Ministère de l'enseignement Supérieur et de la recherche Scientifique
وزارة التعليم العالي والبحث العلمي

Badji Mokhtar Annaba University
Université Badji Mokhtar – Annaba
Faculté de Technologie

Département Hydraulique

جامعة باجي مختار– عنابــة
كـليــة التكـنـولـوجيـا
قســم الـــري

Thèse

Présentée pour obtenir le diplôme de

Doctorat Troisième Cycle

Filière : Hydraulique

Spécialité : Hydraulique urbaine

Par :
BOUGHANDJIOUA Samira

Thème :

Etude d'impact du changement climatique sur la sensibilité hydraulique des systèmes de drainage urbain

Thèse soutenue le 22 Juin 2025 devant le jury composé de :

N°	Nom et prénom	Grade	Etablissement	Qualité
01	HAMMAR Yahia	Prof.	Université Badji Mokhtar - Annaba	Président
02	LAOUACHERIA Fares	Prof.	Université Badji Mokhtar - Annaba	Rapporteur
03	KECHIDA Said	MCA	Université Badji Mokhtar - Annaba	Examinateur
04	BENOUARA Naouel	MCA	Université Badji Mokhtar - Annaba	Examinatrice
05	HEDDAM Salim	Prof.	Université du 20 Août 1955 Skikda	Examinateur
06	HEBAL Aziz	MCA	Université du 20 Août 1955 Skikda	Examinateur
07	AZIZI Nabiha	Prof.	Université Badji Mokhtar - Annaba	Membre Invité

" دراسة تأثير التغير المناخي على الحساسية الهيدروليكية لأنظمة الصرف الحضري"

الملخص:

تشكل الفيضانات تهديدًا كبيرًا للمناطق الحضرية في الجزائر، حيث يعد التوسع الحضري السريع وتغير المناخ من العوامل التي تزيد من تواتر وشدة هطول الأمطار، مما يجعل أنظمة الصرف الصحي غير قادرة على امتصاص مياه الجريان السطحي، مما يشكل تحديات كبيرة لإدارة مياه الأمطار. تدرس هذه الأطروحة تأثير هذه العوامل على نظام الصرف الصحي للمنطقة الحضرية المسماة بئر فارينة ببلدية عزابة ولاية سكيكدة شرق الجزائر، مع التركيز على حدود النموذج الهيدروليكي التقليدي MIKE+ للتنبؤ الدقيق بفيضان شبكة مياه الأمطار. تصف هذه الأطروحة نهجًا مبتكرًا يسمى SWN-ML (شبكة مياه الأمطار - التعلم الآلي)، والذي يجمع بين عمليات المحاكاة الهيدروليكية من نموذج MIKE+ وخوارزميات التعلم الآلي المتقدمة، بما في ذلك تقنيات التعلم الجماعي، مثل تعزيز الانحدار والغابات العشوائية. تم تطوير قاعدة بيانات شاملة تشمل الخصائص الجغرافية والمناخية والهندسية لشبكة مياه الأمطار في منطقة الدراسة. تمت معايرة نموذج MIKE+ بناءً على القياس الوحيد لمستويات المياه عند مخرج شبكة مياه الأمطار أثناء هطول الأمطار في 4 يوم فبراير 2019، وتم استخدام نهج SWN-ML للتنبؤ بمتوسط معدلات الفائض لفترات المطر المختلفة وفترات العودة . تشير النتائج إلى أداء نماذج التعلم الجماعي (ELM) مقارنة بنماذج التعلم الآلي التقليدية في التنبؤ بمعدلات الفائض. ولوحظ وجود علاقة قوية بين تنبؤات طرق المجموعة وعمليات محاكاة MIKE+، مما يؤكد فعاليتها في التقاط ديناميكيات شبكة مياه الأمطار. بالإضافة إلى ذلك، حدد تحليل أهمية الميزة المستند إلى متوسط مربع الخطأ (MSE) المتغيرات الرئيسية التي تؤثر على معدلات الفائض، مما يوفر رؤى قيمة لتحسين استراتيجيات إدارة مياه الأمطار. يسلط هذا البحث الضوء على إمكانية دمج النماذج المادية مع تقنيات التعلم الآلي لتحسين التنبؤ وإدارة فيضانات شبكة مياه الأمطار في المناطق الحضرية. يوفر نهج SWN-ML إطارًا قويًا وموثوقًا لتقييم مخاطر الفيضانات، مما يتيح تطوير أنظمة إنذار مبكر فعالة واستراتيجيات التخفيف من آثار الفيضانات، وهو أمر بالغ الأهمية لبناء بيئات حضرية مرنة.

كلمات مفتاحية: تغير المناخ، أنظمة الصرف الحضري، الفيضانات، التعلم الآلي، التعلم الجماعي.

« Etude d'impact du changement climatique sur la sensibilité hydraulique des systèmes de drainage urbain »

Résumé :

