22.3	7.74	-4.85	48.57	10 %	Low
2026	12	18.5	2.05	-5.95	16.74	14 %	Low
2027	1	18.3	2.44	-8.15	39.93	10 %	Low
2027	2	22.8	43.36	-4.31	33.84	0 %	None
2027	3	19.4	12.2	-7.38	42.97	10 %	Low
2027	4	21.7	37.55	-2.7	25.64	10 %	Low
2027	5	23.8	48.14	-1.76	25.44	10 %	Low
2027	6	30.2	274.11	-2.54	57.97	0 %	None
2027	7	36.5	264.94	-3.28	69.08	50 %	High
2027	8	39.2	222.82	-1.67	60.58	50 %	High
2027	9	23.6	32.64	-9.19	33.28	11 %	Low
2027	10	18.5	12.68	-2.13	33.58	10 %	Low
2027	11	21.4	1.88	-9.7	18.23	10 %	Low
2027	12	21.5	31.95	-3.68	30.93	13 %	Low
2028	1	16.3	2.64	-3.76	20.45	10 %	Low
2028	2	16.3	17.7	-6.54	22.55	0 %	None
2028	3	17.7	34.27	-2.18	38.11	10 %	Low
2028	4	23.6	27.6	-7.78	40.25	10 %	Low
2028	5	22.5	29.79	-1.66	30.31	0 %	None
2028	6	30.5	208.31	-1.95	97.59	0 %	None
2028	7	39.2	218.79	-3.29	93.51	0 %	None
2028	8	32.9	268.63	-1.63	60.78	0 %	None
2028	9	17.9	17.67	-4.89	12.32	10 %	Low
2028	10	17.0	37.51	-3.24	22.72	10 %	Low
2028	11	17.9	3.16	-8.71	12.36	10 %	Low
2028	12	21.8	35.64	-8.13	13.86	13 %	Low
2029	1	19.8	32.62	-5.99	10.15	10 %	Low
2029	2	24.1	48.77	-1.99	43.62	10 %	Low
2029	3	16.8	45.15	-2.78	14.65	10 %	Low
2029	4	19.6	49.86	-1.89	18.89	10 %	Low
2029	5	15.9	7.73	-8.08	36.43	10 %	Low
2029	6	35.2	219.14	-1.23	73.62	48 %	High
2029	7	34.4	247.29	-4.62	96.68	47 %	High
2029	8	32.6	252.0	-3.82	95.01	0 %	None
2029	9	23.9	20.24	-7.77	42.58	14 %	Low
2029	10	17.6	9.83	-1.74	38.47	10 %	Low
2029	11	17.1	1.52	-6.13	22.52	12 %	Low
2029	12	24.9	23.12	-0.91	49.01	12 %	Low
2030	1	16.2	11.39	-2.59	46.75	10 %	Low
2030	2	16.7	44.6	-8.55	37.07	14 %	Low
2030	3	20.3	14.14	-2.97	31.62	10 %	Low
2030	4	16.0	10.52	-0.65	31.25	11 %	Low
2030	5	15.4	19.67	-7.74	22.47	10 %	Low
2030	6	34.0	224.87	-3.76	85.43	50 %	High
2030	7	37.8	299.95	-3.02	79.93	50 %	High
2030	8	31.1	216.73	-3.85	59.32	48 %	High
2030	9	18.7	25.33	-6.68	42.26	13 %	Low
2030	10	19.7	11.63	-2.19	12.52	12 %	Low
2030	11	24.9	46.71	-5.76	24.75	10 %	Low
2030	12	22.8	0.8	-1.78	16.16	10 %	Low

Figure **Error! No text of specified style in document.**.14 Month-wise Prediction of Flood in Punjab for 2025-2030 using Decision Tree

7.4. Flood Prediction in Punjab Using SVM

SVM predictions for Punjab (2025–2030) reveal seasonal flood risk variations (Figure Error! **No text of specified style in document.**.16). In 2025, March and October show high-risk peaks (60% and 50%, respectively), while February, May, and

November exhibit low risks. Summer months (June-August) record heavy rainfall (e.g., 297.93mm in July 2025) but moderate risks, suggesting temperature and vegetation influence flood severity.

Linear Regression Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Linear Regression (%)	Linear Regression Risk
2025	1	24.7	23.24	-8.87	30.78	100 %	High
2025	2	15.3	40.66	-8.31	32.06	97 %	High
2025	3	20.1	26.66	-1.6	20.55	48 %	Moderate
2025	4	16.3	30.64	-0.03	43.91	19 %	Low
2025	5	19.3	29.67	-4.17	43.52	94 %	High
2025	6	37.2	230.5	-2.69	73.72	100 %	High
2025	7	32.2	297.93	-2.69	52.17	100 %	High
2025	8	35.7	296.16	-4.34	99.49	96 %	High
2025	9	21.3	19.16	-1.09	43.76	24 %	Moderate
2025	10	24.6	32.22	-2.04	24.66	47 %	Moderate
2025	11	20.8	1.18	-2.8	36.42	67 %	High
2025	12	16.6	18.1	-9.34	20.88	100 %	High
2026	1	21.2	48.49	-1.62	23.51	53 %	High
2026	2	23.2	6.31	-7.04	30.04	100 %	High
2026	3	18.5	0.11	-1.76	10.44	47 %	Moderate
2026	4	15.5	22.84	-6.12	26.26	100 %	High
2026	5	15.2	16.84	-3.94	39.58	100 %	High
2026	6	39.7	240.58	-3.9	82.72	100 %	High
2026	7	34.0	281.84	-2.04	95.55	98 %	High
2026	8	33.2	231.38	-4.29	75.39	100 %	High
2026	9	21.7	11.99	-5.73	28.93	100 %	High
2026	10	23.0	16.2	-0.1	41.53	10 %	Low
2026	11	22.3	7.74	-4.85	48.57	94 %	High
2026	12	18.5	2.05	-5.95	16.74	93 %	High
2027	1	18.3	2.44	-8.15	39.93	99 %	High
2027	2	22.8	43.36	-4.31	33.84	100 %	High
2027	3	19.4	12.2	-7.38	42.97	100 %	High
2027	4	21.7	37.55	-2.7	25.64	71 %	High
2027	5	23.8	48.14	-1.76	25.44	46 %	Moderate
2027	6	30.2	274.11	-2.54	57.97	100 %	High
2027	7	36.5	264.94	-3.28	69.08	100 %	High
2027	8	39.2	222.82	-1.67	60.58	90 %	High
2027	9	23.6	32.64	-9.19	33.28	91 %	High
2027	10	18.5	12.68	-2.13	33.58	50 %	High
2027	11	21.4	1.88	-9.7	18.23	93 %	High
2027	12	21.5	31.95	-3.68	30.93	96 %	High
2028	1	16.3	2.64	-3.76	20.45	99 %	High
2028	2	16.3	17.7	-6.54	22.55	100 %	High
2028	3	17.7	34.27	-2.18	38.11	70 %	High
2028	4	23.6	27.6	-7.78	40.25	100 %	High
2028	5	22.5	29.79	-1.66	30.31	38 %	Moderate
2028	6	30.5	208.31	-1.95	97.59	98 %	High
2028	7	39.2	218.79	-3.29	93.51	100 %	High
2028	8	32.9	268.63	-1.63	60.78	100 %	High
2028	9	17.9	17.67	-4.89	12.32	100 %	High
2028	10	17.0	37.51	-3.24	22.72	89 %	High
2028	11	17.9	3.16	-8.71	12.36	94 %	High
2028	12	21.8	35.64	-8.13	13.86	100 %	High
2029	1	19.8	32.62	-5.99	10.15	100 %	High
2029	2	24.1	48.77	-1.99	43.62	58 %	High
2029	3	16.8	45.15	-2.78	14.65	77 %	High
2029	4	19.6	49.86	-1.89	18.89	66 %	High
2029	5	15.9	7.73	-8.08	36.43	93 %	High
2029	6	35.2	219.14	-1.23	73.62	77 %	High
2029	7	34.4	247.29	-4.62	96.68	95 %	High
2029	8	32.6	252.0	-3.82	95.01	91 %	High
2029	9	23.9	20.24	-7.77	42.58	94 %	High
2029	10	17.6	9.83	-1.74	38.47	46 %	Moderate
2029	11	17.1	1.52	-6.13	22.52	99 %	High
2029	12	24.9	23.12	-0.91	49.01	20 %	Moderate
2030	1	16.2	11.39	-2.59	46.75	62 %	High
2030	2	16.7	44.6	-8.55	37.07	100 %	High
2030	3	20.3	14.14	-2.97	31.62	67 %	High
2030	4	16.0	10.52	-0.65	31.25	13 %	Low
2030	5	15.4	19.67	-7.74	22.47	100 %	High
2030	6	34.0	224.87	-3.76	85.43	100 %	High
2030	7	37.8	299.95	-3.02	79.93	100 %	High
2030	8	31.1	216.73	-3.85	59.32	96 %	High
2030	9	18.7	25.33	-6.68	42.26	100 %	High
2030	10	19.7	11.63	-2.19	12.52	56 %	High
2030	11	24.9	46.71	-5.76	24.75	100 %	High
2030	12	22.8	0.8	-1.78	16.16	47 %	Moderate

Figure **Error! No text of specified style in document.**.15 Month-wise Prediction of Flood in Punjab for 2025-2030 using Linear Regression

In 2026, March, June, and September are high-risk months, with June and July experiencing heavy rainfall (240.58mm and 281.84mm, respectively). February and December show minimal risks due to low rainfall and moderate temperatures.

By 2027–2028, flood risks decrease slightly, though summer months remain vulnerable. July 2027 records 264.94mm of rainfall with moderate risk, while June-August 2028 show consistent high rainfall and risk levels.

SVM Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	SVM (%)	SVM Risk
2025	1	24.7	23.24	-8.87	30.78	30 %	Moderate
2025	2	15.3	40.66	-8.31	32.06	5 %	Low
2025	3	20.1	26.66	-1.6	20.55	60 %	High
2025	4	16.3	30.64	-0.03	43.91	25 %	Moderate
2025	5	19.3	29.67	-4.17	43.52	5 %	Low
2025	6	37.2	230.5	-2.69	73.72	20 %	Moderate
2025	7	32.2	297.93	-2.69	52.17	5 %	Low
2025	8	35.7	296.16	-4.34	99.49	25 %	Moderate
2025	9	21.3	19.16	-1.09	43.76	20 %	Moderate
2025	10	24.6	32.22	-2.04	24.66	50 %	High
2025	11	20.8	1.18	-2.8	36.42	10 %	Low
2025	12	16.6	18.1	-9.34	20.88	30 %	Moderate
2026	1	21.2	48.49	-1.62	23.51	15 %	Low
2026	2	23.2	6.31	-7.04	30.04	5 %	Low
2026	3	18.5	0.11	-1.76	10.44	10 %	Low
2026	4	15.5	22.84	-6.12	26.26	50 %	High
2026	5	15.2	16.84	-3.94	39.58	25 %	Moderate
2026	6	39.7	240.58	-3.9	82.72	20 %	Moderate
2026	7	34.0	281.84	-2.04	95.55	50 %	High
2026	8	33.2	231.38	-4.29	75.39	30 %	Moderate
2026	9	21.7	11.99	-5.73	28.93	50 %	High
2026	10	23.0	16.2	-0.1	41.53	40 %	Moderate
2026	11	22.3	7.74	-4.85	48.57	15 %	Low
2026	12	18.5	2.05	-5.95	16.74	25 %	Moderate
2027	1	18.3	2.44	-8.15	39.93	10 %	Low
2027	2	22.8	43.36	-4.31	33.84	0 %	Low
2027	3	19.4	12.2	-7.38	42.97	20 %	Moderate
2027	4	21.7	37.55	-2.7	25.64	50 %	High
2027	5	23.8	48.14	-1.76	25.44	20 %	Moderate
2027	6	30.2	274.11	-2.54	57.97	20 %	Moderate
2027	7	36.5	264.94	-3.28	69.08	20 %	Moderate
2027	8	39.2	222.82	-1.67	60.58	30 %	Moderate
2027	9	23.6	32.64	-9.19	33.28	40 %	Moderate
2027	10	18.5	12.68	-2.13	33.58	5 %	Low
2027	11	21.4	1.88	-9.7	18.23	30 %	Moderate
2027	12	21.5	31.95	-3.68	30.93	20 %	Moderate
2028	1	16.3	2.64	-3.76	20.45	40 %	Moderate
2028	2	16.3	17.7	-6.54	22.55	60 %	High
2028	3	17.7	34.27	-2.18	38.11	50 %	High
2028	4	23.6	27.6	-7.78	40.25	0 %	Low
2028	5	22.5	29.79	-1.66	30.31	15 %	Low
2028	6	30.5	208.31	-1.95	97.59	60 %	High
2028	7	39.2	218.79	-3.29	93.51	40 %	Moderate
2028	8	32.9	268.63	-1.63	60.78	40 %	Moderate
2028	9	17.9	17.67	-4.89	12.32	40 %	Moderate
2028	10	17.0	37.51	-3.24	22.72	25 %	Moderate
2028	11	17.9	3.16	-8.71	12.36	30 %	Moderate
2028	12	21.8	35.64	-8.13	13.86	10 %	Low
2029	1	19.8	32.62	-5.99	10.15	25 %	Moderate
2029	2	24.1	48.77	-1.99	43.62	10 %	Low
2029	3	16.8	45.15	-2.78	14.65	5 %	Low
2029	4	19.6	49.86	-1.89	18.89	15 %	Low
2029	5	15.9	7.73	-8.08	36.43	10 %	Low
2029	6	35.2	219.14	-1.23	73.62	0 %	Low
2029	7	34.4	247.29	-4.62	96.68	50 %	High
2029	8	32.6	252.0	-3.82	95.01	15 %	Low
2029	9	23.9	20.24	-7.77	42.58	40 %	Moderate
2029	10	17.6	9.83	-1.74	38.47	5 %	Low
2029	11	17.1	1.52	-6.13	22.52	10 %	Low
2029	12	24.9	23.12	-0.91	49.01	60 %	High
2030	1	16.2	11.39	-2.59	46.75	20 %	Moderate
2030	2	16.7	44.6	-8.55	37.07	10 %	Low
2030	3	20.3	14.14	-2.97	31.62	15 %	Low
2030	4	16.0	10.52	-0.65	31.25	15 %	Low
2030	5	15.4	19.67	-7.74	22.47	40 %	Moderate
2030	6	34.0	224.87	-3.76	85.43	15 %	Low
2030	7	37.8	299.95	-3.02	79.93	5 %	Low
2030	8	31.1	216.73	-3.85	59.32	15 %	Low
2030	9	18.7	25.33	-6.68	42.26	20 %	Moderate
2030	10	19.7	11.63	-2.19	12.52	10 %	Low
2030	11	24.9	46.71	-5.76	24.75	40 %	Moderate
2030	12	22.8	0.8	-1.78	16.16	15 %	Low

Figure **Error! No text of specified style in document.**.16 Month-wise Prediction of Flood in Punjab for 2025-2030 using SVM

In 2029–2030, early months (January-April) exhibit lower risks, but summer months (June-August) see rising rainfall and moderate-to-high risks. December 2029 and 2030 also show moderate risks, highlighting the region's ongoing vulnerability to seasonal flooding.

7.5. Flood Prediction in Sindh Using Random Forest

Figure Error! **No text of specified style in document.**.17 shows flood risk predictions for Sindh from 2025 to 2030, based on variables like temperature, rainfall, ice levels, and vegetation.

Random Forest Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Random Forest (%)	Random Forest Risk
2025	1	22.3	18.1	-8.11	41.5	4 %	Low
2025	2	15.1	49.32	-1.16	39.78	2 %	Low
2025	3	21.8	35.26	-3.05	13.13	2 %	Low
2025	4	16.1	48.63	-5.43	14.75	2 %	Low
2025	5	19.1	12.11	-3.56	40.23	1 %	Low
2025	6	34.5	278.07	-3.01	76.6	18 %	Low
2025	7	34.0	278.21	-2.14	58.12	23 %	Moderate
2025	8	31.6	277.53	-3.63	99.81	23 %	Moderate
2025	9	19.3	22.24	-9.94	17.29	4 %	Low
2025	10	23.9	11.92	-9.04	20.83	1 %	Low
2025	11	15.9	49.84	-3.41	24.92	4 %	Low
2025	12	24.2	19.4	-7.18	16.06	4 %	Low
2026	1	15.5	38.48	-2.96	24.26	4 %	Low
2026	2	18.0	40.3	-1.98	10.12	4 %	Low
2026	3	24.7	5.65	-3.54	13.83	0 %	Low
2026	4	16.6	23.14	-9.57	36.57	2 %	Low
2026	5	25.0	13.29	-9.41	41.53	1 %	Low
2026	6	38.2	262.2	-1.35	97.83	22 %	Moderate
2026	7	33.3	216.65	-1.94	79.47	23 %	Moderate
2026	8	32.9	256.32	-2.52	63.36	23 %	Moderate
2026	9	18.0	11.47	-9.2	25.99	1 %	Low
2026	10	15.1	19.61	-0.55	45.63	4 %	Low
2026	11	19.4	5.4	-3.22	45.97	0 %	Low
2026	12	21.6	5.88	-6.84	20.37	0 %	Low
2027	1	17.5	49.87	-0.18	10.08	2 %	Low
2027	2	16.8	8.03	-2.8	35.15	0 %	Low
2027	3	22.0	31.93	-2.53	28.49	2 %	Low
2027	4	18.2	45.84	-9.83	21.14	2 %	Low
2027	5	16.1	27.2	-3.46	31.74	2 %	Low
2027	6	39.8	280.84	-2.05	88.9	22 %	Moderate
2027	7	31.9	275.34	-1.07	51.23	23 %	Moderate
2027	8	38.0	238.16	-3.31	51.07	27 %	Moderate
2027	9	21.5	39.16	-5.49	27.77	5 %	Low
2027	10	17.8	22.22	-5.34	49.25	4 %	Low
2027	11	16.9	3.02	-1.35	32.11	0 %	Low
2027	12	18.1	47.18	-7.86	32.64	4 %	Low
2028	1	19.0	39.04	-8.64	20.54	4 %	Low
2028	2	16.2	16.43	-0.08	30.53	23 %	Moderate
2028	3	23.0	1.37	-0.23	18.93	0 %	Low
2028	4	16.6	14.37	-2.02	46.64	1 %	Low
2028	5	22.7	4.0	-3.62	25.04	0 %	Low
2028	6	34.1	252.87	-3.12	50.3	18 %	Low
2028	7	34.9	202.45	-3.0	56.93	22 %	Moderate
2028	8	31.0	257.76	-2.46	96.72	23 %	Moderate
2028	9	17.8	34.28	-6.77	37.07	4 %	Low
2028	10	20.7	45.98	-5.92	27.28	4 %	Low
2028	11	15.6	36.19	-9.69	40.06	4 %	Low
2028	12	16.0	8.14	-2.54	38.18	0 %	Low
2029	1	15.0	25.99	-0.57	45.03	2 %	Low
2029	2	23.4	44.44	-2.89	42.32	2 %	Low
2029	3	19.3	43.85	-2.8	26.49	2 %	Low
2029	4	22.2	39.73	-4.65	40.45	4 %	Low
2029	5	18.7	49.2	-4.95	28.75	2 %	Low
2029	6	39.5	241.37	-1.32	65.13	22 %	Moderate
2029	7	38.5	256.26	-4.01	85.87	27 %	Moderate
2029	8	38.2	204.42	-4.43	75.35	27 %	Moderate
2029	9	18.3	46.1	-9.44	28.09	4 %	Low
2029	10	16.9	45.64	-6.41	49.09	4 %	Low
2029	11	21.2	8.63	-2.51	36.26	0 %	Low
2029	12	22.2	42.05	-9.53	23.05	4 %	Low
2030	1	18.8	10.03	-1.26	49.57	0 %	Low
2030	2	20.3	0.75	-2.29	29.38	0 %	Low
2030	3	15.0	15.37	-6.96	22.05	1 %	Low
2030	4	18.0	8.86	-2.75	22.92	0 %	Low
2030	5	17.1	4.05	-4.87	29.63	0 %	Low
2030	6	36.9	202.81	-3.31	82.63	25 %	Moderate
2030	7	30.6	207.12	-3.93	54.02	23 %	Moderate
2030	8	39.7	255.33	-3.18	85.09	27 %	Moderate
2030	9	17.0	17.39	-2.17	49.18	7 %	Low
2030	10	18.3	34.18	-6.58	45.91	4 %	Low
2030	11	22.3	29.26	-0.94	48.99	4 %	Low
2030	12	21.7	30.96	-5.83	29.24	4 %	Low

Figure **Error! No text of specified style in document.**.17 Month-wise Prediction of Flood in Sindh for 2025-2030 using Random Forest

The Random Forest model quantifies flood likelihood using historical data and projected trends. In early 2025, flood risks are low, with risk percentages ranging from 1% to 4%. January to March 2025, for example, show low risk, with temperatures between 15-22°C and modest rainfall (e.g., 18.1mm in January). However, from June to August 2025, temperatures rise to 34-34.5°C, and rainfall peaks (278.07mm in June and 278.21mm in July), increasing flood risk to moderate levels. This seasonal pattern persists, with summer months (June-August) consistently showing higher flood vulnerability due to high temperatures and heavy rainfall. Similar trends are observed in subsequent years. In 2026, June sees 262.2mm of rainfall and temperatures of 38.2°C, highlighting the role of extreme heat in exacerbating flood conditions. By 2027, July and August record rainfall between 275.34mm and 280.84mm, with temperatures around 38°C, maintaining moderate flood risks. In 2028, January to March show lower risks, but June to August remain moderate. Notably, February 2028 predicts moderate flood risk despite low rainfall (16.43mm), possibly due to other factors like ice levels or vegetation. By 2029 and 2030, early months continue to show low flood risk, while June to August maintain moderate risks, with temperatures exceeding 34°C and rainfall ranging from 202.45mm to 256.26mm. The data confirms that flood risks are highest during the warmer, wetter months, aligning with Sindh's monsoon season.

7.6. Flood Prediction in Sindh Using Decision Tree

Figure 4.14 presents Decision Tree predictions for flood risk in Sindh from 2025 to 2030, analyzing temperature, rainfall, ice, vegetation, and associated risks. June to August consistently show moderate to low flood risks despite high rainfall (200mm–280mm) and temperatures exceeding 30°C. For example, June and July 2025 record 278.07mm and 278.21mm of rainfall, respectively, yet flood risk remains "Low." This suggests regional flood mitigation infrastructure or adaptive measures may help manage these conditions. In contrast, months with lower rainfall, such as February and March 2025 (49.32mm and 5.65mm, respectively), show "None" or "Low" flood risks. Vegetation also plays a key role; higher vegetation levels (e.g., 76.6 in June 2025) correlate with lower flood risks, likely due to improved water absorption.

Decision Tree Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Decision Tree (%)	Decision Tree Risk
2025	1	22.3	18.1	-8.11	41.5	10 %	Low
2025	2	15.1	49.32	-1.16	39.78	0 %	None
2025	3	21.8	35.26	-3.05	13.13	10 %	Low
2025	4	16.1	48.63	-5.43	14.75	0 %	None
2025	5	19.1	12.11	-3.56	40.23	11 %	Low
2025	6	34.5	278.07	-3.01	76.6	10 %	Low
2025	7	34.0	278.21	-2.14	58.12	14 %	Low
2025	8	31.6	277.53	-3.63	99.81	10 %	Low
2025	9	19.3	22.24	-9.94	17.29	11 %	Low
2025	10	23.9	11.92	-9.04	20.83	10 %	Low
2025	11	15.9	49.84	-3.41	24.92	10 %	Low
2025	12	24.2	19.4	-7.18	16.06	0 %	None
2026	1	15.5	38.48	-2.96	24.26	10 %	Low
2026	2	18.0	40.3	-1.98	10.12	0 %	None
2026	3	24.7	5.65	-3.54	13.83	0 %	None
2026	4	16.6	23.14	-9.57	36.57	0 %	None
2026	5	25.0	13.29	-9.41	41.53	13 %	Low
2026	6	38.2	262.2	-1.35	97.83	13 %	Low
2026	7	33.3	216.65	-1.94	79.47	10 %	Low
2026	8	32.9	256.32	-2.52	63.36	11 %	Low
2026	9	18.0	11.47	-9.2	25.99	12 %	Low
2026	10	15.1	19.61	-0.55	45.63	10 %	Low
2026	11	19.4	5.4	-3.22	45.97	10 %	Low
2026	12	21.6	5.88	-6.84	20.37	11 %	Low
2027	1	17.5	49.87	-0.18	10.08	10 %	Low
2027	2	16.8	8.03	-2.8	35.15	10 %	Low
2027	3	22.0	31.93	-2.53	28.49	11 %	Low
2027	4	18.2	45.84	-9.83	21.14	10 %	Low
2027	5	16.1	27.2	-3.46	31.74	10 %	Low
2027	6	39.8	280.84	-2.05	88.9	10 %	Low
2027	7	31.9	275.34	-1.07	51.23	13 %	Low
2027	8	38.0	238.16	-3.31	51.07	10 %	Low
2027	9	21.5	39.16	-5.49	27.77	14 %	Low
2027	10	17.8	22.22	-5.34	49.25	10 %	Low
2027	11	16.9	3.02	-1.35	32.11	10 %	Low
2027	12	18.1	47.18	-7.86	32.64	11 %	Low
2028	1	19.0	39.04	-8.64	20.54	10 %	Low
2028	2	16.2	16.43	-0.08	30.53	0 %	None
2028	3	23.0	1.37	-0.23	18.93	13 %	Low
2028	4	16.6	14.37	-2.02	46.64	0 %	None
2028	5	22.7	4.0	-3.62	25.04	13 %	Low
2028	6	34.1	252.87	-3.12	50.3	10 %	Low
2028	7	34.9	202.45	-3.0	56.93	10 %	Low
2028	8	31.0	257.76	-2.46	96.72	0 %	None
2028	9	17.8	34.28	-6.77	37.07	10 %	Low
2028	10	20.7	45.98	-5.92	27.28	11 %	Low
2028	11	15.6	36.19	-9.69	40.06	0 %	None
2028	12	16.0	8.14	-2.54	38.18	10 %	Low
2029	1	15.0	25.99	-0.57	45.03	12 %	Low
2029	2	23.4	44.44	-2.89	42.32	10 %	Low
2029	3	19.3	43.85	-2.8	26.49	0 %	None
2029	4	22.2	39.73	-4.65	40.45	11 %	Low
2029	5	18.7	49.2	-4.95	28.75	13 %	Low
2029	6	39.5	241.37	-1.32	65.13	0 %	None
2029	7	38.5	256.26	-4.01	85.87	11 %	Low
2029	8	38.2	204.42	-4.43	75.35	10 %	Low
2029	9	18.3	46.1	-9.44	28.09	10 %	Low
2029	10	16.9	45.64	-6.41	49.09	0 %	None
2029	11	21.2	8.63	-2.51	36.26	13 %	Low
2029	12	22.2	42.05	-9.53	23.05	11 %	Low
2030	1	18.8	10.03	-1.26	49.57	14 %	Low
2030	2	20.3	0.75	-2.29	29.38	13 %	Low
2030	3	15.0	15.37	-6.96	22.05	13 %	Low
2030	4	18.0	8.86	-2.75	22.92	10 %	Low
2030	5	17.1	4.05	-4.87	29.63	0 %	None
2030	6	36.9	202.81	-3.31	82.63	14 %	Low
2030	7	30.6	207.12	-3.93	54.02	10 %	Low
2030	8	39.7	255.33	-3.18	85.09	10 %	Low
2030	9	17.0	17.39	-2.17	49.18	10 %	Low
2030	10	18.3	34.18	-6.58	45.91	10 %	Low
2030	11	22.3	29.26	-0.94	48.99	14 %	Low
2030	12	21.7	30.96	-5.83	29.24	13 %	Low

Figure **Error! No text of specified style in document.**.14 Month-wise Prediction of Flood in Sindh for 2025-2030 using Decision Tree

Temperature trends show gradual increases, particularly in later months, aligning with long-term warming. For instance, June 2029 reaches 38.2°C, yet flood risk remains "Low," possibly due to reduced runoff from drying effects. However, heavy rains during these periods could still pose risks if mitigation systems are inadequate.

7.7. Flood Prediction in Sindh Using Linear Regression

Figure Error! **No text of specified style in document.**.18 shows Linear Regression predictions for flood risks in Sindh from 2025 to 2030, focusing on temperature, rainfall, and other environmental variables. From 2025 to 2027, flood risks remain low, with risk percentages often below 10%. For example, January 2025 records 18.1mm of rainfall and a 5% risk. However, summer months (June-July) show significant rainfall spikes (e.g., 278.07mm in June 2025) and slightly higher risks, consistent with monsoon patterns. By 2028–2030, temperatures rise, with June and July reaching 38.2°C (2029) and 39.7°C (2030). These higher temperatures may intensify evaporation and rainfall, potentially overwhelming drainage systems. For instance, June 2029 records 241.37mm of rainfall but a 0% risk, suggesting other mitigating factors like vegetation or efficient water management. Months with moderate rainfall, such as February 2026 (40.3mm) and February 2029 (44.44mm), show minimal flood risks, indicating the influence of variables like soil moisture and vegetation. Higher vegetation levels, for example, help retain soil water, reducing runoff and flood risks.(*Integrating Geospatial and Machine Learning Approaches for Flood Risk Assessment and Disaster Management: A Case Study of Sheikhupura District, Pakistan (2022 – 2024)*)

7.8. Flood Prediction in Sindh Using SVM

Figure Error! **No text of specified style in document.**.19 shows SVM predictions for flood risks in Sindh from 2025 to 2030, analyzing temperature, rainfall, ice cover, vegetation, and risk classifications ("Low," "Moderate," or "High"). In 2025, temperatures range from 15.1°C to 34.5°C, with rainfall peaking in June (278.07mm) and July (278.21mm).

Linear Regression Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Linear Regression (%)	Linear Regression Risk
2025	1	22.3	18.1	-8.11	41.5	5 %	Low
2025	2	15.1	49.32	-1.16	39.78	3 %	Low
2025	3	21.8	35.26	-3.05	13.13	0 %	Low
2025	4	16.1	48.63	-5.43	14.75	2 %	Low
2025	5	19.1	12.11	-3.56	40.23	0 %	Low
2025	6	34.5	278.07	-3.01	76.6	7 %	Low
2025	7	34.0	278.21	-2.14	58.12	0 %	Low
2025	8	31.6	277.53	-3.63	99.81	6 %	Low
2025	9	19.3	22.24	-9.94	17.29	3 %	Low
2025	10	23.9	11.92	-9.04	20.83	0 %	Low
2025	11	15.9	49.84	-3.41	24.92	0 %	Low
2025	12	24.2	19.4	-7.18	16.06	0 %	Low
2026	1	15.5	38.48	-2.96	24.26	3 %	Low
2026	2	18.0	40.3	-1.98	10.12	1 %	Low
2026	3	24.7	5.65	-3.54	13.83	6 %	Low
2026	4	16.6	23.14	-9.57	36.57	8 %	Low
2026	5	25.0	13.29	-9.41	41.53	0 %	Low
2026	6	38.2	262.2	-1.35	97.83	2 %	Low
2026	7	33.3	216.65	-1.94	79.47	0 %	Low
2026	8	32.9	256.32	-2.52	63.36	7 %	Low
2026	9	18.0	11.47	-9.2	25.99	0 %	Low
2026	10	15.1	19.61	-0.55	45.63	0 %	Low
2026	11	19.4	5.4	-3.22	45.97	6 %	Low
2026	12	21.6	5.88	-6.84	20.37	0 %	Low
2027	1	17.5	49.87	-0.18	10.08	7 %	Low
2027	2	16.8	8.03	-2.8	35.15	1 %	Low
2027	3	22.0	31.93	-2.53	28.49	3 %	Low
2027	4	18.2	45.84	-9.83	21.14	6 %	Low
2027	5	16.1	27.2	-3.46	31.74	0 %	Low
2027	6	39.8	280.84	-2.05	88.9	4 %	Low
2027	7	31.9	275.34	-1.07	51.23	19 %	Low
2027	8	38.0	238.16	-3.31	51.07	0 %	Low
2027	9	21.5	39.16	-5.49	27.77	2 %	Low
2027	10	17.8	22.22	-5.34	49.25	0 %	Low
2027	11	16.9	3.02	-1.35	32.11	9 %	Low
2027	12	18.1	47.18	-7.86	32.64	2 %	Low
2028	1	19.0	39.04	-8.64	20.54	1 %	Low
2028	2	16.2	16.43	-0.08	30.53	13 %	Low
2028	3	23.0	1.37	-0.23	18.93	0 %	Low
2028	4	16.6	14.37	-2.02	46.64	0 %	Low
2028	5	22.7	4.0	-3.62	25.04	7 %	Low
2028	6	34.1	252.87	-3.12	50.3	8 %	Low
2028	7	34.9	202.45	-3.0	56.93	0 %	Low
2028	8	31.0	257.76	-2.46	96.72	6 %	Low
2028	9	17.8	34.28	-6.77	37.07	0 %	Low
2028	10	20.7	45.98	-5.92	27.28	2 %	Low
2028	11	15.6	36.19	-9.69	40.06	0 %	Low
2028	12	16.0	8.14	-2.54	38.18	0 %	Low
2029	1	15.0	25.99	-0.57	45.03	9 %	Low
2029	2	23.4	44.44	-2.89	42.32	0 %	Low
2029	3	19.3	43.85	-2.8	26.49	0 %	Low
2029	4	22.2	39.73	-4.65	40.45	6 %	Low
2029	5	18.7	49.2	-4.95	28.75	4 %	Low
2029	6	39.5	241.37	-1.32	65.13	0 %	Low
2029	7	38.5	256.26	-4.01	85.87	1 %	Low
2029	8	38.2	204.42	-4.43	75.35	1 %	Low
2029	9	18.3	46.1	-9.44	28.09	2 %	Low
2029	10	16.9	45.64	-6.41	49.09	0 %	Low
2029	11	21.2	8.63	-2.51	36.26	8 %	Low
2029	12	22.2	42.05	-9.53	23.05	0 %	Low
2030	1	18.8	10.03	-1.26	49.57	0 %	Low
2030	2	20.3	0.75	-2.29	29.38	2 %	Low
2030	3	15.0	15.37	-6.96	22.05	8 %	Low
2030	4	18.0	8.86	-2.75	22.92	5 %	Low
2030	5	17.1	4.05	-4.87	29.63	5 %	Low
2030	6	36.9	202.81	-3.31	82.63	0 %	Low
2030	7	30.6	207.12	-3.93	54.02	0 %	Low
2030	8	39.7	255.33	-3.18	85.09	3 %	Low
2030	9	17.0	17.39	-2.17	49.18	7 %	Low
2030	10	18.3	34.18	-6.58	45.91	0 %	Low
2030	11	22.3	29.26	-0.94	48.99	0 %	Low
2030	12	21.7	30.96	-5.83	29.24	3 %	Low