Les inondations constituent une menace majeure pour les zones urbaines en Algérie, car l'urbanisation rapide et le changement climatique sont des facteurs qui augmentent la fréquence et l'intensité des pluies, rendant les systèmes de drainage incapable d'absorber les eaux de ruissellement, posant des défis importants pour la gestion des eaux pluviales. Cette thèse étudie l'impact de ces facteurs sur le système de drainage de la zone urbaine appelée Bir Farina, commune d'Azzaba, wilaya de Skikda à l'Est de l'Algérie, en se concentrant sur les limites du modèle hydraulique traditionnel MIKE+ pour la prévision précise des débordements du réseau d'eau pluviale. Cette thèse décrit une approche innovante appelée SWN-ML (Storm Water Network - Machine Learning), qui combine les simulations hydrauliques du modèle MIKE+ avec des algorithmes avancés d'apprentissage automatique, notamment des techniques d'apprentissage d'ensemble, telles que Gradient Boosting et Random Forests. Une base de données complète englobant les caractéristiques géographiques, climatiques et géométriques du réseau d'eaux pluviales de la région d'étude a été développée. Le modèle MIKE+ a été calibré en se basant sur l'unique mesure des hauteurs d'eau à l'exutoire du réseau d'eau pluviale lors de l'événement pluvieux du 4 février 2019, et l'approche SWN-ML a été utilisée pour prédire les taux de débordement moyens pour différentes durées de pluie et périodes de retour. Les résultats indiquent la performance des modèles d'apprentissage par ensemble (ELM) par rapport aux modèles d'apprentissage automatique classiques dans la prédiction des taux de débordement. Une forte corrélation a été observée entre les prédictions des méthodes d'ensemble et les simulations du MIKE+, confirmant leur efficacité à capturer la dynamique du réseau d'eaux pluviales. De plus, une analyse de l'importance des caractéristiques basée sur l'erreur quadratique moyenne (MSE) a permet d'identifier les principales variables influençant le taux de débordement, offrant ainsi des informations précieuses pour améliorer les stratégies de gestion des eaux pluviales. Cette recherche met en lumière le potentiel d'intégrer des modèles basés sur des principes physiques avec des techniques d'apprentissage automatique pour améliorer la prédiction et la gestion des débordements des réseaux d'eaux pluviales dans les zones urbaines. L'approche SWN-ML offre un cadre robuste et fiable pour l'évaluation des risques d'inondation, permettant le développement de systèmes d'alerte précoce efficaces et de stratégies d'atténuation des inondations, cruciales pour construire des environnements urbains résilients.

Mots clés : Changement climatique, Systèmes de drainage urbains, Inondation, Apprentissage automatique, Apprentissage par ensemble

«Study of the impact of climate change on the hydraulic sensitivity of urban drainage systems»

Abstract:

Floods are a major threat to urban areas in Algeria, as rapid urbanization and climate change are factors that increase the frequency and intensity of rainfall, making drainage systems unable to absorb runoff, posing significant challenges for stormwater management. This thesis studies the impact of these factors on the drainage system of the urban area called Bir Farina in Azzaba city, Skikda province in eastern Algeria, focusing on the limitations of the traditional MIKE+ hydraulic model for the accurate prediction of stormwater network overflows. This thesis describes an innovative approach called SWN-ML (Storm Water Network - Machine Learning), which combines the hydraulic simulations of the MIKE+ model with advanced machine learning algorithms, including ensemble learning techniques, such as Gradient Boosting and Random Forests. A comprehensive database encompassing the geographical, climatic and geometric characteristics of the stormwater network in the study region was developed. The MIKE+ model was calibrated based on the unique measurement of water levels at the stormwater network outlet during the rainfall event of February 4, 2019, and the SWN-ML approach was used to predict the average overflow rates for different rainfall durations and return periods. The results indicate the performance of ensemble learning models (ELM) compared to classical machine learning models in predicting overflow rates. A strong correlation was observed between the predictions of the ensemble methods and the MIKE+ simulations, confirming their effectiveness in capturing the dynamics of the stormwater network. In addition, a feature importance analysis based on Mean Square Error (MSE) identified the main variables influencing overflow rates, providing valuable insights to improve stormwater management strategies. This research highlights the potential of integrating physically based models with machine learning techniques to improve the prediction and management of stormwater network overflows in urban areas. The SWN-ML approach provides a robust and reliable framework for flood risk assessment, enabling the development of effective early warning systems and flood mitigation strategies, crucial for building resilient urban environments.

Key words: Climate change, Urban drainage systems, Flood, Machine Learning, Ensemble Learning

Dedication

I dedicate this work to those who have been a continuous source of inspiration and encouragement along my journey:

To my dear parents, whose love, prayers, and support have been my greatest strength. Your steadfast belief in me has been the foundation of all my accomplishments.

To my family, who stood by me with love and patience, enduring my long hours of study and research with unwavering understanding.

To my advisors and professors, whose guidance and expertise have enriched my academic journey, helping me grow both personally and professionally.

And to my friends and colleagues, for their constant encouragement, insightful discussions, and shared moments of joy and challenges.

This work is a tribute to all of you, who have been my strength and inspiration.

Finally, I dedicate this work to Allah. All that I have, all that I am, and all that I do is because of and for You.

Acknowledgement

First and foremost, I express my deepest gratitude to Allah, the Almighty, for granting me the will, health, and patience throughout these long years.

I wish to sincerely thank all those who have contributed to the completion of this research and supported me throughout my academic journey.

Firstly, I extend my heartfelt appreciation to my supervisor, Dr. Fares Laouacheria, for his invaluable guidance, continuous support, and patience throughout this process. His expertise, insightful feedback, and encouragement have been pivotal in shaping my work, and I am truly grateful for the opportunity to learn under his mentorship.

A special thanks to Professor Azizi Nabiha for his assistance and valuable support throughout this journey. His effective guidance and insightful comments have significantly influenced the development of my ideas and research. His ongoing support and belief in my abilities have had a profoundly positive impact on enhancing my confidence, and I am truly grateful for all the advice and inspiration he has provided.
I would also like to extend my gratitude to Professor Meksaouine, Director of the laboratory of soils and hydraulic, for his support and guidance. His assistance and the resources of the laboratory were instrumental in facilitating my research.

I am also deeply thankful to Dr. Kechida Said for his encouragement and guidance during this journey. His support and advice have been invaluable, and I am sincerely grateful for his contributions.

Additionally, I am grateful to DHI for providing free software powered by MIKE for hydrological and hydraulic analysis. Their commitment to supporting research in this field has been invaluable to my work, allowing me to conduct thorough analyses and simulations.

I am deeply thankful to the faculty and staff of the Faculty of Technology at the University of Badji Mokhtar Annaba. Their knowledge, resources, and the supportive learning environment they provided have been essential in advancing my academic progress.

I would like to express my profound gratitude to the jury members for agreeing to evaluate this modest work. Your time and dedication are greatly appreciated.

To my beloved parents, who have been my unwavering pillars of strength: your prayers, sacrifices, and encouragement have been my greatest motivation. Without your love and support, this journey would not have been possible. To my family and friends, thank you for your patience, understanding, and for believing in me even when I doubted myself.

I am also grateful to my colleagues and friends who shared this journey with me, offering companionship, ideas, and moral support. Your encouragement has lightened this challenging journey and made it all the more fulfilling.

Finally, I acknowledge and thank everyone who, in one way or another, contributed to the completion of this work. Your support and encouragement will always be remembered with gratitude.

Chapter 1: General Introduction

Chapter 2: Impacts of Climate Change and Flooding on Urban Drainage Systems

Chapter 3 : Overview of Urban Drainage System Modelling: Hydrological, Hydraulic, and Machine Learning Approaches

Chapter 4 : Study Area and Methodology

Chapter 5: Results and Discussion

List of Figures

List of Tables

List of Abbreviations

A - Cross-sectional area of flow (m^2)
Af - Equatorial Climate
Am - Monsoon Climate (Köppen clima
ANN - Artificial Neural Network
Aw - Tropical Savanna Climate (Köppen climate classification)
AMC - Antecedent Moisture Conditions (conditions before rainfall or runoff event)
BSh - Warm Semi-Arid Climate
BSk - Cold Semi-Arid Climate
BWk - Cold Desert Climate
Ca - Humid Subtropical Climate
Cfa - Warm Oceanic Climate
Csa - Warm Mediterranean Climate
Csb - Temperate Mediterranean Climate
Cwa - Humid Subtropical Climate
Cwb - Subtropical Oceanic Highland Climate
Cfb - Temperate Oceanic Climate
CNN - Convolutional Neural Network
CN - Curve Number
DDPG - Deep Deterministic Policy Gradient
DHI - Danish Hydraulic Institute
DRL - Deep Reinforcement Learning
ELM - Ensemble Learning Model
EM-DAT - Emergency Events Database
EML - Ensemble Machine Learning
EPA - Environmental Protection Agency
FAO - Food and Agriculture Organization
GB - Gradient Boosting
GIS - Geographic Information System
GPD - Generalized Pareto Distribution
GUI - Graphical User Interface
HBV - Hydrologiska Byråns Vattenbalansavdelning model
HEC-HMS - Hydrologic Engineering Center's Hydrologic Modeling System
HEC-RAS - Hydrologic Engineering Center's River Analysis System
HML - The Hybrid Machine Learning
Ia - Initial Abstraction
IDF - Intensity-Duration-Frequency
ICM - Integrated Catchment Modeling
IPCC - Intergovernmental Panel on Climate Change
KNN - K-Nearest Neighbors
LID - Low Impact Development
LSTM - Long Short-Term Memory Neural Network
MCC - Matthews Correlation Coefficient
MAE - Mean Absolute Error
ML - Machine Learning
MODFLOW - Modular Finite-Difference Flow Model
MIKE+ - Hydrological and hydraulic modeling software by DHI
MSE - Mean Squared Error
N - Number of Observations

NSCE - Nash-Sutcliffe Efficiency
P - Average Rainfall (mm)
PCA - Principal Component Analysis
PCSWMM - Personal Computer Storm Water Management Model
Q - Actual Runoff (mm)
QDF - Flood-Duration-Frequency
R - Rainfall (mm)
R^2 - Coefficient of Determination
r - Pearson Correlation Coefficient
RF - Random Forest
RL - Reinforcement Learning
RNN - Recurrent Neural Network
RMSE - Root Mean Squared Error
S - Potential Infiltration
SCS-CN - Soil Conservation Service Curve Number
S_f - Slope of the Energy Grade Line
S_0 - Slope of the Channel Bed
SWAT - Soil and Water Assessment Tool
SWMM - Storm Water Management Model
SWMM5 - Version 5 of the Storm Water Management Model
SWN-ML - Stormwater Network Machine Learning
SVM - Support Vector Machine
T - Temperature (°C)
UDSs - Urban Drainage Systems
UH - Unit Hydrograph
USD - United States Dollar
USEPA - United States Environmental Protection Agency
X - Feature Set or Predictors
X_i - Individual Data Point
$\bar{X}$ - Mean of X_i
Y - Response Variable or Output
Y_i - Individual Data Point
$\bar{Y}$ - Mean of Y_i
Yi - Observed Value
$\bar{Y}$ - Corresponding Predicted Value
g - Acceleration due to gravity (m/s^2)
θ - Angle or parameter
x - Longitudinal distance along the channel (m)
t - Time (s)
t-SNE - t-Distributed Stochastic Neighbor Embedding

Chapter 1 : General Introduction

1.1 Background

1.1.1 Importance of Studying the Impact of Climate Change on Urban Drainage Systems

The study of climate change's impact on urban drainage systems (UDSs) in Algeria is crucial due to several interrelated factors, including rapid urbanization, increased flooding risks, and the unique climatic conditions of the region. Algeria has experienced significant urban growth, with its population increasing from about 11.3 million in 1960 to nearly 43 million in 2019. This rapid urbanization has led to a concentration of people in coastal and valley areas, where 90% of the population resides on just 13% of the country's land. Such urban expansion has resulted in increased impervious surfaces, which alter natural drainage patterns and exacerbate flood risks. The encroachment into flood-prone areas, particularly riverbeds, has further heightened vulnerability to flooding events, which are often triggered by heavy rainfall.

Algeria is highly sensitive to climate change, with projections indicating a 5-13% reduction in rainfall and a temperature increase of 0.6-1.1ºC by 2020(Schilling et al., 2020). These changes are expected to intensify the frequency and severity of flooding, particularly in urban areas where drainage systems may be inadequate to handle increased runoff. Historical data shows that floods have caused significant casualties and damages, with major events recorded in Algiers and other regions, leading to substantial human and material losses (Hassan et al., 2022).