Figure **Error! No text of specified style in document.**.18 Month-wise Prediction of Flood in Sindh for 2025-2030 using Linear Regression

Despite heavy rainfall, flood risks remain moderate, suggesting factors like vegetation and temperature mitigate risks. By 2026, temperatures reach 38.2°C, and rainfall exceeds 250mm in June and July, resulting in “High” flood risks.

SVM Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	SVM (%)	SVM Risk
2025	1	22.3	18.1	-8.11	41.5	30 %	Moderate
2025	2	15.1	49.32	-1.16	39.78	40 %	Moderate
2025	3	21.8	35.26	-3.05	13.13	10 %	Low
2025	4	16.1	48.63	-5.43	14.75	25 %	Moderate
2025	5	19.1	12.11	-3.56	40.23	25 %	Moderate
2025	6	34.5	278.07	-3.01	76.6	15 %	Low
2025	7	34.0	278.21	-2.14	58.12	30 %	Moderate
2025	8	31.6	277.53	-3.63	99.81	30 %	Moderate
2025	9	19.3	22.24	-9.94	17.29	20 %	Moderate
2025	10	23.9	11.92	-9.04	20.83	20 %	Moderate
2025	11	15.9	49.84	-3.41	24.92	25 %	Moderate
2025	12	24.2	19.4	-7.18	16.06	40 %	Moderate
2026	1	15.5	38.48	-2.96	24.26	50 %	High
2026	2	18.0	40.3	-1.98	10.12	20 %	Moderate
2026	3	24.7	5.65	-3.54	13.83	20 %	Moderate
2026	4	16.6	23.14	-9.57	36.57	0 %	Low
2026	5	25.0	13.29	-9.41	41.53	60 %	High
2026	6	38.2	262.2	-1.35	97.83	60 %	High
2026	7	33.3	216.65	-1.94	79.47	0 %	Low
2026	8	32.9	256.32	-2.52	63.36	50 %	High
2026	9	18.0	11.47	-9.2	25.99	50 %	High
2026	10	15.1	19.61	-0.55	45.63	60 %	High
2026	11	19.4	5.4	-3.22	45.97	15 %	Low
2026	12	21.6	5.88	-6.84	20.37	0 %	Low
2027	1	17.5	49.87	-0.18	10.08	20 %	Moderate
2027	2	16.8	8.03	-2.8	35.15	60 %	High
2027	3	22.0	31.93	-2.53	28.49	0 %	Low
2027	4	18.2	45.84	-9.83	21.14	0 %	Low
2027	5	16.1	27.2	-3.46	31.74	50 %	High
2027	6	39.8	280.84	-2.05	88.9	0 %	Low
2027	7	31.9	275.34	-1.07	51.23	10 %	Low
2027	8	38.0	238.16	-3.31	51.07	40 %	Moderate
2027	9	21.5	39.16	-5.49	27.77	0 %	Low
2027	10	17.8	22.22	-5.34	49.25	10 %	Low
2027	11	16.9	3.02	-1.35	32.11	5 %	Low
2027	12	18.1	47.18	-7.86	32.64	25 %	Moderate
2028	1	19.0	39.04	-8.64	20.54	30 %	Moderate
2028	2	16.2	16.43	-0.08	30.53	10 %	Low
2028	3	23.0	1.37	-0.23	18.93	10 %	Low
2028	4	16.6	14.37	-2.02	46.64	0 %	Low
2028	5	22.7	4.0	-3.62	25.04	10 %	Low
2028	6	34.1	252.87	-3.12	50.3	25 %	Moderate
2028	7	34.9	202.45	-3.0	56.93	5 %	Low
2028	8	31.0	257.76	-2.46	96.72	50 %	High
2028	9	17.8	34.28	-6.77	37.07	10 %	Low
2028	10	20.7	45.98	-5.92	27.28	50 %	High
2028	11	15.6	36.19	-9.69	40.06	60 %	High
2028	12	16.0	8.14	-2.54	38.18	10 %	Low
2029	1	15.0	25.99	-0.57	45.03	10 %	Low
2029	2	23.4	44.44	-2.89	42.32	15 %	Low
2029	3	19.3	43.85	-2.8	26.49	0 %	Low
2029	4	22.2	39.73	-4.65	40.45	60 %	High
2029	5	18.7	49.2	-4.95	28.75	60 %	High
2029	6	39.5	241.37	-1.32	65.13	50 %	High
2029	7	38.5	256.26	-4.01	85.87	0 %	Low
2029	8	38.2	204.42	-4.43	75.35	50 %	High
2029	9	18.3	46.1	-9.44	28.09	20 %	Moderate
2029	10	16.9	45.64	-6.41	49.09	40 %	Moderate
2029	11	21.2	8.63	-2.51	36.26	15 %	Low
2029	12	22.2	42.05	-9.53	23.05	15 %	Low
2030	1	18.8	10.03	-1.26	49.57	15 %	Low
2030	2	20.3	0.75	-2.29	29.38	20 %	Moderate
2030	3	15.0	15.37	-6.96	22.05	15 %	Low
2030	4	18.0	8.86	-2.75	22.92	50 %	High
2030	5	17.1	4.05	-4.87	29.63	5 %	Low
2030	6	36.9	202.81	-3.31	82.63	30 %	Moderate
2030	7	30.6	207.12	-3.93	54.02	40 %	Moderate
2030	8	39.7	255.33	-3.18	85.09	15 %	Low
2030	9	17.0	17.39	-2.17	49.18	40 %	Moderate
2030	10	18.3	34.18	-6.58	45.91	0 %	Low
2030	11	22.3	29.26	-0.94	48.99	0 %	Low
2030	12	21.7	30.96	-5.83	29.24	25 %	Moderate

Figure **Error! No text of specified style in document.**.19 Month-wise Prediction of Flood in Sindh for 2025-2030 using SVM

Cooler months like November and December show "Low" risks, highlighting the seasonal nature of flood vulnerability. In 2027, June and July see temperatures up to 39.8°C and significant rainfall, maintaining "High" flood risks. Conversely, cooler months (January-March) show reduced risks due to lower rainfall and temperatures. By 2028–2030, summer months (June-August) continue to show “Moderate” to “High” risks, while drier months exhibit “Low” risks. Notably, June, July, and October 2030 show elevated risks despite reduced rainfall, suggesting increased sensitivity to temperature and vegetation changes. The SVM model highlights the strong link between summer rainfall, rising temperatures, and flood risks in Sindh.

7.9. Flood Prediction in Khyber Pakhtunkhwa Using Random Forest

Figure Error! **No text of specified style in document.**.20 illustrates future flood risk predictions for Sindh from 2025 to 2030 based on the Random Forest model. These predictions highlight significant seasonal variations in both environmental factors and flood risks, categorized from low to high. The model uses temperature, rainfall, ice, and vegetation data to estimate flooding probabilities. From January to April, temperatures remain low (16.4°C to 25°C), and rainfall is minimal (0.5 mm to 45 mm). During these months, flood risks are consistently low, generally below 5%, due to stable temperatures and limited rainfall, which prevent significant water accumulation and runoff. In contrast, May to September sees rising temperatures, with peaks reaching 39.1°C in June. Rainfall increases substantially, especially from June to August, where it ranges from 230 mm to 291 mm. Flood risks rise to moderate levels in early summer, peaking in July with rainfall of 279 mm, high temperatures, and considerable vegetation growth. These conditions, characterized by heavy precipitation and enhanced runoff, result in a heightened risk of flooding. From October to December, temperatures decrease and rainfall drops below 50 mm, causing flood risks to return to low levels. This pattern reflects stable conditions after the monsoon. Between 2025 and 2030, flood risks fluctuate based on changing weather patterns. Years like 2025 and 2026 show consistent low to moderate risks, while 2027 and 2029 experience slightly higher risks, especially in July, with rainfall reaching 279 mm to 291 mm.

Random Forest Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Random Forest (%)	Random Forest Risk
2025	1	24.0	45.52	-2.46	25.41	4 %	Low
2025	2	21.8	45.02	-7.56	23.86	5 %	Low
2025	3	16.6	38.28	-8.21	29.95	5 %	Low
2025	4	18.0	43.7	-9.07	12.17	5 %	Low
2025	5	20.6	27.49	-1.68	14.98	4 %	Low
2025	6	32.8	267.37	-1.81	84.64	36 %	Moderate
2025	7	33.2	230.48	-2.82	86.36	45 %	Moderate
2025	8	34.7	245.87	-3.49	77.41	45 %	Moderate
2025	9	22.9	33.17	-7.85	17.24	3 %	Low
2025	10	24.8	0.9	-3.23	33.8	2 %	Low
2025	11	24.1	40.48	-0.36	13.36	3 %	Low
2025	12	21.1	26.95	-5.49	19.55	2 %	Low
2026	1	21.1	3.72	-6.74	46.7	3 %	Low
2026	2	16.9	36.81	-3.93	30.07	5 %	Low
2026	3	18.5	7.91	-3.8	14.02	4 %	Low
2026	4	15.7	44.22	-7.03	14.16	5 %	Low
2026	5	23.2	32.53	-3.82	25.91	6 %	Low
2026	6	36.2	228.7	-3.16	53.61	43 %	Moderate
2026	7	35.7	201.76	-4.04	59.35	51 %	High
2026	8	33.9	276.24	-1.16	69.39	45 %	Moderate
2026	9	15.6	41.96	-9.11	46.1	3 %	Low
2026	10	24.3	3.03	-9.64	46.06	2 %	Low
2026	11	15.4	37.68	-5.27	24.95	3 %	Low
2026	12	18.8	17.98	-1.92	36.43	2 %	Low
2027	1	23.6	38.86	-6.58	11.1	4 %	Low
2027	2	15.5	1.22	-4.36	32.2	4 %	Low
2027	3	19.3	22.25	-6.2	29.32	4 %	Low
2027	4	25.0	3.49	-4.46	24.6	4 %	Low
2027	5	15.6	13.31	-4.41	19.7	4 %	Low
2027	6	38.2	208.99	-1.56	93.89	43 %	Moderate
2027	7	38.0	279.63	-2.43	99.34	51 %	High
2027	8	32.5	235.12	-1.48	56.49	40 %	Moderate
2027	9	16.5	13.69	-9.3	41.95	2 %	Low
2027	10	23.7	48.0	-4.63	38.85	3 %	Low
2027	11	22.3	31.6	-7.31	22.51	3 %	Low
2027	12	17.3	28.47	-0.54	36.91	2 %	Low
2028	1	20.4	33.52	-7.86	36.28	4 %	Low
2028	2	20.3	24.86	-2.92	11.65	4 %	Low
2028	3	19.8	25.09	-9.56	20.75	4 %	Low
2028	4	17.1	25.35	-7.89	19.35	4 %	Low
2028	5	24.4	47.16	-3.71	29.37	5 %	Low
2028	6	39.1	277.86	-2.77	91.78	43 %	Moderate
2028	7	38.4	275.23	-4.97	89.04	51 %	High
2028	8	32.4	224.37	-2.18	88.38	40 %	Moderate
2028	9	21.0	47.1	-0.93	37.36	3 %	Low
2028	10	15.8	41.57	-0.07	29.76	4 %	Low
2028	11	24.7	45.71	-3.49	25.0	3 %	Low
2028	12	18.4	47.45	-0.21	45.82	3 %	Low
2029	1	16.4	20.5	-5.07	19.35	3 %	Low
2029	2	18.1	37.79	-5.24	33.04	5 %	Low
2029	3	21.4	34.74	-4.91	34.11	5 %	Low
2029	4	21.9	46.86	-7.61	18.16	5 %	Low
2029	5	18.9	19.05	-3.83	46.03	4 %	Low
2029	6	36.6	210.35	-2.95	94.41	43 %	Moderate
2029	7	36.2	241.47	-3.57	72.7	51 %	High
2029	8	31.2	291.98	-2.01	82.4	40 %	Moderate
2029	9	23.0	0.53	-4.5	40.19	2 %	Low
2029	10	17.5	26.7	-5.21	10.81	2 %	Low
2029	11	22.7	7.95	-2.06	12.32	2 %	Low
2029	12	17.3	32.66	-3.38	47.97	3 %	Low
2030	1	17.0	12.46	-5.05	16.23	3 %	Low
2030	2	22.6	24.01	-5.96	31.88	4 %	Low
2030	3	21.9	26.21	-8.27	15.83	4 %	Low
2030	4	22.8	44.74	-0.73	30.68	5 %	Low
2030	5	18.7	15.96	-8.86	18.51	4 %	Low
2030	6	30.6	239.08	-4.31	65.93	38 %	Moderate
2030	7	35.7	233.08	-4.3	96.56	51 %	High
2030	8	36.4	264.67	-4.08	65.02	45 %	Moderate
2030	9	18.1	30.56	-6.18	27.58	3 %	Low
2030	10	18.9	34.78	-0.61	42.12	3 %	Low
2030	11	19.2	3.13	-8.76	10.96	2 %	Low
2030	12	21.9	46.67	-7.17	36.13	3 %	Low

Figure **Error! No text of specified style in document.**.20 Month-wise Prediction of Flood in KPK for 2025-2030 using Random Forest

Notably, the intensity of summer rainfall increases in 2027, 2028, and 2029, contributing to higher flood risks. Vegetation can mitigate flood risks, but during intense rainfall, its ability to absorb water is overwhelmed, as seen in June 2027, when rainfall reached 279.63 mm with 99.34% vegetation coverage, yet the flood risk remained high. The data shows an increasing trend in flood risk during warmer months, particularly between 2027 and 2029. This likely reflects broader climate changes, such as rising temperatures and more unpredictable rainfall, both of which drive the growing flood risks.

7.10. Flood Prediction in Khyber Pakhtunkhwa Using Decision Tree

Figure Error! **No text of specified style in document.**.21 presents future flood predictions for Khyber Pakhtunkhwa (KPK) from 2025 to 2030, based on the decision tree model. The model considers factors like temperature, rainfall, ice, vegetation, and flood risk, offering predictions categorized by likelihood percentages. The highest flood risks consistently occur between June and August, corresponding with the monsoon season. For example, June 2025 experiences 267.37 mm of rainfall and 32.8°C temperatures, resulting in a 46% flood risk. Similarly, July 2026 shows 201.76 mm of rainfall and 35.7°C temperatures, leading to a 50% flood risk. These trends continue in subsequent years, with rainfall exceeding 200 mm, correlating with higher flood risks. In August 2029, rainfall surges to 291.98 mm, further validating the link between heavy rainfall and increased flood risk. Conversely, January, February, and November show lower flood risks (0% to 14%), as rainfall remains below 50 mm, and temperatures average in the low 20s. For instance, January 2025 shows a rainfall of 45.52 mm and a temperature of 24°C, leading to no flood risk. Other periods, like January 2029 and November 2029, also experience no flood risk due to minimal rainfall. A trend of increasing temperatures is evident, particularly in 2026, when June reaches 36.2°C, and rainfall increases to

228.7 mm, leading to a 50% flood risk. This warming may exacerbate snowmelt from mountainous regions, contributing to runoff and higher flood risks.

Decision Tree Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Decision Tree (%)	Decision Tree Risk
2025	1	24.0	45.52	-2.46	25.41	0 %	None
2025	2	21.8	45.02	-7.56	23.86	11 %	Low
2025	3	16.6	38.28	-8.21	29.95	10 %	Low
2025	4	18.0	43.7	-9.07	12.17	10 %	Low
2025	5	20.6	27.49	-1.68	14.98	10 %	Low
2025	6	32.8	267.37	-1.81	84.64	46 %	High
2025	7	33.2	230.48	-2.82	86.36	0 %	None
2025	8	34.7	245.87	-3.49	77.41	50 %	High
2025	9	22.9	33.17	-7.85	17.24	10 %	Low
2025	10	24.8	0.9	-3.23	33.8	11 %	Low
2025	11	24.1	40.48	-0.36	13.36	13 %	Low
2025	12	21.1	26.95	-5.49	19.55	14 %	Low
2026	1	21.1	3.72	-6.74	46.7	0 %	None
2026	2	16.9	36.81	-3.93	30.07	10 %	Low
2026	3	18.5	7.91	-3.8	14.02	14 %	Low
2026	4	15.7	44.22	-7.03	14.16	10 %	Low
2026	5	23.2	32.53	-3.82	25.91	10 %	Low
2026	6	36.2	228.7	-3.16	53.61	50 %	High
2026	7	35.7	201.76	-4.04	59.35	50 %	High
2026	8	33.9	276.24	-1.16	69.39	45 %	High
2026	9	15.6	41.96	-9.11	46.1	11 %	Low
2026	10	24.3	3.03	-9.64	46.06	14 %	Low
2026	11	15.4	37.68	-5.27	24.95	0 %	None
2026	12	18.8	17.98	-1.92	36.43	10 %	Low
2027	1	23.6	38.86	-6.58	11.1	0 %	None
2027	2	15.5	1.22	-4.36	32.2	13 %	Low
2027	3	19.3	22.25	-6.2	29.32	10 %	Low
2027	4	25.0	3.49	-4.46	24.6	10 %	Low
2027	5	15.6	13.31	-4.41	19.7	12 %	Low
2027	6	38.2	208.99	-1.56	93.89	50 %	High
2027	7	38.0	279.63	-2.43	99.34	50 %	High
2027	8	32.5	235.12	-1.48	56.49	46 %	High
2027	9	16.5	13.69	-9.3	41.95	10 %	Low
2027	10	23.7	48.0	-4.63	38.85	10 %	Low
2027	11	22.3	31.6	-7.31	22.51	11 %	Low
2027	12	17.3	28.47	-0.54	36.91	0 %	None
2028	1	20.4	33.52	-7.86	36.28	12 %	Low
2028	2	20.3	24.86	-2.92	11.65	14 %	Low
2028	3	19.8	25.09	-9.56	20.75	10 %	Low
2028	4	17.1	25.35	-7.89	19.35	13 %	Low
2028	5	24.4	47.16	-3.71	29.37	0 %	None
2028	6	39.1	277.86	-2.77	91.78	50 %	High
2028	7	38.4	275.23	-4.97	89.04	0 %	None
2028	8	32.4	224.37	-2.18	88.38	0 %	None
2028	9	21.0	47.1	-0.93	37.36	10 %	Low
2028	10	15.8	41.57	-0.07	29.76	12 %	Low
2028	11	24.7	45.71	-3.49	25.0	10 %	Low
2028	12	18.4	47.45	-0.21	45.82	10 %	Low
2029	1	16.4	20.5	-5.07	19.35	0 %	None
2029	2	18.1	37.79	-5.24	33.04	10 %	Low
2029	3	21.4	34.74	-4.91	34.11	14 %	Low
2029	4	21.9	46.86	-7.61	18.16	13 %	Low
2029	5	18.9	19.05	-3.83	46.03	13 %	Low
2029	6	36.6	210.35	-2.95	94.41	0 %	None
2029	7	36.2	241.47	-3.57	72.7	48 %	High
2029	8	31.2	291.98	-2.01	82.4	50 %	High
2029	9	23.0	0.53	-4.5	40.19	0 %	None
2029	10	17.5	26.7	-5.21	10.81	10 %	Low
2029	11	22.7	7.95	-2.06	12.32	10 %	Low
2029	12	17.3	32.66	-3.38	47.97	10 %	Low
2030	1	17.0	12.46	-5.05	16.23	10 %	Low
2030	2	22.6	24.01	-5.96	31.88	10 %	Low
2030	3	21.9	26.21	-8.27	15.83	11 %	Low
2030	4	22.8	44.74	-0.73	30.68	10 %	Low
2030	5	18.7	15.96	-8.86	18.51	14 %	Low
2030	6	30.6	239.08	-4.31	65.93	50 %	High
2030	7	35.7	233.08	-4.3	96.56	46 %	High
2030	8	36.4	264.67	-4.08	65.02	46 %	High
2030	9	18.1	30.56	-6.18	27.58	0 %	None
2030	10	18.9	34.78	-0.61	42.12	0 %	None
2030	11	19.2	3.13	-8.76	10.96	14 %	Low
2030	12	21.9	46.67	-7.17	36.13	10 %	Low

Figure **Error! No text of specified style in document.**.21 Month-wise Prediction of Flood in KPK for 2025-2030 using Decision Tree

While the overall flood risk remains elevated in the summer, moderate risks occur during transition months like March, October, and December, thanks to moderate rainfall and stable temperatures. However, localized climatic events can significantly alter these predictions, emphasizing the need for adaptive flood management. Future flood predictions for KPK highlight the significant seasonal fluctuations, with peak risks from June to August due to heavy rainfall and rising temperatures. The increasing frequency and intensity of extreme weather events stress the need for effective flood mitigation strategies.

7.11. Flood Prediction in Khyber Pakhtunkhwa Using Linear Regression

Figure Error! **No text of specified style in document.**.22 displays flood risk predictions for KPK from 2025 to 2030 using linear regression. The analysis considers various environmental factors, including temperature, rainfall, ice coverage, vegetation, and calculated risk levels. The data reveals fluctuations in flood risk, with higher risks during the warmer months. For example, June and July 2025 show moderate risks with rainfall of 267.37 mm and 230.48 mm, alongside temperatures of 32.8°C and 33.2°C. This trend of moderate flood risk continues in subsequent years, particularly in the summer months, due to rising temperatures and increased precipitation. The highest flood risk occurs in 2028, with July experiencing 275.23 mm of rainfall and a 50% flood risk. This intense summer rainfall combined with high temperatures (38.4°C in July) underscores the region's vulnerability to flooding during this period. Similar patterns are seen in 2029, with high rainfall recorded in June (210.35 mm) and July (241.47 mm), maintaining moderate risks. In contrast, flood risks are generally lower during cooler months, from November to February, with rainfall significantly tapering off. For instance, January 2025 shows a low risk of 3%, due to moderate rainfall (45.52 mm) and cooler temperatures (24°C).

Cooler months like November and December show low flood risk unless unusual rainfall patterns or sudden ice melts occur. The trend toward increased flood risk in summer months suggests that rising global temperatures may be a key factor in extreme weather events.

Linear Regression Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Linear Regression (%)	Linear Regression Risk
2025	1	24.0	45.52	-2.46	25.41	3 %	Low
2025	2	21.8	45.02	-7.56	23.86	18 %	Low
2025	3	16.6	38.28	-8.21	29.95	36 %	Moderate
2025	4	18.0	43.7	-9.07	12.17	31 %	Moderate
2025	5	20.6	27.49	-1.68	14.98	0 %	Low
2025	6	32.8	267.37	-1.81	84.64	28 %	Moderate
2025	7	33.2	230.48	-2.82	86.36	21 %	Moderate
2025	8	34.7	245.87	-3.49	77.41	40 %	Moderate
2025	9	22.9	33.17	-7.85	17.24	18 %	Low
2025	10	24.8	0.9	-3.23	33.8	8 %	Low
2025	11	24.1	40.48	-0.36	13.36	0 %	Low
2025	12	21.1	26.95	-5.49	19.55	8 %	Low
2026	1	21.1	3.72	-6.74	46.7	13 %	Low
2026	2	16.9	36.81	-3.93	30.07	18 %	Low
2026	3	18.5	7.91	-3.8	14.02	0 %	Low
2026	4	15.7	44.22	-7.03	14.16	27 %	Moderate
2026	5	23.2	32.53	-3.82	25.91	8 %	Low
2026	6	36.2	228.7	-3.16	53.61	40 %	Moderate
2026	7	35.7	201.76	-4.04	59.35	29 %	Moderate
2026	8	33.9	276.24	-1.16	69.39	21 %	Moderate
2026	9	15.6	41.96	-9.11	46.1	33 %	Moderate
2026	10	24.3	3.03	-9.64	46.06	38 %	Moderate
2026	11	15.4	37.68	-5.27	24.95	18 %	Low
2026	12	18.8	17.98	-1.92	36.43	6 %	Low
2027	1	23.6	38.86	-6.58	11.1	18 %	Low
2027	2	15.5	1.22	-4.36	32.2	14 %	Low
2027	3	19.3	22.25	-6.2	29.32	26 %	Moderate
2027	4	25.0	3.49	-4.46	24.6	13 %	Low
2027	5	15.6	13.31	-4.41	19.7	3 %	Low
2027	6	38.2	208.99	-1.56	93.89	13 %	Low
2027	7	38.0	279.63	-2.43	99.34	35 %	Moderate
2027	8	32.5	235.12	-1.48	56.49	31 %	Moderate
2027	9	16.5	13.69	-9.3	41.95	32 %	Moderate
2027	10	23.7	48.0	-4.63	38.85	22 %	Moderate
2027	11	22.3	31.6	-7.31	22.51	30 %	Moderate
2027	12	17.3	28.47	-0.54	36.91	0 %	Low
2028	1	20.4	33.52	-7.86	36.28	19 %	Low
2028	2	20.3	24.86	-2.92	11.65	10 %	Low
2028	3	19.8	25.09	-9.56	20.75	25 %	Moderate
2028	4	17.1	25.35	-7.89	19.35	26 %	Moderate
2028	5	24.4	47.16	-3.71	29.37	9 %	Low
2028	6	39.1	277.86	-2.77	91.78	26 %	Moderate
2028	7	38.4	275.23	-4.97	89.04	50 %	High
2028	8	32.4	224.37	-2.18	88.38	32 %	Moderate
2028	9	21.0	47.1	-0.93	37.36	0 %	Low
2028	10	15.8	41.57	-0.07	29.76	8 %	Low
2028	11	24.7	45.71	-3.49	25.0	19 %	Low
2028	12	18.4	47.45	-0.21	45.82	8 %	Low
2029	1	16.4	20.5	-5.07	19.35	10 %	Low
2029	2	18.1	37.79	-5.24	33.04	10 %	Low
2029	3	21.4	34.74	-4.91	34.11	17 %	Low
2029	4	21.9	46.86	-7.61	18.16	35 %	Moderate
2029	5	18.9	19.05	-3.83	46.03	2 %	Low
2029	6	36.6	210.35	-2.95	94.41	18 %	Low
2029	7	36.2	241.47	-3.57	72.7	31 %	Moderate
2029	8	31.2	291.98	-2.01	82.4	39 %	Moderate
2029	9	23.0	0.53	-4.5	40.19	6 %	Low
2029	10	17.5	26.7	-5.21	10.81	19 %	Low
2029	11	22.7	7.95	-2.06	12.32	0 %	Low
2029	12	17.3	32.66	-3.38	47.97	3 %	Low
2030	1	17.0	12.46	-5.05	16.23	11 %	Low
2030	2	22.6	24.01	-5.96	31.88	22 %	Moderate
2030	3	21.9	26.21	-8.27	15.83	25 %	Moderate
2030	4	22.8	44.74	-0.73	30.68	8 %	Low
2030	5	18.7	15.96	-8.86	18.51	29 %	Moderate
2030	6	30.6	239.08	-4.31	65.93	30 %	Moderate
2030	7	35.7	233.08	-4.3	96.56	41 %	Moderate
2030	8	36.4	264.67	-4.08	65.02	47 %	Moderate
2030	9	18.1	30.56	-6.18	27.58	12 %	Low
2030	10	18.9	34.78	-0.61	42.12	5 %	Low
2030	11	19.2	3.13	-8.76	10.96	26 %	Moderate
2030	12	21.9	46.67	-7.17	36.13	32 %	Moderate

Figure **Error! No text of specified style in document.**.22 Month-wise Prediction of Flood in KPK for 2025-2030 using Linear Regression

Additionally, diminishing ice coverage, particularly in high-altitude areas, may contribute to flooding when snowmelt combines with heavy rainfall. Vegetation levels fluctuate, with higher vegetation correlating with a higher risk of floods, as it retains more surface water. For example, June 2025 shows high vegetation (84.64%) alongside significant rainfall, leading to moderate flood risk. The combination of rising temperatures, intense rainfall, diminishing ice coverage, and vegetation patterns indicates a growing vulnerability to flooding in KPK, especially during the warmer months. Strategic flood management and early warning systems are crucial to mitigate the impacts of these predicted risks.

7.12. Flood Prediction in Khyber Pakhtunkhwa Using SVM

Figure Error! **No text of specified style in document.**.23 provides flood predictions for KPK from 2025 to 2030 based on support vector machine (SVM) analysis. This model incorporates temperature, rainfall, ice levels, and vegetation, which all significantly influence flood risks. In early 2025, January shows a high flood risk (60%), with a temperature of 24°C and rainfall of 45.52 mm. This pattern continues into early 2026, with high flood risks in January, February, and March, driven by increased temperatures, significant rainfall, and changes in ice levels. These months typically exhibit higher rainfall and lower ice levels, suggesting the potential for rapid runoff and higher flood risks. As 2026 and 2027 progress, flood risk moderates in April and May but remains substantial in June and July due to extreme temperatures and heavy rainfall. For example, June and July in 2026 and 2027 show high temperatures (35-38°C) and rainfall (200-270 mm), although the risk may be low due to mitigating factors like vegetation growth, which helps absorb excess water. In 2028, the flood risk pattern becomes more variable. February, March, and April show relatively high risks, but these decrease in May due to changes in temperature and rainfall.

Despite moderate temperatures in February and April, rainfall intensity and ice levels suggest higher risks. These predictions highlight a fluctuating yet growing trend of flood risks in KPK, driven by rising temperatures, extreme rainfall, and diminishing ice coverage.