1.1.2 Overview of Climate Change Impacts on Hydraulic Systems

Climate change has a significant impact on hydraulic systems, affecting their design, operation, and environmental interactions. One major consequence is the increased risk of flooding due to more frequent and severe extreme weather events, such as heavy rainfall. Many hydraulic structures, such as bridges and culverts, are designed using historical climate data, which may no longer be accurate under changing conditions. For instance, studies like one in Rwanda have shown that rising precipitation levels could exceed the design capacity of existing infrastructure, potentially leading to failures during extreme flood events(Iradukunda et al., 2024) . Furthermore, research in Costa Rica has demonstrated that climate change alters hydraulic parameters, such as river flow velocity and depth, leading to reduced water supply and increased flooding risks(Watson-Hernández et al., 2022). Additionally, temperature extremes pose challenges for the operation of hydraulic equipment, as fluctuations can affect system reliability and efficiency, such as changes in oil viscosity leading to mechanical failures. These impacts highlight the need to reassess hydraulic system designs to accommodate the evolving climate (Power & Motion. 2024).

1.1.3 Introduction to Classical Hydrological and Machine Learning Models

In the study of hydrological processes, classical hydrological models and machine learning models represent two distinct yet complementary approaches. Classical hydrological models are grounded in the physical principles of the water cycle, relying on mathematical equations to simulate interactions between precipitation, evapotranspiration, runoff, and other hydrological variables. These models, including well-established tools like HEC-HMS, SWAT, and MODFLOW, provide a detailed understanding based on physical characteristics of the landscape, but require extensive data for calibration and validation (Xu & Liang, 2021).

In contrast, machine learning models are data-driven and utilize algorithms such as Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and Random Forests to capture complex, non-linear relationships within observed data (Bhasme et al., 2022). These models excel in handling large datasets and are especially useful in applications where the physical processes are difficult to model, such as streamflow prediction, flood forecasting, and water quality management(Torres et al., 2024).

Recently, researchers have explored integrating the strengths of both classical and machine learning models to improve hydrological predictions (Bhasme et al., 2022). This hybrid approach leverages the physical understanding of classical models alongside the predictive power of machine learning, creating more robust and efficient models to address challenges like climate change and urbanization impacts on water systems(Xu & Liang, 2021). This thesis will delve into the potential of combining these methods to advance hydrological modeling and enhance decision-making for water resource management(Lange & Sippel, 2020).

1.2 Problem Statement

1.2.1 Challenges Faced Due to Increased Rainfall Intensity and Frequency

The increasing intensity and frequency of rainfall, driven by climate change, present several significant challenges, particularly for Algeria. Urban areas are experiencing heightened flood risks, as intensified short-duration rainfall leads to more frequent and severe flash floods, especially in cities with non-permeable surfaces and inadequate drainage systems. Rapid urbanization near riverbeds further exacerbates this vulnerability, causing substantial infrastructure damage and economic disruptions (World Bank, 2023). Despite the increased rainfall, Algeria grapples with critical water scarcity, with variability in rainfall patterns and rising temperatures threatening to worsen drought conditions. This complicates water resource management and agricultural productivity (World Health Organization, 2015; Dept. 2024). Additionally, climate change is projected to intensify existing arid conditions, with rising temperatures and increased extreme weather events potentially leading to more severe flooding and reduced water availability due to heightened evaporation and runoff (Dept., 2024; Hamitouche et al., 2024). The economic and social consequences are profound, with agriculture vital to Algeria's economy being particularly vulnerable to extreme weather, leading to crop failures, food insecurity, and infrastructure damage that strains public resources and hinders development (World Bank, 2023) Addressing these challenges necessitates comprehensive disaster risk management, updated infrastructure standards, and localized climate data to improve resilience against future climatic shocks.

1.2.2 Limitations of Classical Models in Predicting Stormwater Overflows

Classical models for predicting stormwater overflows, such as the Storm Water Management Model (SWMM) and Kinematic Wave Model, face several limitations, particularly in the context of climate change and urbanization. These models often rely on simplifying assumptions about rainfall distribution, soil characteristics, and catchment responses, which can lead to discrepancies between predictions and actual events (Lundy et al., 2012). Additionally, the accuracy of these models depends on the quality and granularity of input data, with uncertainties in rainfall measurements being a significant source of error (Lundy et al., 2012). Classical models also struggle to adapt to extreme weather events, like increased rainfall intensity, making it difficult to predict combined sewer overflows (Chiaburu & Biessmann, 2024; Marie et al., 2022). Moreover, they generally lack integration with modern techniques, such as machine learning and hybrid statistical methods, which have shown potential for improving predictive accuracy (T. Liu et al., 2021; Szeląg et al., 2018). These models frequently overlook environmental and systemic factors, such as the impact of urban green infrastructure, resulting in an underestimation of its effectiveness in stormwater managemen (Szeląg et al., 2018) . Therefore, there is an increasing need for more advanced models that can integrate diverse data sources, leverage real-time data, and more accurately capture the complexities of urban ecosystems.

1.2.3 The Need for Advanced Prediction Models for Overflow Management

The increasing frequency and severity of Combined Sewer Overflows, fueled by climate change and rapid urbanization, highlights the urgent need for advanced predictive models. Traditional methods for managing these overflows often struggle to accurately forecast events, especially during extreme weather conditions, leading to unregulated spills and environmental harm. Advanced models, leveraging machine learning and hybrid approaches, provide a more effective solution by enhancing real-time predictive capabilities. These models help utilities anticipate overflow risks, optimize drainage operations, and ensure timely interventions, ultimately protecting public health and minimizing environmental damage (Rosin et al. 2021; Yin et al. 2022; Liu et al. 2022a; Boughandjioua et al. 2024; Ma et al. 2024).