SVM Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	SVM (%)	SVM Risk
2025	1	24.0	45.52	-2.46	25.41	60 %	High
2025	2	21.8	45.02	-7.56	23.86	30 %	Moderate
2025	3	16.6	38.28	-8.21	29.95	40 %	Moderate
2025	4	18.0	43.7	-9.07	12.17	20 %	Moderate
2025	5	20.6	27.49	-1.68	14.98	30 %	Moderate
2025	6	32.8	267.37	-1.81	84.64	20 %	Moderate
2025	7	33.2	230.48	-2.82	86.36	5 %	Low
2025	8	34.7	245.87	-3.49	77.41	5 %	Low
2025	9	22.9	33.17	-7.85	17.24	20 %	Moderate
2025	10	24.8	0.9	-3.23	33.8	10 %	Low
2025	11	24.1	40.48	-0.36	13.36	25 %	Moderate
2025	12	21.1	26.95	-5.49	19.55	20 %	Moderate
2026	1	21.1	3.72	-6.74	46.7	50 %	High
2026	2	16.9	36.81	-3.93	30.07	50 %	High
2026	3	18.5	7.91	-3.8	14.02	50 %	High
2026	4	15.7	44.22	-7.03	14.16	5 %	Low
2026	5	23.2	32.53	-3.82	25.91	50 %	High
2026	6	36.2	228.7	-3.16	53.61	25 %	Moderate
2026	7	35.7	201.76	-4.04	59.35	15 %	Low
2026	8	33.9	276.24	-1.16	69.39	15 %	Low
2026	9	15.6	41.96	-9.11	46.1	0 %	Low
2026	10	24.3	3.03	-9.64	46.06	15 %	Low
2026	11	15.4	37.68	-5.27	24.95	60 %	High
2026	12	18.8	17.98	-1.92	36.43	40 %	Moderate
2027	1	23.6	38.86	-6.58	11.1	30 %	Moderate
2027	2	15.5	1.22	-4.36	32.2	40 %	Moderate
2027	3	19.3	22.25	-6.2	29.32	20 %	Moderate
2027	4	25.0	3.49	-4.46	24.6	25 %	Moderate
2027	5	15.6	13.31	-4.41	19.7	10 %	Low
2027	6	38.2	208.99	-1.56	93.89	5 %	Low
2027	7	38.0	279.63	-2.43	99.34	10 %	Low
2027	8	32.5	235.12	-1.48	56.49	25 %	Moderate
2027	9	16.5	13.69	-9.3	41.95	0 %	Low
2027	10	23.7	48.0	-4.63	38.85	25 %	Moderate
2027	11	22.3	31.6	-7.31	22.51	10 %	Low
2027	12	17.3	28.47	-0.54	36.91	40 %	Moderate
2028	1	20.4	33.52	-7.86	36.28	0 %	Low
2028	2	20.3	24.86	-2.92	11.65	60 %	High
2028	3	19.8	25.09	-9.56	20.75	10 %	Low
2028	4	17.1	25.35	-7.89	19.35	60 %	High
2028	5	24.4	47.16	-3.71	29.37	15 %	Low
2028	6	39.1	277.86	-2.77	91.78	10 %	Low
2028	7	38.4	275.23	-4.97	89.04	30 %	Moderate
2028	8	32.4	224.37	-2.18	88.38	15 %	Low
2028	9	21.0	47.1	-0.93	37.36	20 %	Moderate
2028	10	15.8	41.57	-0.07	29.76	0 %	Low
2028	11	24.7	45.71	-3.49	25.0	50 %	High
2028	12	18.4	47.45	-0.21	45.82	10 %	Low
2029	1	16.4	20.5	-5.07	19.35	40 %	Moderate
2029	2	18.1	37.79	-5.24	33.04	40 %	Moderate
2029	3	21.4	34.74	-4.91	34.11	5 %	Low
2029	4	21.9	46.86	-7.61	18.16	10 %	Low
2029	5	18.9	19.05	-3.83	46.03	30 %	Moderate
2029	6	36.6	210.35	-2.95	94.41	25 %	Moderate
2029	7	36.2	241.47	-3.57	72.7	30 %	Moderate
2029	8	31.2	291.98	-2.01	82.4	10 %	Low
2029	9	23.0	0.53	-4.5	40.19	15 %	Low
2029	10	17.5	26.7	-5.21	10.81	5 %	Low
2029	11	22.7	7.95	-2.06	12.32	20 %	Moderate
2029	12	17.3	32.66	-3.38	47.97	0 %	Low
2030	1	17.0	12.46	-5.05	16.23	20 %	Moderate
2030	2	22.6	24.01	-5.96	31.88	40 %	Moderate
2030	3	21.9	26.21	-8.27	15.83	50 %	High
2030	4	22.8	44.74	-0.73	30.68	50 %	High
2030	5	18.7	15.96	-8.86	18.51	40 %	Moderate
2030	6	30.6	239.08	-4.31	65.93	60 %	High
2030	7	35.7	233.08	-4.3	96.56	0 %	Low
2030	8	36.4	264.67	-4.08	65.02	10 %	Low
2030	9	18.1	30.56	-6.18	27.58	25 %	Moderate
2030	10	18.9	34.78	-0.61	42.12	5 %	Low
2030	11	19.2	3.13	-8.76	10.96	40 %	Moderate
2030	12	21.9	46.67	-7.17	36.13	5 %	Low

Figure **Error! No text of specified style in document.**.23 Month-wise Prediction of Flood in KPK for 2025-2030 using SVM

The data underscores the need for comprehensive flood management strategies to mitigate the increasing flood risks over the next decade.

7.13. Flood Prediction in Balochistan Using Random Forest

The flood prediction data for Balochistan from 2025 to 2030, based on the Random Forest model, outlines key factors influencing flood risks, as shown in Figure Error! No text of specified style in document.**.24.** Key parameters such as temperature, rainfall, ice, vegetation, and risk predictions offer valuable insights into future flood risks.

Flood risk tends to fluctuate, showing moderate levels in the first quarter each year. For instance, January 2025 starts with a moderate risk of 24%, with temperatures of 17.5°C and rainfall of 32.33 mm. As the year progresses, risk remains moderate, especially in summer with temperatures above 37°C and substantial rainfall in June, July, and August. The highest rainfall projections, 253.21 mm in July 2025 and 289.1 mm in August 2025, push the risk to 42%. This trend of moderate to high flood risk during the warmer months continues through 2030.

Low-risk months generally occur during winter with cooler temperatures and less rainfall. For example, February and November 2025 have low risks, 16% and 18%, respectively, with rainfall of 2.56 mm in February and 11.97 mm in November. This pattern is consistent through 2026-2030, with lower risks in cooler months.

Years 2026-2027 show moderate risks, particularly in June-August, but risk decreases slightly by late 2027 and early 2028 due to stable rainfall and temperatures. Notable peaks in rainfall (262.01 mm in August 2026 and 265.59 mm in June 2028) still correlate with higher flood risks. By December 2030, the risk drops to 12%, the lowest observed throughout the 2025-2030 period, likely due to cooler temperatures and reduced rainfall.

A key trend is the rise in flood risk from May to September due to increased rainfall and peak temperatures. The risk consistently hovers between 40% and 42% during these months. Summer months are critical for flood preparedness, especially with high rainfall and temperature patterns.

Random Forest Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Random Forest (%)	Random Forest Risk
2025	1	17.5	32.33	-6.7	35.67	24 %	Moderate
2025	2	18.0	2.56	-8.95	32.29	16 %	Low
2025	3	18.0	35.26	-4.47	10.72	34 %	Moderate
2025	4	23.7	15.04	-2.99	41.11	17 %	Low
2025	5	19.4	31.66	-7.94	22.99	28 %	Moderate
2025	6	37.3	218.63	-3.04	53.5	40 %	Moderate
2025	7	38.7	253.21	-1.85	97.75	42 %	Moderate
2025	8	37.8	289.1	-3.93	60.44	40 %	Moderate
2025	9	17.2	32.88	-2.67	19.11	31 %	Moderate
2025	10	21.6	36.92	-8.33	30.16	33 %	Moderate
2025	11	22.8	11.97	-3.41	13.65	18 %	Low
2025	12	22.8	5.66	-3.25	44.48	11 %	Low
2026	1	17.5	39.69	-0.54	24.46	28 %	Moderate
2026	2	21.9	18.8	-8.78	16.7	23 %	Moderate
2026	3	16.6	24.97	-7.93	35.59	23 %	Moderate
2026	4	23.3	25.03	-7.4	12.4	23 %	Moderate
2026	5	15.6	46.51	-8.14	29.23	35 %	Moderate
2026	6	35.5	239.07	-2.94	79.65	39 %	Moderate
2026	7	35.3	211.29	-3.59	88.98	40 %	Moderate
2026	8	39.8	262.01	-4.55	90.18	40 %	Moderate
2026	9	24.1	5.35	-0.49	21.94	6 %	Low
2026	10	24.0	46.05	-0.5	21.33	28 %	Moderate
2026	11	17.7	24.27	-1.49	31.2	24 %	Moderate
2026	12	15.2	45.03	-7.53	14.72	32 %	Moderate
2027	1	17.3	42.63	-7.2	44.63	28 %	Moderate
2027	2	22.5	9.48	-9.5	38.21	18 %	Low
2027	3	20.8	45.15	-6.76	27.7	35 %	Moderate
2027	4	23.3	9.44	-6.15	40.83	13 %	Low
2027	5	16.2	33.42	-6.35	12.86	31 %	Moderate
2027	6	39.0	209.35	-4.41	87.24	41 %	Moderate
2027	7	39.9	220.11	-2.39	85.1	42 %	Moderate
2027	8	31.3	280.22	-2.91	56.06	38 %	Moderate
2027	9	24.5	30.89	-3.0	45.51	21 %	Moderate
2027	10	16.6	24.46	-4.24	27.77	23 %	Moderate
2027	11	24.1	40.03	-1.03	24.9	28 %	Moderate
2027	12	24.3	1.02	-6.63	19.62	6 %	Low
2028	1	17.3	29.67	-3.82	13.43	22 %	Moderate
2028	2	22.0	20.26	-9.68	17.52	24 %	Moderate
2028	3	18.1	24.63	-6.81	49.48	23 %	Moderate
2028	4	22.6	27.83	-3.96	31.28	20 %	Moderate
2028	5	21.4	36.44	-2.45	14.87	36 %	Moderate
2028	6	39.6	265.59	-1.27	66.52	41 %	Moderate
2028	7	39.2	260.08	-1.9	54.83	42 %	Moderate
2028	8	37.6	241.56	-1.72	94.69	40 %	Moderate
2028	9	15.6	30.63	-3.16	27.11	28 %	Moderate
2028	10	24.3	43.04	-7.7	22.67	28 %	Moderate
2028	11	20.9	26.87	-4.01	33.61	23 %	Moderate
2028	12	19.5	44.76	-7.58	35.14	32 %	Moderate
2029	1	20.5	14.2	-8.08	37.74	20 %	Moderate
2029	2	17.5	34.46	-1.65	34.76	31 %	Moderate
2029	3	18.3	22.62	-6.7	23.35	23 %	Moderate
2029	4	24.5	19.37	-3.96	25.59	14 %	Low
2029	5	19.0	11.8	-3.88	39.76	20 %	Moderate
2029	6	39.8	238.53	-3.09	58.44	41 %	Moderate
2029	7	37.4	244.74	-1.88	99.33	41 %	Moderate
2029	8	37.7	223.58	-3.39	57.0	40 %	Moderate
2029	9	21.2	9.23	-9.77	44.6	14 %	Low
2029	10	21.5	24.39	-0.47	47.4	24 %	Moderate
2029	11	16.2	3.32	-3.51	27.21	13 %	Low
2029	12	16.8	40.22	-5.51	39.96	32 %	Moderate
2030	1	20.4	4.38	-4.73	14.55	17 %	Low
2030	2	19.2	17.95	-2.47	13.58	23 %	Moderate
2030	3	16.5	17.08	-0.13	16.6	45 %	Moderate
2030	4	17.7	30.74	-9.98	46.79	28 %	Moderate
2030	5	23.1	38.02	-2.23	28.46	35 %	Moderate
2030	6	37.0	253.88	-2.3	76.42	40 %	Moderate
2030	7	32.6	216.89	-3.26	86.51	40 %	Moderate
2030	8	33.7	221.06	-4.76	64.38	39 %	Moderate
2030	9	19.4	19.51	-0.45	25.5	23 %	Moderate
2030	10	19.3	38.21	-7.71	18.37	33 %	Moderate
2030	11	19.8	36.45	-5.49	47.83	33 %	Moderate
2030	12	23.1	9.12	-3.4	22.49	12 %	Low

Figure Error! No text of specified style in document.**.24** Month-wise Prediction of Flood in Balochistan for 2025-2030 using Random Forest

This data indicates Balochistan will likely experience periodic floods, with high risks during the summer monsoon season. Year-to-year consistency suggests that flood mitigation efforts should prioritize the summer months, accounting for increased rainfall and temperatures.

7.14. Flood Prediction in Balochistan Using Decision Tree

The decision tree model for flood predictions in Balochistan from 2025-2030 reveals low to medium risks, with some months showing higher risks, as shown in Figure Error! **No text of specified style in document.**.25. A notable trend is that flood risk remains low for most of the period, with some months showing “None” risk, particularly in dry months like September 2025 and February 2026. Even during high rainfall months (June-August), the flood risk remains low, with rainfall exceeding 200 mm, particularly in July and August 2025, peaking at 289.1 mm. This suggests that heavy rainfall alone may not necessarily trigger higher risks. However, a significant rise in risk occurs in March 2030, where rainfall of 17.08 mm leads to a “High” risk, reaching 50%. This spike highlights the unpredictable nature of floods in the region and the need for vigilance. Temperature trends show higher temperatures from March to August, peaking at 39.8°C in August 2026 and 39.9°C in July 2027. Despite high temperatures, the decision tree model indicates that flood risks remain low due to other factors such as terrain and infrastructure. While low risks dominate, occasional spikes in risk – especially during high rainfall months and extreme temperatures – emphasize the need for ongoing monitoring and preparedness.

7.15. Flood Prediction in Balochistan Using Linear Regression

The climate data for Balochistan from 2025 to 2030 shows fluctuations in temperature and rainfall, impacting flood risks, as seen in Figure Error! No text of specified style in document.**.26**. Higher temperatures, particularly in the summer, contribute to increased evaporation and potential flash floods. From June to August, temperatures remain above 35°C, peaking at 39°C in

July and August. Rainfall spikes, especially in June-August, correlate with moderate to high flood risks.

Decision Tree Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Decision Tree (%)	Decision Tree Risk
2025	1	17.5	32.33	-6.7	35.67	10 %	Low
2025	2	18.0	2.56	-8.95	32.29	10 %	Low
2025	3	18.0	35.26	-4.47	10.72	10 %	Low
2025	4	23.7	15.04	-2.99	41.11	10 %	Low
2025	5	19.4	31.66	-7.94	22.99	10 %	Low
2025	6	37.3	218.63	-3.04	53.5	10 %	Low
2025	7	38.7	253.21	-1.85	97.75	10 %	Low
2025	8	37.8	289.1	-3.93	60.44	12 %	Low
2025	9	17.2	32.88	-2.67	19.11	0 %	None
2025	10	21.6	36.92	-8.33	30.16	13 %	Low
2025	11	22.8	11.97	-3.41	13.65	14 %	Low
2025	12	22.8	5.66	-3.25	44.48	11 %	Low
2026	1	17.5	39.69	-0.54	24.46	14 %	Low
2026	2	21.9	18.8	-8.78	16.7	0 %	None
2026	3	16.6	24.97	-7.93	35.59	10 %	Low
2026	4	23.3	25.03	-7.4	12.4	10 %	Low
2026	5	15.6	46.51	-8.14	29.23	10 %	Low
2026	6	35.5	239.07	-2.94	79.65	11 %	Low
2026	7	35.3	211.29	-3.59	88.98	11 %	Low
2026	8	39.8	262.01	-4.55	90.18	10 %	Low
2026	9	24.1	5.35	-0.49	21.94	10 %	Low
2026	10	24.0	46.05	-0.5	21.33	10 %	Low
2026	11	17.7	24.27	-1.49	31.2	0 %	None
2026	12	15.2	45.03	-7.53	14.72	10 %	Low
2027	1	17.3	42.63	-7.2	44.63	10 %	Low
2027	2	22.5	9.48	-9.5	38.21	10 %	Low
2027	3	20.8	45.15	-6.76	27.7	14 %	Low
2027	4	23.3	9.44	-6.15	40.83	10 %	Low
2027	5	16.2	33.42	-6.35	12.86	13 %	Low
2027	6	39.0	209.35	-4.41	87.24	10 %	Low
2027	7	39.9	220.11	-2.39	85.1	10 %	Low
2027	8	31.3	280.22	-2.91	56.06	0 %	None
2027	9	24.5	30.89	-3.0	45.51	10 %	Low
2027	10	16.6	24.46	-4.24	27.77	10 %	Low
2027	11	24.1	40.03	-1.03	24.9	13 %	Low
2027	12	24.3	1.02	-6.63	19.62	10 %	Low
2028	1	17.3	29.67	-3.82	13.43	10 %	Low
2028	2	22.0	20.26	-9.68	17.52	10 %	Low
2028	3	18.1	24.63	-6.81	49.48	0 %	None
2028	4	22.6	27.83	-3.96	31.28	10 %	Low
2028	5	21.4	36.44	-2.45	14.87	0 %	None
2028	6	39.6	265.59	-1.27	66.52	10 %	Low
2028	7	39.2	260.08	-1.9	54.83	10 %	Low
2028	8	37.6	241.56	-1.72	94.69	0 %	None
2028	9	15.6	30.63	-3.16	27.11	0 %	None
2028	10	24.3	43.04	-7.7	22.67	10 %	Low
2028	11	20.9	26.87	-4.01	33.61	0 %	None
2028	12	19.5	44.76	-7.58	35.14	10 %	Low
2029	1	20.5	14.2	-8.08	37.74	10 %	Low
2029	2	17.5	34.46	-1.65	34.76	10 %	Low
2029	3	18.3	22.62	-6.7	23.35	0 %	None
2029	4	24.5	19.37	-3.96	25.59	10 %	Low
2029	5	19.0	11.8	-3.88	39.76	10 %	Low
2029	6	39.8	238.53	-3.09	58.44	10 %	Low
2029	7	37.4	244.74	-1.88	99.33	10 %	Low
2029	8	37.7	223.58	-3.39	57.0	14 %	Low
2029	9	21.2	9.23	-9.77	44.6	10 %	Low
2029	10	21.5	24.39	-0.47	47.4	10 %	Low
2029	11	16.2	3.32	-3.51	27.21	10 %	Low
2029	12	16.8	40.22	-5.51	39.96	12 %	Low
2030	1	20.4	4.38	-4.73	14.55	11 %	Low
2030	2	19.2	17.95	-2.47	13.58	0 %	None
2030	3	16.5	17.08	-0.13	16.6	50 %	High
2030	4	17.7	30.74	-9.98	46.79	11 %	Low
2030	5	23.1	38.02	-2.23	28.46	10 %	Low
2030	6	37.0	253.88	-2.3	76.42	0 %	None
2030	7	32.6	216.89	-3.26	86.51	11 %	Low
2030	8	33.7	221.06	-4.76	64.38	10 %	Low
2030	9	19.4	19.51	-0.45	25.5	10 %	Low
2030	10	19.3	38.21	-7.71	18.37	0 %	None
2030	11	19.8	36.45	-5.49	47.83	10 %	Low
2030	12	23.1	9.12	-3.4	22.49	0 %	None

Figure **Error! No text of specified style in document.**.25 Month-wise Prediction of Flood in Balochistan for 2025-2030 using Decision Tree

For instance, June 2025 sees 218.63 mm of rainfall, and July 2026 reaches 253.21 mm, with flood risk peaking at 92%.