1.3 Objectives

This thesis explores the improvement of urban drainage systems in the context of climate change and stormwater management. The objective of the research is to assess how climate change affects urban drainage by examining the impact of shifting climate patterns on the performance and efficiency of stormwater infrastructure. To achieve this, the study will analyze the sensitivity of hydraulic systems using MIKE+ software, which allows for in-depth examination of various rainfall durations and return periods. Additionally, the thesis introduces the SWN-ML approach, which integrates hydraulic simulations from MIKE+ with machine learning models to predict average overflow rates across different rainfall scenarios. The research will focus on advanced machine learning algorithms, particularly ensemble learning methods such as Gradient Boosting and Random Forest, to develop robust models that improve overflow rate predictions in stormwater networks. By combining hydraulic simulations with machine learning techniques, this study aims to provide valuable insights into enhancing the performance of urban drainage systems in the face of evolving climatic conditions.

1.4 Research Questions

The research questions for this thesis are designed to address critical aspects of urban drainage systems and their performance under varying conditions. Firstly, the study aims to explore how rainfall duration affects overflow rates in urban drainage systems, examining the relationship between different rainfall durations and the resulting overflow. Secondly, the impact of various return periods on the sensitivity of hydraulic systems will be investigated to understand how different frequencies of rainfall events influence system performance. Thirdly, the research will compare the predictions of ensemble learning models with those of classical models to assess their relative effectiveness in forecasting overflow rates in stormwater networks. Finally, the thesis will investigate the effectiveness and accuracy of different ensemble learning techniques, such as Gradient Boosting and Random Forest, in predicting overflow rates, aiming to identify the most robust methods for improving stormwater management.

1.5 Thesis Structure

This section provides a concise overview of each chapter in the thesis, summarizing the key content and purpose of each. Chapter 2 explores the impact of climate change on urban drainage systems, with a focus on Africa and Algeria. It analyzes shifting rainfall patterns, the increasing frequency of extreme weather events, and the pressures of urbanization, all of which contribute to heightened flood risks. By using Algeria as a case study, the chapter highlights the vulnerability of existing drainage infrastructures and the need for adaptive stormwater management strategies to ensure resilience. It also reviews previous studies to set the stage for the innovative methods that follow.

Chapter 3 then examines traditional hydrological and hydraulic models alongside emerging machine learning (ML) approaches in urban drainage system analysis. It contrasts the limitations of conventional physical-based models with the advantages of ML techniques, particularly ensemble learning methods such as bagging and boosting, which offer improved flood prediction accuracy through enhanced data analysis. The chapter explores the application of supervised and unsupervised learning, as well as the potential of deep learning and reinforcement learning for stormwater management.

In Chapter 4, the study focuses on the case of Bir Farina in Algeria, detailing its geographical, climatic, and urban characteristics. It introduces the SWN-ML (Stormwater Network-Machine Learning) methodology, which integrates MIKE+ hydrological and hydraulic models with ensemble learning techniques like Random Forest and Gradient Boosting. This integrated approach is used to predict stormwater overflow rates and demonstrates the success of combining traditional and ML-based models to improve flood risk management and system resilience.

Chapter 5 presents the results and discussion of applying the SWN-ML methodology, demonstrating that integrating MIKE+ models with ensemble learning techniques significantly enhances flood prediction accuracy. The study evaluates model performance using metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). A feature importance analysis, conducted using the MSE method with 10 permutations, identifies critical factors influencing stormwater dynamics, including rainfall duration, intensity, and drainage network capacity. These findings provide valuable insights for improved planning and flood mitigation strategies.

Finally, the research concludes by summarizing the key contributions of integrating traditional hydrological models with machine learning techniques to improve stormwater overflow predictions. The study highlights the potential of ensemble learning models, such as Random Forest and Gradient Boosting, for effective flood risk management, especially in vulnerable regions like Algeria. The limitations of the study are acknowledged, and suggestions for future research are provided, including the exploration of advanced techniques like deep learning and reinforcement learning to enhance predictive models. Furthermore, the applicability of the SWN-ML approach to other urban areas with diverse climatic conditions is proposed as a promising avenue for improving urban drainage systems on a global scale.

Chapter 2 : Impacts of Climate Change and Flooding on Urban Drainage Systems

2.1 Introduction

Urban drainage systems are critical infrastructures for managing stormwater and minimizing flood risks in cities. However, climate change is increasing their vulnerability by altering rainfall patterns, intensifying storms, and raising the frequency of extreme weather events. With rising global temperatures, urban areas are now dealing with both extended droughts and sudden heavy rains, straining older drainage networks that were not designed for these conditions. Adapting these systems to changing climates is now crucial.

Africa, as a continent, presents a unique case in understanding climate change impacts on urban drainage due to its diverse climatic zones and rapid urbanization. From the equatorial rainforests to the arid deserts, climate variability across the continent leads to distinct challenges for stormwater management. The region's increasing population and expanding urban areas exacerbate the pressure on existing drainage infrastructures, leading to more frequent and severe flood events.

In Algeria, climate change significantly impacts urban drainage systems. The country experiences diverse climatic conditions, ranging from Mediterranean climates along the northern coast to arid and semi-arid conditions in the interior and southern regions. This variability poses significant challenges in managing urban stormwater, as the northern cities are prone to heavy rainfall and flash floods, while the southern regions face prolonged droughts followed by sudden, intense rainfalls. Rapid urbanization, coupled with inadequate drainage infrastructure, has led to increased flood risks in cities like Algiers and Oran, where stormwater networks are often overwhelmed during extreme weather events. Addressing these vulnerabilities requires a comprehensive understanding of climate projections and adaptive strategies to ensure the resilience of urban drainage systems.

This review will focus on the effects of climate change on urban drainage systems, particularly in the African and Algerian contexts, examining historical trends, climate projections, and the implications for urban stormwater management. By understanding these factors, we can better prepare urban areas for the challenges posed by climate change and ensure the resilience of vital infrastructure.