Linear Regression Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	Linear Regression (%)	Linear Regression Risk
2025	1	17.5	32.33	-6.7	35.67	0 %	Low
2025	2	18.0	2.56	-8.95	32.29	0 %	Low
2025	3	18.0	35.26	-4.47	10.72	0 %	Low
2025	4	23.7	15.04	-2.99	41.11	2 %	Low
2025	5	19.4	31.66	-7.94	22.99	0 %	Low
2025	6	37.3	218.63	-3.04	53.5	30 %	Moderate
2025	7	38.7	253.21	-1.85	97.75	92 %	High
2025	8	37.8	289.1	-3.93	60.44	10 %	Low
2025	9	17.2	32.88	-2.67	19.11	5 %	Low
2025	10	21.6	36.92	-8.33	30.16	1 %	Low
2025	11	22.8	11.97	-3.41	13.65	4 %	Low
2025	12	22.8	5.66	-3.25	44.48	0 %	Low
2026	1	17.5	39.69	-0.54	24.46	43 %	Moderate
2026	2	21.9	18.8	-8.78	16.7	0 %	Low
2026	3	16.6	24.97	-7.93	35.59	0 %	Low
2026	4	23.3	25.03	-7.4	12.4	0 %	Low
2026	5	15.6	46.51	-8.14	29.23	0 %	Low
2026	6	35.5	239.07	-2.94	79.65	28 %	Moderate
2026	7	35.3	211.29	-3.59	88.98	0 %	Low
2026	8	39.8	262.01	-4.55	90.18	5 %	Low
2026	9	24.1	5.35	-0.49	21.94	24 %	Moderate
2026	10	24.0	46.05	-0.5	21.33	52 %	High
2026	11	17.7	24.27	-1.49	31.2	6 %	Low
2026	12	15.2	45.03	-7.53	14.72	0 %	Low
2027	1	17.3	42.63	-7.2	44.63	0 %	Low
2027	2	22.5	9.48	-9.5	38.21	0 %	Low
2027	3	20.8	45.15	-6.76	27.7	0 %	Low
2027	4	23.3	9.44	-6.15	40.83	0 %	Low
2027	5	16.2	33.42	-6.35	12.86	0 %	Low
2027	6	39.0	209.35	-4.41	87.24	9 %	Low
2027	7	39.9	220.11	-2.39	85.1	45 %	Moderate
2027	8	31.3	280.22	-2.91	56.06	61 %	High
2027	9	24.5	30.89	-3.0	45.51	9 %	Low
2027	10	16.6	24.46	-4.24	27.77	2 %	Low
2027	11	24.1	40.03	-1.03	24.9	16 %	Low
2027	12	24.3	1.02	-6.63	19.62	7 %	Low
2028	1	17.3	29.67	-3.82	13.43	0 %	Low
2028	2	22.0	20.26	-9.68	17.52	0 %	Low
2028	3	18.1	24.63	-6.81	49.48	4 %	Low
2028	4	22.6	27.83	-3.96	31.28	8 %	Low
2028	5	21.4	36.44	-2.45	14.87	0 %	Low
2028	6	39.6	265.59	-1.27	66.52	99 %	High
2028	7	39.2	260.08	-1.9	54.83	91 %	High
2028	8	37.6	241.56	-1.72	94.69	91 %	High
2028	9	15.6	30.63	-3.16	27.11	3 %	Low
2028	10	24.3	43.04	-7.7	22.67	4 %	Low
2028	11	20.9	26.87	-4.01	33.61	3 %	Low
2028	12	19.5	44.76	-7.58	35.14	0 %	Low
2029	1	20.5	14.2	-8.08	37.74	0 %	Low
2029	2	17.5	34.46	-1.65	34.76	9 %	Low
2029	3	18.3	22.62	-6.7	23.35	3 %	Low
2029	4	24.5	19.37	-3.96	25.59	8 %	Low
2029	5	19.0	11.8	-3.88	39.76	0 %	Low
2029	6	39.8	238.53	-3.09	58.44	41 %	Moderate
2029	7	37.4	244.74	-1.88	99.33	98 %	High
2029	8	37.7	223.58	-3.39	57.0	9 %	Low
2029	9	21.2	9.23	-9.77	44.6	4 %	Low
2029	10	21.5	24.39	-0.47	47.4	50 %	High
2029	11	16.2	3.32	-3.51	27.21	0 %	Low
2029	12	16.8	40.22	-5.51	39.96	0 %	Low
2030	1	20.4	4.38	-4.73	14.55	2 %	Low
2030	2	19.2	17.95	-2.47	13.58	0 %	Low
2030	3	16.5	17.08	-0.13	16.6	43 %	Moderate
2030	4	17.7	30.74	-9.98	46.79	0 %	Low
2030	5	23.1	38.02	-2.23	28.46	0 %	Low
2030	6	37.0	253.88	-2.3	76.42	78 %	High
2030	7	32.6	216.89	-3.26	86.51	8 %	Low
2030	8	33.7	221.06	-4.76	64.38	4 %	Low
2030	9	19.4	19.51	-0.45	25.5	42 %	Moderate
2030	10	19.3	38.21	-7.71	18.37	0 %	Low
2030	11	19.8	36.45	-5.49	47.83	0 %	Low
2030	12	23.1	9.12	-3.4	22.49	7 %	Low

Figure Error! No text of specified style in document.**.26** Month-wise Prediction of Flood in Balochistan for 2025-2030 using Linear Regression

The summer months of June, July, and August consistently show high rainfall and risk, with some months nearing a 100% risk threshold. Conversely, months with low rainfall and moderate temperatures, like January, February, and December, show lower flood risks, typically categorized as low or moderate. For example, January 2026 reports 39.69 mm of rain but only a 43% risk. By 2028-2029, the region experiences consistent high-risk months with rainfall over 200 mm. This, coupled with rising temperatures, suggests a greater intensity of flood events in the coming years. Additionally, decreasing ice levels may further contribute to increased water flow and potential flooding.

Changes in vegetation could affect soil permeability, influencing flood risk, although this relationship remains unclear. These trends highlight the growing need for enhanced flood management strategies, particularly during the summer months.

4.16. Flood Prediction in Balochistan Using SVM

Flood risk predictions in Balochistan from 2025-2030, based on the Support Vector Machine (SVM) model, reveal varying environmental conditions impacting flood risk, as shown in Figure Error! No text of specified style in document..**27**. The year 2025 starts with moderate flood risks, but a spike occurs in August when rainfall reaches 289.1 mm, and temperatures hit 37.8°C, pushing the risk to 50%. In 2026, similar patterns are observed, with peak risks in June-August due to extreme rainfall events. Rainfall exceeds 200 mm in several months, reflecting the region's vulnerability to heavy rainfall.

In 2027, a notable anomaly occurs in December, where minimal rainfall (1.02 mm) results in a high-risk rating of 60%. This suggests other factors, such as ice or vegetation conditions, may influence the risk. Despite general consistency in risk patterns, the monsoon season remains a critical period. By 2028, a slight reduction in flood risk is observed, particularly due to lower rainfall in August.

However, high temperatures and moderate rainfall still lead to moderate flood risks.

SVM Predictions (2025-2030) Month-wise

Year	Month	Temp	Rain(mm)	Ice	veg	SVM (%)	SVM Risk
2025	1	17.5	32.33	-6.7	35.67	15 %	Low
2025	2	18.0	2.56	-8.95	32.29	25 %	Moderate
2025	3	18.0	35.26	-4.47	10.72	40 %	Moderate
2025	4	23.7	15.04	-2.99	41.11	10 %	Low
2025	5	19.4	31.66	-7.94	22.99	5 %	Low
2025	6	37.3	218.63	-3.04	53.5	0 %	Low
2025	7	38.7	253.21	-1.85	97.75	10 %	Low
2025	8	37.8	289.1	-3.93	60.44	50 %	High
2025	9	17.2	32.88	-2.67	19.11	15 %	Low
2025	10	21.6	36.92	-8.33	30.16	20 %	Moderate
2025	11	22.8	11.97	-3.41	13.65	25 %	Moderate
2025	12	22.8	5.66	-3.25	44.48	20 %	Moderate
2026	1	17.5	39.69	-0.54	24.46	25 %	Moderate
2026	2	21.9	18.8	-8.78	16.7	10 %	Low
2026	3	16.6	24.97	-7.93	35.59	25 %	Moderate
2026	4	23.3	25.03	-7.4	12.4	50 %	High
2026	5	15.6	46.51	-8.14	29.23	5 %	Low
2026	6	35.5	239.07	-2.94	79.65	30 %	Moderate
2026	7	35.3	211.29	-3.59	88.98	40 %	Moderate
2026	8	39.8	262.01	-4.55	90.18	60 %	High
2026	9	24.1	5.35	-0.49	21.94	30 %	Moderate
2026	10	24.0	46.05	-0.5	21.33	20 %	Moderate
2026	11	17.7	24.27	-1.49	31.2	25 %	Moderate
2026	12	15.2	45.03	-7.53	14.72	25 %	Moderate
2027	1	17.3	42.63	-7.2	44.63	15 %	Low
2027	2	22.5	9.48	-9.5	38.21	25 %	Moderate
2027	3	20.8	45.15	-6.76	27.7	25 %	Moderate
2027	4	23.3	9.44	-6.15	40.83	30 %	Moderate
2027	5	16.2	33.42	-6.35	12.86	30 %	Moderate
2027	6	39.0	209.35	-4.41	87.24	15 %	Low
2027	7	39.9	220.11	-2.39	85.1	40 %	Moderate
2027	8	31.3	280.22	-2.91	56.06	0 %	Low
2027	9	24.5	30.89	-3.0	45.51	40 %	Moderate
2027	10	16.6	24.46	-4.24	27.77	20 %	Moderate
2027	11	24.1	40.03	-1.03	24.9	15 %	Low
2027	12	24.3	1.02	-6.63	19.62	60 %	High
2028	1	17.3	29.67	-3.82	13.43	10 %	Low
2028	2	22.0	20.26	-9.68	17.52	10 %	Low
2028	3	18.1	24.63	-6.81	49.48	15 %	Low
2028	4	22.6	27.83	-3.96	31.28	5 %	Low
2028	5	21.4	36.44	-2.45	14.87	30 %	Moderate
2028	6	39.6	265.59	-1.27	66.52	5 %	Low
2028	7	39.2	260.08	-1.9	54.83	15 %	Low
2028	8	37.6	241.56	-1.72	94.69	5 %	Low
2028	9	15.6	30.63	-3.16	27.11	0 %	Low
2028	10	24.3	43.04	-7.7	22.67	25 %	Moderate
2028	11	20.9	26.87	-4.01	33.61	25 %	Moderate
2028	12	19.5	44.76	-7.58	35.14	0 %	Low
2029	1	20.5	14.2	-8.08	37.74	10 %	Low
2029	2	17.5	34.46	-1.65	34.76	10 %	Low
2029	3	18.3	22.62	-6.7	23.35	25 %	Moderate
2029	4	24.5	19.37	-3.96	25.59	30 %	Moderate
2029	5	19.0	11.8	-3.88	39.76	30 %	Moderate
2029	6	39.8	238.53	-3.09	58.44	15 %	Low
2029	7	37.4	244.74	-1.88	99.33	0 %	Low
2029	8	37.7	223.58	-3.39	57.0	0 %	Low
2029	9	21.2	9.23	-9.77	44.6	40 %	Moderate
2029	10	21.5	24.39	-0.47	47.4	40 %	Moderate
2029	11	16.2	3.32	-3.51	27.21	0 %	Low
2029	12	16.8	40.22	-5.51	39.96	50 %	High
2030	1	20.4	4.38	-4.73	14.55	15 %	Low
2030	2	19.2	17.95	-2.47	13.58	20 %	Moderate
2030	3	16.5	17.08	-0.13	16.6	25 %	Moderate
2030	4	17.7	30.74	-9.98	46.79	40 %	Moderate
2030	5	23.1	38.02	-2.23	28.46	50 %	High
2030	6	37.0	253.88	-2.3	76.42	10 %	Low
2030	7	32.6	216.89	-3.26	86.51	15 %	Low
2030	8	33.7	221.06	-4.76	64.38	20 %	Moderate
2030	9	19.4	19.51	-0.45	25.5	25 %	Moderate
2030	10	19.3	38.21	-7.71	18.37	20 %	Moderate
2030	11	19.8	36.45	-5.49	47.83	30 %	Moderate
2030	12	23.1	9.12	-3.4	22.49	60 %	High

Figure Error! No text of specified style in document.**.27** Month-wise Prediction of Flood in Balochistan for 2025-2030 using SVM

The flood risk predictions for Balochistan from 2025 to 2030 show a clear seasonal pattern, with the monsoon months consistently exhibiting the highest flood risks. However, the variability in rainfall, alongside changing temperature and vegetation conditions, suggests that flood risks are increasingly influenced by a combination of factors beyond just precipitation. This emphasizes the need for continuous monitoring and adaptive management strategies to mitigate the effects of extreme weather events in the region. The predictions also indicate that while some years show a decrease in risk, the overall trend towards higher rainfall and fluctuating temperature patterns signals a growing concern for future flood risks in Balochistan.

4.17 Comparative Analysis

The comparative analysis of various studies on flood prediction and risk assessment reveals significant insights into the methodologies, input parameters, and accuracy levels achieved across different regions. Aziz *et al.* (2024) focused on predicting flood-induced displacement and identifying the most affected Tehsils in Sindh, Pakistan, using demographic and geographical variables from the 'Sindh Flood Data Analysis and Prediction 2022' dataset. They employed models such as Gradient Boosting Classifier, Logistic Regression, Decision Trees, and Lasso Regression, achieving an overall accuracy of 77.42%. While this study provided valuable insights into localized flood impacts, its accuracy was relatively lower compared to other studies, possibly due to the complexity of demographic and geographical factors.

In contrast, Rajab *et al.* (2023) utilized climatic data, including temperature, rainfall, wind speed, and pressure, to forecast flood risk in Bangladesh. Their models, Random Forest and Support Vector Machines (SVM), achieved a higher average accuracy of 90.5%. This suggests that climatic variables, when effectively modeled, can yield more precise flood risk predictions. Similarly, Khan *et al.* (2024)

incorporated a broader range of climatic and hydrological data, such as rainfall, temperature, river discharge, soil moisture, and humidity, to predict flood risk in the Indus Basin, Pakistan. Their use of Random Forest, SVM, and Decision Trees resulted in an average accuracy of 88%, further emphasizing the importance of integrating multiple environmental factors for robust flood prediction.

Antwi-Agyakwa *et al.* (2023) conducted a global review of flood risk prediction models, incorporating variables like rainfall, ice, soil moisture, and historical flood data. Their analysis of various models, including Random Forest, SVM, and regression models, achieved an average accuracy of 90%. This study highlights the effectiveness of machine learning models in diverse geographical contexts, demonstrating their adaptability to different environmental conditions. Farman *et al.* (2024) focused on predicting rainfall trends in Pakistan using historical rainfall data, temperature, humidity, and other climatic variables. Their use of ARIMA, LSTM, and Random Forest models achieved an accuracy of 88%, underscoring the potential of time-series and machine learning models in understanding climatic patterns.

The current study stands out for its high accuracy levels across four provinces of Pakistan: Sindh, Punjab, KPK, and Balochistan. By utilizing historical flood data on rainfall, vegetation (NDVI), temperature changes, and ICE (Individual Conditional Expectation), the study achieved remarkable accuracy rates: 96.98% for Sindh, 99.15% for Punjab, 99.92% for KPK, and 99% for Balochistan. The best-performing models varied by region, with Random Forest and SVM being effective in Punjab and Balochistan, Decision Trees and Random Forest in Sindh, and Decision Trees in KPK. This study demonstrates the importance of tailoring models to regional characteristics and leveraging a combination of environmental and climatic variables for highly accurate flood risk prediction.(Mosavi et al.)

Table **Error! No text of specified style in document.**.2 Comparative Analysis of this study with other similar studies

Study	Input Parameters	Output Parameters	Models Used	Overall Accuracy	Location
(Aziz *et al.*, 2024)	Demographic and geographical variables from the 'Sindh Flood Data Analysis and Prediction 2022' dataset	Prediction of flood-induced displacement and identification of the most affected Tehsils	Gradient Boosting Classifier, Logistic Regression, Decision Trees, Lasso Regression	77.42% (Average of all models used)	Sindh, Pakistan
(Rajab *et al.*, 2023)	Temperature, Rainfall, Wind Speed, and Pressure (Climatic Data from historic records)	Flood risk forecasting based on climatic variables	Random Forest, Support Vector Machines (SVM)	90.5% (Average of all models used)	Bangladesh
(Khan *et al.*, 2024)	Rainfall, Temperature, River Discharge, Soil Moisture, and Humidity (Climatic and Hydrological Data)	Flood risk prediction and identification of high-risk areas	Random Forest, Support Vector Machine (SVM), and Decision Trees	88% (Average across the models used)	Indus Basin, Pakistan
(Antwi-Agyakwa *et al.*, 2023)	Rainfall, Ice, Soil Moisture, and Historical Flood Data	Flood risk prediction and forecast of potential flood events	Various (Review of multiple models, including Machine Learning models like Random Forest, SVM, and Regression Models)	90% (Average of all models used	Global (Various locations)
(Farman *et al.*, 2024)	Historical rainfall data, temperature, humidity, and climatic variables	Predicted rainfall trends	ARIMA, LSTM, Random Forest	88%	Pakistan
This Study	Historical Flood Data on Rain (mm/month), Veg NDVI, Temperature changes, and ICE: Individual Conditional Expectation	Flood risk prediction	Best Performing Models: Punjab: Random Forest, SVM Sindh: Decision Tree, Random Forest, Linear Regression, SVM KPK: Decision Tree Baluchistan: Random Forest, SVM	Sindh: 96.98% Punjab: 99.15% KPK: 99.92% Baluchistan: 99%	Four provinces of Pakistan "Sindh" "Punjab" "KPK" "Baluchistan"

4.18. Mitigation Insights and Recommendations

Based on the analysis of historical flood data, several mitigation strategies were identified:

- **Flood Control Infrastructure:** Based on historical flood data experts suggest building new or repairing flood control infrastructure in flood-prone areas.
- **Land-Use Planning**: Past flood records allow cities to develop zoning rules that block new development in regions that often flood.

- **Community Awareness**: Flood history shows why neighborhood training and planning efforts must continue to lessen flood risks and help residents handle storms effectively.
- **Enhanced Drainage Systems**: Past flood records show us which drainage areas require upgrades to handle water properly and lower flood damage.

CHAPTER FIVE

Conclusion

Fatima Abbas
Riphah International University Faisalabad

*Corresponding Author: fatima.abbas@riphahfsd.edu.pk

The application of machine learning tools in natural disaster prediction and disaster risk management brings about a revolution in the approach taken by societies in the preparation, prediction, and mitigation of natural disasters. Natural disasters, especially floods, continue to be a serious challenge to humanity across the globe. These disasters lead to the death of people, destruction of essential infrastructural assets, economic volatility, and environmental deterioration. Developing countries become highly susceptible to such disasters because of various reasons like their location, high population rate, rapid urbanization, and weak structure in their ability to handle disasters. Pakistan becomes a vital example of such countries because of its vast river networks, monsoon weather conditions, accelerated melting glaciers, and socioeconomic structures which make the country prone to flood incidents. The country has depended on traditional hydrological models, trend analysis, and statistical methods in predicting and responding to flood occurrences.

Although such models have contributed to the initial comprehension of river behavior and environmental characteristics, they contain certain structural weaknesses that prevent them from becoming effective tools for our modern, rapidly evolving world. Such conventional forecasting models do not possess the ability to analyze vast data sources in real-time conditions, nor do they reflect the highly non-linear nature of relationships between changing environmental factors. In addition to this, due to the fact that conventional models are based on static premises regarding climatic change and land-use patterns, they demonstrate poor adaptation to the climate change and harsh weather conditions characteristic of the current global warming epoch. The result of these structural inadequacies is the delay in warnings and a reduction in the efficiency of warning systems.

The recent developments in technology in terms of artificial intelligence and data-driven analytical models have created revolutionary possibilities that allow for transcending these traditional constraints and creating a new level of disaster prediction and management frameworks. While traditional mathematical models fail at certain areas of modeling, machine learning excels at precisely these areas with the ability to analyze large amounts of complex data, find complex underlying patterns, and create non-linear models without the need for predefined assumptions about the physical world. The data-driven nature of machine learning makes it ideal for flood prediction, as this environment consists of multiple dynamic interactions between different climatic, hydrological, and geographical factors. By being trained on huge datasets, the machine learning system can understand complex patterns of future floods and accurately predict them. In order to explore the possibility of implementing machine learning into a practical framework for flood prediction on a regional scale, a detailed implementation of machine learning models was conducted on flood prediction in Pakistan. The extensive historical dataset with a duration of twenty-two years from 2000 to 2021 was used for this analysis.

In order to address the potential inadequacies arising out of nation-wide modeling, the analysis was carried out separately within the four unique regions of Pakistan, namely, Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan. Through such regionalized approach, it was possible to account for the unique factors prevailing in each region, which include individual river systems, differential distribution of monsoon rains, unique topography, as well as the history of flood incidents within these regions. In doing so, it is possible to arrive at forecast results that will be highly practical and useful for managing disasters. Based on this historic data, forecasts were made for predicting the likelihood of floods occurring during every month between 2025 and 2030.

The practical implementation of this framework included the training, testing, and comparison of four commonly used machine learning algorithms, namely, Decision Tree, Random Forest, Linear Regression, and Support Vector Machine. The unique predictive ability, mathematical background, and learning efficiency of each model makes them suitable for different applications in the environmental and classification sciences. In order to optimize these aspects of machine learning, thorough consideration had to be given to data preprocessing, imputation of missing values, transformation of categorical values, and feature selection. Environmental variables such as rain intensity, river discharges, seasonal climate variables, temperature variations, and NDVI vegetation values were singled out and used to train the models. Comparing these models revealed remarkable insights about their functioning, showing that machine learning approaches considerably outperform traditional hydrological models in terms of prediction accuracy and computational speed.

Among the analyzed models, the Random Forest model proved particularly effective due to the use of ensemble learning techniques that allow for combining different decision-making models in order to reduce variability and effectively avoid the problem of overfitting that is prevalent in all machine learning models. Whereas the Decision Tree model showed high efficacy in terms of mapping features of local interest, it was vulnerable to potential random errors and variations associated with the data stream. In comparison, the Linear Regression model offered a simple baseline that could help to identify linear trends but failed to take into account the complexity of the data associated with the highly volatile non-linear transformations of the environment caused by climate change and urbanization. The Support Vector Machine showed particular effectiveness in handling multi-index calculations and classifications, allowing us to remove subjectivity from regional vulnerability assessments.

In particular, the results of the province-specific model have shown the exceptional effectiveness of data-oriented AI systems in representing the environmental conditions of each region of Pakistan. The machine learning models trained for this purpose proved their ability to make highly accurate predictions regarding the studied parameters in the case of all four provinces. In particular, the prediction system showed an accuracy rate of 96.98 percent concerning the difficult environmental situation in the province of Sindh, 99.15 percent in Punjab with its extensive fields and river valleys, 99.92 percent in mountainous Khyber Pakhtunkhwa, and 99.00 percent in the dry region of Balochistan. Moreover, the high level of accuracy has been confirmed by the fact that the actual changes in vegetation were always very close to those predicted by the models. In particular, Punjab and Sindh proved highly similar predicted and actual environmental trends in the sense that they showed some discrepancy when there was an overprediction of vegetation density; however, it remained close to the actual trend. Even more interestingly, in the highly mountainous regions of Khyber Pakhtunkhwa, where there is a lot of vegetation, the prediction remained rather close to actual figures.

These results illustrate the extremely high robustness of the machine learning approach, its capability to operate across different landscapes, as well as ability to derive future-oriented environmental information out of past history. Apart from their impressive mathematical success, the results produced by this study emphasize the revolutionary potential of machine learning as a fundamental decision-making tool in the context of national disasters' mitigation efforts. Precise predictions of monthly flooding risks in the 2025-2030 timeframe allow creating actionable knowledge that can be used within different planning, policy making, and infrastructure improvement initiatives. Using the reliable long-range forecasts, it becomes possible for the disaster management teams to move from the reactive model of handling crises to a more proactive approach focused on preparations for the adverse environmental events.

From an operational point of view, effective flood risk maps contribute to the rational land use planning, prevent the construction of new settlements and buildings on flood-prone territories, help urban planners develop appropriate drainage systems within cities, and offer climate resilient agricultural schedules. Moreover, by offering objective estimates that cannot be obtained through historical interpretation, machine learning models can offer better cooperation between different sectors, from governmental agencies and NGOs to community-based preparedness measures. At the end of the day, the ultimate importance of such technology is assessed in both human and financial terms, since the lead time that can be provided thanks to these models helps save more lives and avoid displacement, ensure agricultural safety, save resources on restoration and protect vital infrastructure.

Although the benefits outlined by these studies are clearly evident, the ongoing application of machine learning towards disaster prediction will need to actively respond to a number of technological and structural obstacles over a sustained period of time. Machine learning algorithms are essentially data-driven systems, with the reliability of predictions being reliant upon the accuracy of the data input into the system. In less developed areas, for instance, the lack of monitoring systems may lead to incomplete data sets, adding an element of uncertainty to algorithmic predictions. Furthermore, environmental processes are constantly changing, and in the context of fast urbanization in underdeveloped areas, the expansion of impervious surfaces has been disrupting traditional drainage flows and shifting the nature of flood risk factors within relatively short periods of time. Global warming, on the other hand, continues to affect temperature levels, increasing glacial melt and monsoon rains beyond established norms.

However, if machine learning models are not constantly updated, their predictive accuracy will deteriorate in the long term as real-life environments will become more different from those represented by historical data collected during initial training procedures. In order to deal with such obstacles and keep the framework operating at an optimum level, its design needs to include the ability to acquire new data, train the existing machine learning models, and synchronize them with the current technological advancements in environmental monitoring. The path towards the future development of sophisticated disaster predictions involves the synergetic fusion of machine learning technologies and evolving technological ecosystems. In combination with Geographic Information Systems technology, machine learning techniques can produce hazard maps, which will be useful in guiding rescue operations. At the same time, IoT sensor networks established within major rivers and reservoirs as well as municipal drainage systems will provide continuous real-time hydrological data streams.

It is possible to achieve all this through cloud computing, since this will facilitate swift computation, analytics, and visualization of data. The above research makes a significant contribution to literature by showing how advanced artificial intelligence can be effectively applied in highly vulnerable areas through customization in an area that faces extreme disaster risk. The methodology, provincial customization, and results of this study act as a fundamental foundation that should be followed in order to achieve successful artificial intelligence forecasting in regions facing similar environmental and socioeconomic challenges. In summary, it can be concluded that the application of historical baseline, provincial, and machine learning algorithms can facilitate highly accurate and efficient flood disaster prediction. It is very crucial for such advanced tools and technologies to be embraced in order to create a solid base for sustainable and evidence-based policies, institution readiness, preservation of human life, reduction of economic losses, and sustainable development.

References

Abdalzaher, M.S., H.A. Elsayed, M.M. Foudaand M.M. Salim. 2023. Employing machine learning and iot for earthquake early warning system in smart cities. Energies 16:495-507.

Abdollahzadeh, B.and F.S. Gharehchopogh. 2022. A multi-objective optimization algorithm for feature selection problems. Engineering with Computers 38:1845-1863.

Abrahams, M., J. Price, F. Whitlockand G. Williams. 1976. The Brisbane floods, January 1974: their impact on health. Medical journal of Australia 2:936-939.

Abt, S., R. Wittier, A. Taylorand D. Love. 1989. Human stability in a high flood hazard zone 1. JAWRA Journal of the American Water Resources Association 25:881-890.

Ahmad, I., X. Wang, M. Waseem, M. Zaman, F. Aziz, R.Z.N. Khanand M. Ashraf. 2022. Flood Management, Characterization and Vulnerability Analysis Using an Integrated RS-GIS and 2D Hydrodynamic Modelling Approach: The Case of Deg Nullah, Pakistan. Remote Sensing 14:781-799.

Ahmadlou, M., A.k. Al-Fugara, A.R. Al-Shabeeb, A. Arora, R. Al-Adamat, Q.B. Pham, N. Al-Ansari, N.T.T. Linhand H. Sajedi. 2021. Flood susceptibility mapping and assessment using a novel deep learning model combining multilayer perceptron and autoencoder neural networks. Journal of Flood Risk Management 14:12683:12701.

Ahmed, M.K., M. Rahmanand J. Van Ginneken. 1999. Epidemiology of child deaths due to drowning in Matlab, Bangladesh. International journal of epidemiology 28:306-311.

Albahri, A., Y.L. Khaleel, M.A. Habeeb, R.D. Ismael, Q.A. Hameed, M. Deveci, R.Z. Homod, O. Albahri, A. Alamoodiand L. Alzubaidi. 2024. A systematic review of trustworthy artificial intelligence applications in natural disasters. Computers and Electrical Engineering 118:109409:109521.

Ali, S., M.J.M. Cheema, M.M. Waqas, M. Waseem, M.K. Leta, M.U. Qamar, U.K. Awan, M. Bilaland M.H.u. Rahman. 2021. Flood mitigation in the transboundary chenab river basin: a basin-wise approach from flood forecasting to management. Remote Sensing 13:3916:1937.

Althuwaynee, O.F., B. Pradhanand S. Lee. 2012. Application of an evidential belief function model in landslide susceptibility mapping. Computers & Geosciences 44:120-135.

Ansari, M.P.A.and M.S. Patil. 2022. Real Time Flood Management and Control with Early Warning System using Artificial Neural Network. International Journal of Creative Research Thoughts 10:912-917.

Antwi-Agyakwa, K.T., M.K. Afenyoand D.B. Angnuureng. 2023. Know to Predict, Forecast to Warn: A Review of Flood Risk Prediction Tools. Water 15:1-34.

Aziz, A., M. Ali, S. Muhammad, H. Zaidi, F. Nawazand F. Sheikh. 2024. Flood Risk Analysis in Sindh , Pakistan : Predicting the Most Affected Tehsil Using Statistical and Machine Learning Models with Comprehensive Data. 4:53-73.

Baydargil, H.B., S. Serdaroglu, J.-S. Park, K.-H. Parkand H.-S. Shin. 2018. Flood detection and control using deep convolutional encoder-decoder architecture. In: 2018 International conference on information and communication technology robotics (ICT-ROBOT). p 1-3.

Beltaos, S. 1995. River Ice Jams. 1st ed. Water Resources Publication, LLC, Colarado.

Bhatta, N., S.H. Tanimand P. Murray-Tuite. 2024. Dynamics of Link Importance through Normal Conditions, Flood Response, and Recovery. Sustainability 16:819-842.

Broto, V.C. 2017. The United Nations Conference on Housing and Sustainable Urban Development. Journal of Housing & Community Development 74:10-37.