2.2 Preliminary Review of Climate Change Impacts on Urban Drainage Systems

2.2.1 Overview of Climate in Africa

Historically, Africa has experienced diverse climatic conditions, ranging from humid equatorial climates to arid desert environments. Since the early 20th century, the continent has warmed significantly, with average temperatures increasing by about 1°C. Projections indicate that temperatures could rise by an additional 2°C to as much as 6°C by the end of the 21st century, depending on greenhouse gas emissions scenarios(Hulme et al., 2001; Leal Filho et al. 2024) Rainfall patterns in Africa are highly variable, with substantial inter-annual and multi-decadal fluctuations. These variations can lead to extreme weather events, including droughts and floods, which have become more frequent and severe in recent decades. The Sahel region, in particular, has been affected by erratic rainfall, impacting agriculture and water resources (Hulme et al., 2001; Leal Filho et al. 2024; Giordano et al, 2019).The continent spans an area of 11,677,240 square miles (30,244,050 square kilometers) and is home to over 500 million people, representing roughly 10% of the global population (Chang et al., 2023).

Climatic classification methods are mainly categorized into empirical and genetic approaches. Empirical methods rely on observed environmental data, such as temperature, humidity, and precipitation (Fredholm, 1985), or derived measures like evaporation. On the other hand, genetic methods classify climate based on its underlying causal factors, including circulation systems, fronts, jet streams, solar radiation, and topography, which influence the spatial and temporal patterns of climatic data. While empirical methods provide a descriptive view of climate, genetic methods offer explanations. Nevertheless, empirical classifications are commonly used in practical applications.

The Köppen-Geiger climate classification system is the most commonly employed approach for categorizing climates (Yin et al., 2022). This system, developed by Wladimir Köppen in the late 19th century, divides the world's climates into five main groups (A, B, C, D, and E) based on average temperature and precipitation. The system has been further refined over the years, with the latest version, Köppen-Geiger-Peel, using a combination of temperature, precipitation, and vegetation data to classify climate (Irvin & Nakashima, 2023).

In Africa, only three of the main Köppen climate types (A, B, and C) are present, as shown in Figure 2.1. Among these, the dominant climate type by land area is the arid B, covering 57.2% of Africa, followed by the tropical A at 31.0%, and the temperate C at 11.8%(Nanda et al.2018). The figure illustrating the Köppen climate classification of Africa shows the different climate zones, each represented by a distinct color.

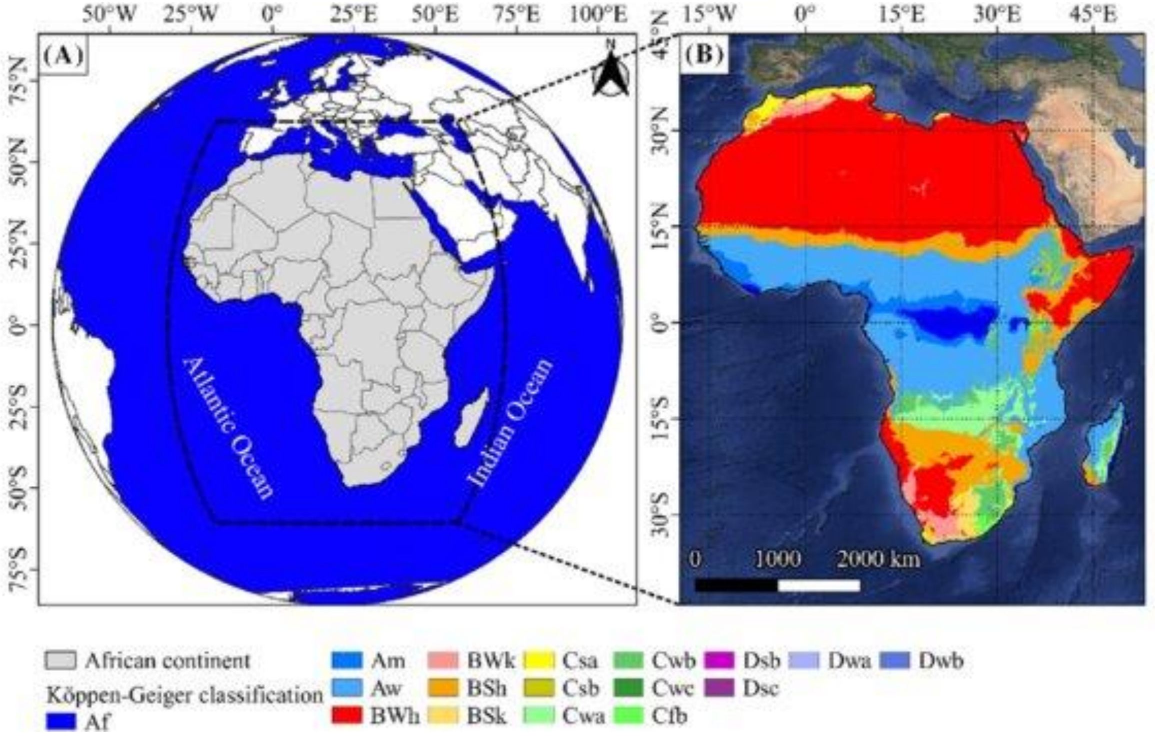


Figure 2.1: Africa map of koppen climate classification (Beck et al., 2018)

This study introduces global maps of the Köppen-Geiger climate classification with a remarkable resolution of 1 km, offering insights into both current conditions (1980–2016) and future projections (2071–2100) influenced by climate change (Beck et al., 2018; de Oliveira-Júnior et al., 2024).

.

Africa is characterized by various climate zones, including:

- **Equatorial climate (Af):** This climate is characterized by high temperatures and heavy rainfall throughout the year. It is found in the equatorial regions of Africa, such as the Congo Basin.
- **Monsoon climate (Am):** This climate is similar to the equatorial climate, but it has a distinct dry season. It is found in parts of West Africa and East Africa.
- Tropical savanna climate (Aw): This climate is characterized by warm temperatures and a distinct dry season. It is found in the savanna regions of Africa, such as the Serengeti National Park in Tanzania.

- **Warm desert climate (BWh):** This climate is characterized by very hot temperatures and very low rainfall. It is found in the deserts of Africa, such as the Sahara Desert.
- **Cold desert climate (BWk):** This climate is similar to the warm desert climate, but it has colder temperatures. It is found in the high-altitude deserts of Africa, such as the Namib Desert.
- **Warm semi-arid climate (BSh):** This climate is characterized by warm temperatures and low rainfall. It is found in the semi-arid regions of Africa, such as the Sahel.
- **Cold semi-arid climate (BSk):** This climate is similar to the warm semi-arid climate, but it has colder temperatures. It is found in the high-altitude semi-arid regions of Africa.
- **Warm mediterranean climate (Csa):** This climate is characterized by warm, dry summers and mild, wet winters. It is found in the Mediterranean region of Africa, such as Morocco and Algeria.
- **Temperate mediterranean climate (Csb):** This climate is similar to the warm mediterranean climate, but it has cooler summers. It is found in the higher-altitude Mediterranean regions of Africa.
- **Humid subtropical climate (Cwa):** This climate is characterized by warm, humid summers and mild, wet winters. It is found in the subtropical regions of Africa, such as South Africa.
- **Humid subtropical climate/Subtropical oceanic highland climate (Cwb):** This climate is similar to the humid subtropical climate, but it has cooler summers due to higher altitude. It is found in the mountainous regions of Africa, such as the Drakensberg Mountains in South Africa.
- **Warm oceanic climate (Cfa):** This climate is characterized by warm, humid summers and mild, wet winters. It is found in the coastal regions of Africa, such as the Atlantic coast of Morocco.
- **Humid subtropical climate (Ca):** This climate is similar to the warm oceanic climate, but it has a more pronounced dry season. It is found in the coastal regions of Africa, such as the Mediterranean coast of Egypt.
- **Temperate oceanic climate (Cfb):** This climate is characterized by mild temperatures and abundant rainfall throughout the year. It is found in the coastal regions of Africa, such as the Atlantic coast of South Africa.

The Köppen climate classification is a useful tool for understanding the climate of Africa. It helps us to understand the different climate zones in Africa and the factors that influence these climatesThe Intergovernmental Panel on Climate Change (IPCC) reports that the frequency and intensity of both droughts and floods have risen in recent years and are expected to continue increasing in many areas around the globe. This trend has significant consequences for the demand for and availability of high-quality freshwater resources (Ogallo, 2009).

In conclusion, classifying climate in the field of hydrological and urban drainage systems is crucial for understanding and adapting to the impacts of climate change (Kang et al., 2016). By incorporating climate change projections into urban drainage system planning and design, and implementing adaptation strategies, cities can become more resilient to the challenges posed by a changing climate(Acharki et al., 2017; Cea & Costabile, 2022; Ducharne, 2007; Duncker, 2019; Kourtis & Tsihrintzis, 2021.; Kushawaha et al., 2021; Tramblay et al., 2021) **.**

2.2.2 Overview of Climate Classification in Algeria

Algeria's climate is characterized by its significant diversity due to the country's vast size and varied geography. In the northern coastal regions, the climate is Mediterranean, featuring mild, wet winters with temperatures ranging from 10-15°C, and hot, dry summers with temperatures between 25-30°C.

The analysis of climate data from 1931 to 1990 in northern Algeria reveals a rise in temperature of 0.5°C, which is expected to increase by 1°C by 2020 and 2°C by 2050 (Climates to Travel, 2024; Sahnoune et al. 2013; Jianping et al. 2014). These regions receive moderate rainfall, particularly from November to January, with the wettest period lasting from September to May (Houta & Gadiyatov, 2023) . However, precipitation patterns are highly variable, and models predict that rainfall duration will become less frequent but more intense, with droughts becoming more common and prolonged(Sahnoune et al., 2013).

Moving inland to the central regions, the climate becomes more extreme, with hot, dry summers where temperatures can reach 35-40°C, and mild, dry winters with temperatures dropping to 5-10°C. Rainfall in these areas is sparse, often less than 400mm annually. In the southern Saharan regions, the climate shifts to a hot desert type, with scorching summer temperatures soaring up to 50°C and cool, dry winters with nighttime temperatures that can fall to 0°C(Miloudi et al. 2021; Climates to Travel, 2024). Rainfall is exceedingly rare, usually under 100mm per year. The Atlas Mountains in the northern part of the country present a cooler, more temperate climate with higher rainfall and even snowfall during winter months, with temperatures sometimes dipping below freezing at higher elevations.

The Algerian coast is shaped by a range of meteorological and hydrological factors, including winds, storms, and sea currents. Predominantly, winds blow from the west or northwest, sometimes shifting to the north-northeast, which can lead to hazardous storms impacting navigation. Frequent storms generate destructive waves from the northern sector (Houta & Gadiyatov, 2023). Algeria's climate overall is strongly influenced by its position between the Mediterranean Sea and the Sahara Desert, resulting in a climatic transition from Mediterranean conditions in the north to Saharan climates in the far south.(Houta & Gadiyatov, 2023).

Climate change poses significant challenges for Algeria, exacerbating existing climatic conditions. The country has already experienced a rise in average temperatures, which are projected to increase by 1.1°C by 2039 and by 2.8°C by the end of the century under high-emission scenarios. Rainfall patterns are also shifting, with a general decrease in total annual precipitation and an increase in the intensity of rainfall events. These changes heighten the risks of drought, desertification, and water scarcity, critically affecting agriculture and water resources. Overall, Algeria's climate is complex and varies significantly across regions, influenced by geographical features and climate change, resulting in distinct weather patterns and environmental challenges.