Cazabat, C. 2022. Displacement, natural hazards, and health consequences, Oxford Research Encyclopedia of Natural Hazard Science No. 12. p. 436-455.

Chaudhary, M.T.and A. Piracha. 2021. Natural Disasters—Origins, Impacts, Management. Encyclopedia 1:1101-1131.

Chen, J., Q. Li, H. Wangand M. Deng. 2020. A machine learning ensemble approach based on random forest and radial basis function neural network for risk evaluation of regional flood disaster: a case study of the Yangtze River Delta, China. International journal of environmental research and public health 17:49:68.

Cirilo, J.A., L.F.d.M. Verçosa, M.M.d.A. Gomes, M.A.B. Feitoza, G.d.F. Ferrazand B.d.M. Silva. 2020. Development and application of a rainfall-runoff model for semi-arid regions. Revista Brasileira de Recursos Hidricos 25:1-19.

Cosgrove, W.J.and F.R. Rijsberman. 2000. World water vision: making water everybody's business. 1st ed. Routledge, London.

De Wrachien, D., S. Mambrettiand B. Schultz. 2011. Flood management and risk assessment in flood-prone areas: Measures and solutions. Irrigation and Drainage 60:229-240.

Drakonakis, G.I., G. Tsagkatakis, K. Fotiadouand P. Tsakalides. 2022. Ombrianet—supervised flood mapping via convolutional neural networks using multitemporal sentinel-1 and sentinel-2 data fusion. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 15:2341-2356.

Ekka, P., S. Patra, M. Upreti, G. Kumar, A. Kumarand P. Saikia. 2023. Land Degradation and its impacts on Biodiversity and Ecosystem services. Land and Environmental Management through Forestry 6:77-101.

El-Haddad, B.A., A.M. Youssef, H.R. Pourghasemi, B. Pradhan, A.-H. El-Shaterand M.H. El-Khashab. 2021. Flood susceptibility prediction using four machine learning techniques and comparison of their performance at Wadi Qena Basin, Egypt. Natural Hazards 105:83-114.

Fan, Y., H. Liand G. Miguez-Macho. 2013. Global patterns of groundwater table depth. Science 339:940-943.

Farman, H., N. Islam, S.A. Ali, D. Khan, H.A. Khan, M. Ahmedand A. Farman. 2024. Advancing Rainfall Prediction in Pakistan : A Fusion of Machine Learning and Time Series Forecasting Models. International Journal of Emerging Engineering and Technology 3:17-24.

Glago, F.J. 2021. Flood disaster hazards; causes, impacts and management: a state-of-the-art review. Natural hazards-impacts, adjustments and resilience 8:29-37.

Goel, R. 2020. Flood damage analysis using machine learning techniques. Procedia Computer Science 173:78-85.

Handelman, G.S., H.K. Kok, R.V. Chandra, A.H. Razavi, S. Huang, M. Brooks, M.J. Leeand H. Asadi. 2019. Peering into the black box of artificial intelligence: evaluation metrics of machine learning methods. American Journal of Roentgenology 212:38-43.

Hashi, A.O., A.A. Abdirahman, M.A. Elmi, S.Z.M. Hashiand O.E.R. Rodriguez. 2021. A real-time flood detection system based on machine learning algorithms with emphasis on deep learning. International Journal of Engineering Trends and Technology 69:249-256.

He, Y., D. Ma, J. Xiong, W. Cheng, H. Jia, N. Wang, L. Guo, Y. Duan, J. Liuand G. Yang. 2022. Flash flood vulnerability assessment of roads in China based on support vector machine. Geocarto International 37:6141-6164.

Hussain, M.A., Z. Shuai, M.A. Moawwez, T. Umar, M.R. Iqbal, M. Kamranand M. Muneer. 2023. A review of spatial variations of multiple natural hazards and risk management strategies in Pakistan. Water 15:407-430.

Ighile, E.H., H. Shirakawaand H. Tanikawa. 2022. Application of GIS and machine learning to predict flood areas in Nigeria. Sustainability 14:5039-5056.

Jato-Espino, D., N. Sillanpää, I. Andrés-Doménechand J. Rodriguez-Hernandez. 2018. Flood risk assessment in urban catchments using multiple regression analysis. Journal of Water Resources Planning and Management 144:04017085-04017098.

Jehanzaib, M., M. Ajmal, M. Achiteand T.W. Kim. 2022. Comprehensive Review: Advancements in Rainfall-Runoff Modelling for Flood Mitigation. Climate 10:1-17.

Jiang, R., Z. Cai, Z. Wang, C. Yang, Z. Fan, Q. Chen, K. Tsubouchi, X. Songand R. Shibasaki. 2021. DeepCrowd: A deep model for large-scale citywide crowd density and flow prediction. IEEE Transactions on Knowledge and Data Engineering 35:276-290.

Jin, W., J. Yangand Y. Fang. 2020. Application methodology of big data for emergency management. In: 2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS). p 326-330.

Jordan, M.I.and T.M. Mitchell. 2015. Machine learning: Trends, perspectives, and prospects. Science 349:255-260.

Jurafsky, D.and J. Martin. 2012. Logistic regression Logistic regression Logistic regression. Speech and Language Processing 404:731-735.

Keum, H.J., K.Y. Hanand H.I. Kim. 2020. Real-time flood disaster prediction system by applying machine learning technique. KSCE Journal of Civil Engineering 24:2835-2848.

Khalaf, M., A.J. Hussain, D. Al-Jumeily, T. Baker, R. Keight, P. Lisboa, P. Fergusand A.S. Al Kafri. 2018. A data science methodology based on machine learning algorithms for flood severity prediction. In: 2018 IEEE Congress on Evolutionary Computation (CEC). p 1-8.

Khan, M., U. Ali, N. Khan, S. Hussainand A. Ahmad. 2022. Hydraulic Model for Flood Forecasting of Tajabad Khwar in Hayatabad Phase III Peshawar, Pakistan: A Case Study. Journal of Advanced Research in Fluid Mechanics and Thermal Sciences 89:160-174.

Khan, M., A.U. Khan, B. Ullahand S. Khan. 2024. Developing a machine learning-based flood risk prediction model for the Indus Basin in Pakistan. Water Practice and Technology 19:2213-2225.

Krichen, M., M.S. Abdalzaher, M. Elwekeiland M.M. Fouda. 2024. Managing natural disasters: An analysis of technological advancements, opportunities, and challenges. Internet of Things and Cyber-Physical Systems 4:99-109.

Kumar, V., H.M. Azamathulla, K.V. Sharma, D.J. Mehtaand K.T. Maharaj. 2023. The State of the Art in Deep Learning Applications, Challenges, and Future Prospects: A Comprehensive Review of Flood Forecasting and Management. Sustainability (Switzerland) 15

Lantz, B. 2019. Machine learning with R: expert techniques for predictive modeling. Packt publishing ltd.

Lawal, Z.K., H. Yassinand R.Y. Zakari. 2021. Flood prediction using machine learning models: a case study of Kebbi state

Nigeria. In: 2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE). p 1-6.

Lee, J., D. Perera, T. Glickmanand L. Taing. 2020. Water-related disasters and their health impacts: A global review. Progress in Disaster Science 8:100123-100145.

Lei, X., W. Chen, M. Panahi, F. Falah, O. Rahmati, E. Uuemaa, Z. Kalantari, C.S.S. Ferreira, F. Rezaieand J.P. Tiefenbacher. 2021. Urban flood modeling using deep-learning approaches in Seoul, South Korea. Journal of Hydrology 601:126684-126702.

Letessier, C., J. Cardi, A. Dussel, I. Ebtehajand H. Bonakdari. 2023. Enhancing Flood Prediction Accuracy through Integration of Meteorological Parameters in River Flow Observations: A Case Study Ottawa River. Hydrology 10:835-864.

Li, X., D. Yan, K. Wang, B. Weng, T. Qinand S. Liu. 2019. Flood risk assessment of global watersheds based on multiple machine learning models. Water 11:1654-1677.

Lisboa, P.J., S. Saralajew, A. Vellido, R. Fernández-Domenechand T. Villmann. 2023. The coming of age of interpretable and explainable machine learning models. Neurocomputing 535:25-39.

Londhe, S.N.and V. Panchang. 2018. ANN techniques: A survey of coastal applications. Advances in Coastal Hydraulics 16:199-234.

Mahmood, S., A.-u. Rahmanand A. Sajjad. 2019. Assessment of 2010 flood disaster causes and damages in district Muzaffargarh, Central Indus Basin, Pakistan. Environmental Earth Sciences 78:1-11.

Makker, M., R. Ramanathanand S.B. Dinesh. 2019. Post disaster management using satellite imagery and social media data. In: 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS). p 1-6.

Mateo-Garcia, G., J. Veitch-Michaelis, L. Smith, S.V. Oprea, G. Schumann, Y. Gal, A.G. Baydinand D. Backes. 2021. Towards global flood mapping onboard low cost satellites with machine learning. Scientific reports 11:7249-7281.

McAlpine, E.D., P. Michelowand T. Celik. 2022. The utility of unsupervised machine learning in anatomic pathology. American Journal of Clinical Pathology 157:5-14.
Mishra, S., M.K. Jenaand A.K. Tripathy. 2022. Towards the development of disaster management tailored machine learning systems. In: 2022 IEEE India Council International Subsections Conference (INDISCON). p 1-6.
Moishin, M., R.C. Deo, R. Prasad, N. Rajand S. Abdulla. 2021. Designing deep-based learning flood forecast model with ConvLSTM hybrid algorithm. IEEE Access 9:50982-50993.
Moon, H., S. Yoonand Y. Moon. 2023. Urban flood forecasting using a hybrid modeling approach based on a deep learning technique. Journal of Hydroinformatics 25:593-610.
Mosavi, A., P. Ozturkand K.W. Chau. 2018. Flood prediction using machine learning models: Literature review. Water (Switzerland) 10:1-40.
Mosqueira-Rey, E., E. Hernández-Pereira, D. Alonso-Ríos, J. Bobes-Bascaránand Á. Fernández-Leal. 2023. Human-in-the-loop machine learning: a state of the art. Artificial Intelligence Review 56:3005-3054.
Motta, M., M. de Castro Netoand P. Sarmento. 2021. A mixed approach for urban flood prediction using Machine Learning and GIS. International Journal of Disaster Risk Reduction 56:7053-7082.
Moussaoui, H.and M. Benslimane. 2023. Reinforcement learning: A review. International Journal of Computing and Digital Systems 13:1-27.
Moustafa, S.S., M.S. Abdalzaherand H. Abdelhafiez. 2022. Seismo-lineaments in Egypt: analysis and implications for active tectonic structures and earthquake magnitudes. Remote Sensing 14:6151-6173.
Munawar, H.S., A.W. Hammadand S.T. Waller. 2021. A review on flood management technologies related to image processing and machine learning. Automation in Construction 132:103916-103939.
Murthy, P.and S. Bobba. 2021. AI-Powered Predictive Scaling in Cloud Computing: Enhancing Efficiency through Real-Time Workload Forecasting. IRE Journals 5:143-144.

Naidu, G., T. Zuvaand E.M. Sibanda. 2023. A review of evaluation metrics in machine learning algorithms. In: Computer Science On-line Conference. p 15-25.

Nevo, S., E. Morin, A. Gerzi Rosenthal, A. Metzger, C. Barshai, D. Weitzner, D. Voloshin, F. Kratzert, G. Elidan, G. Dror, G. Begelman, G. Nearing, G. Shalev, H. Noga, I. Shavitt, L. Yuklea, M. Royz, N. Giladi, N. Peled Levi, O. Reich, O. Gilon, R. Maor, S. Timnat, T. Shechter, V. Anisimov, Y. Gigi, Y. Levin, Z. Moshe, Z. Ben-Haim, A. Hassidimand Y. Matias. 2022. Flood forecasting with machine learning models in an operational framework. Hydrology and Earth System Sciences 26:4013-4032.

Ng, K.K., C.-H. Chen, C.K. Lee, J.R. Jiaoand Z.-X. Yang. 2021. A systematic literature review on intelligent automation: Aligning concepts from theory, practice, and future perspectives. Advanced Engineering Informatics 47:101246-101274.

Pham Quang, M.and K. Tallam. 2022. Predicting Flood Hazards in the Vietnam Central Region: An Artificial Neural Network Approach. Sustainability 14.

Puttinaovarat, S.and P. Horkaew. 2020. Flood forecasting system based on integrated big and crowdsource data by using machine learning techniques. IEEE Access 8:5885-5905.

Rajab, A., H. Farman, N. Islam, D. Syed, M.A. Elmagzoub, A. Shaikh, M. Akramand M. Alrizq. 2023. Flood Forecasting by Using Machine Learning: A Study Leveraging Historic Climatic Records of Bangladesh. Water (Switzerland) 15

Reed, C., W. Anderson, A. Kruczkiewicz, J. Nakamura, D. Gallo, R. Seagerand S.S. McDermid. 2022. The impact of flooding on food security across Africa. Proceedings of the National Academy of Sciences 119:2119399119-2119399131.

Riza, H., E.W. Santoso, I.G. Tejakusumaand F. Prawiradisastra. 2020. Advancing flood disaster mitigation in Indonesia using machine learning methods. In: 2020 International Conference on ICT for Smart Society (ICISS). p 1-4.

Rosenzweig, B.R., L. McPhillips, H. Chang, C. Cheng, C. Welty, M. Matsler, D. Iwaniecand C.I. Davidson. 2018. Pluvial flood risk and opportunities for resilience. Wiley Interdisciplinary Reviews: Water 5:1302-1324.

Sarker, I.H. 2021. Machine learning: Algorithms, real-world applications and research directions. SN computer science 2:160-179.

Serre, D.and C. Heinzlef. 2018. Assessing and mapping urban resilience to floods with respect to cascading effects through critical infrastructure networks. International Journal of Disaster Risk Reduction 30:235-243.

Shah, S.A., D.Z. Seker, S. Hameedand D. Draheim. 2019. The rising role of big data analytics and IoT in disaster management: recent advances, taxonomy and prospects. IEEE Access 7:54595-54614.

Sun, X., K. Jin, H. Tao, Z. Duanand C. Gao. 2023. Flood Risk Assessment Based on Hydrodynamic Model—A Case of the China–Pakistan Economic Corridor. Water (Switzerland) 15:167-193.

Syifa, M., S.J. Park, A.R. Achmad, C.-W. Leeand J. Eom. 2019. Flood mapping using remote sensing imagery and artificial intelligence techniques: a case study in Brumadinho, Brazil. Journal of coastal research 90:197-204.

Tan, X.Z., Y. Li, X.X. Wu, C. Dai, X.L. Zhangand Y.P. Cai. 2024. Identification of the key driving factors of flash flood based on different feature selection techniques coupled with random forest method. Journal of Hydrology: Regional Studies 51:457-489.

Tanim, A.H., C.B. McRae, H. Tavakol-Davaniand E. Goharian. 2022. Flood detection in urban areas using satellite imagery and machine learning. Water 14:1140-1164.

Tehrany, M.S., M.-J. Lee, B. Pradhan, M.N. Jeburand S. Lee. 2014. Flood susceptibility mapping using integrated bivariate and multivariate statistical models. Environmental earth sciences 72:4001-4015.

Wajid, M., M.K. Abid, A. Asif Raza, M. Haroonand A.Q. Mudasar. 2024. Flood Prediction System Using IOT & Artificial Neural Network. VFAST Transactions on Software Engineering 12:210-224.

Watik, N.and L.M. Jaelani. 2019. Flood evacuation routes mapping based on derived-flood impact analysis from landsat 8 imagery using network analyst method. The International

Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 42:455-460.
Yang, Q., Y. Zhang, W. Daiand S.J. Pan. 2020. Transfer learning. 1st ed. Cambridge University Press.
Yang, X., Z. Song, I. Kingand Z. Xu. 2022. A survey on deep semi-supervised learning. IEEE Transactions on Knowledge and Data Engineering 35:8934-8954.
Zhong, S., K. Zhang, M. Bagheri, J.G. Burken, A. Gu, B. Li, X. Ma, B.L. Marrone, Z.J. Renand J. Schrier. 2021. Machine learning: new ideas and tools in environmental science and engineering. Environmental science & technology 55:12741-12754.
Zhu, H., J. Leandroand Q. Lin. 2021. Optimization of artificial neural network (Ann) for maximum flood inundation forecasts. Water (Switzerland) 13:1-15.
Zhu, X., Y. Zhang, W. Qi, Y. Liang, X. Zhao, W. Caiand Y. Li. 2022. Flood forecasting methods for a semi-arid and semi-humid area in Northern China. Journal of Flood Risk Management 15:1-16.