Figure 2.2 illustrates the climatic diversity in Algeria, ranging from the arid deserts in the south to the more temperate Mediterranean climate along the northern coast (Ciello, 2020). It clearly shows that the warm desert climate is the most extensive, covering most of the country's land area. This climate type predominates in the Sahara Desert region of Algeria and is characterized by extremely high temperatures and very low precipitation. Summers are intensely hot, while winters are mild to warm, with minimal rainfall throughout the year. This climate is typical of the southern and central parts of the country (Peel et al., 2007).

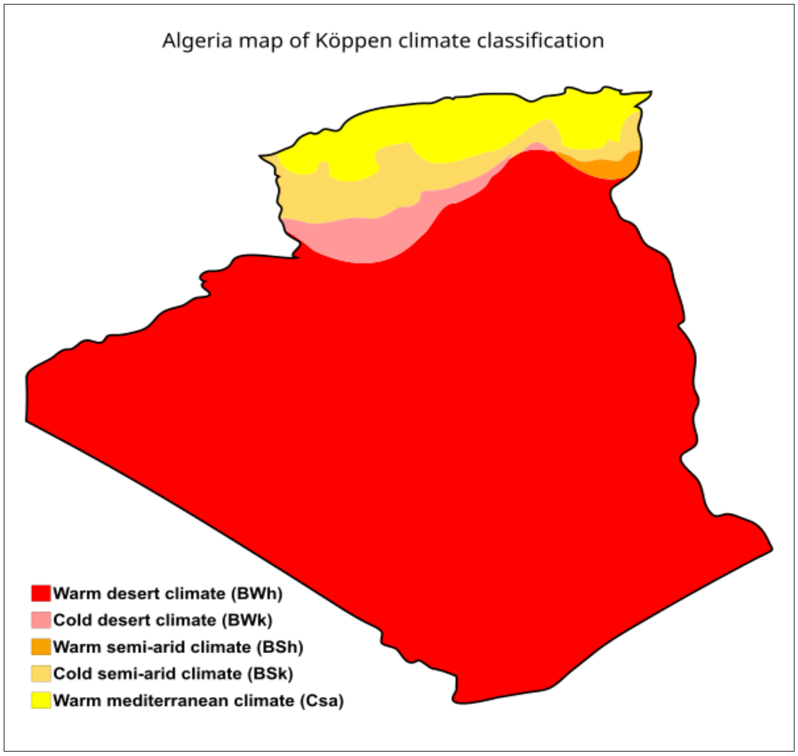


Figure 2.2: Köppen Climate Zones of Algeria: A Visual Classification Map (Ciello, 2020)

Algeria's diverse climates foster a variety of ecosystems, ranging from the arid Sahara to the verdant Mediterranean coast. However, like many regions worldwide, Algeria is experiencing the effects of climate change, which include rising temperatures, shifting precipitation patterns, and an increase in extreme weather events. These developments present considerable challenges for water resources, agriculture, and human communities, highlighting the need for adaptive strategies to lessen their impacts (IPCC, 2014).

2.2.3 Impact of Climate Change on Urban Drainage Systems

Stormwater drainage networks in urban areas are vital for managing water flow and preventing flooding, with their performance heavily influenced by regional precipitation patterns. Climate change poses a significant challenge by altering hydrological components, including increased rainfall intensity and frequency. According to the IPCC (2018), rising global temperatures will lead to more frequent and intense rainfall events, which may overwhelm existing drainage systems designed based on outdated rainfall patterns, threatening their effectiveness.

In Algeria, climate change has led to shifts in rainfall patterns, contributing to more intense and unpredictable storm events. While the frequency of moderate rainfall events has decreased, extreme rainfall events have surged by approximately 17% for 100-year storms (Hassan et al., 2022). This shift increases the vulnerability of urban drainage systems, which are not always designed to handle such extremes. Tranet al. (2011) proposed a framework for urban drainage systems in Australia that accounted for hydraulic considerations, urbanization effects, and changes in precipitation, highlighting the heightened risk of flooding due to climate change. Similarly, Djordjević et al. (2011) identified urban watershed flooding causes, emphasizing the necessity for innovative methodologies and financial resilience for flood management. The Danish Integrated Assessment System introduced by Kaspersen and Halsnæs (2017)) evaluated flood risks from excessive rainfall in Odense City, Denmark. Ahmed et al. (2018) examined urbanization and climate change effects in Dhaka City, emphasizing the need for adaptation-based spatial planning. Furthermore, Gu et al. (2019) assessed the impact of urbanization on nonstationary precipitation extremes in China, reinforcing the need to integrate climate change considerations into stormwater management. These studies collectively underscore the urgency of adapting urban drainage systems to better predict runoff, manage floods, and mitigate damage in light of changing climatic conditions.

2.3 Overview of Floods Risk at global and local scale

2.3.1 Global Overview of Floods Risk

Floods have been a significant global hazard, causing substantial economic damage and loss of life over the past few decades. Between 1992 and 2012, flood-related damages amounted to nearly 0.6 trillion USD, representing 28% of the total damages from all disasters during that period (Güneralp et al., 2015). From 1980 to 2013, global flood losses exceeded 1 trillion USD, and approximately 220,000 people lost their lives due to flood-related incidents (Dottori et al., 2016) . According to the international disaster database EM-DAT, between 1999 and 2009, there were 2,470 recorded flood events worldwide. These floods resulted in 147,457 fatalities and caused an estimated 372.5 billion USD in damages (EM-DAT, 2015). The occurrence and impact of floods vary significantly from country to country, influenced by various geological, spatial, developmental, and administrative factors, as illustrated in Figure 2.3. This variability highlights the intricate nature of flood risk and emphasizes the importance of customized flood management approaches for different regions (Berhail & Katipoğlu, 2024).

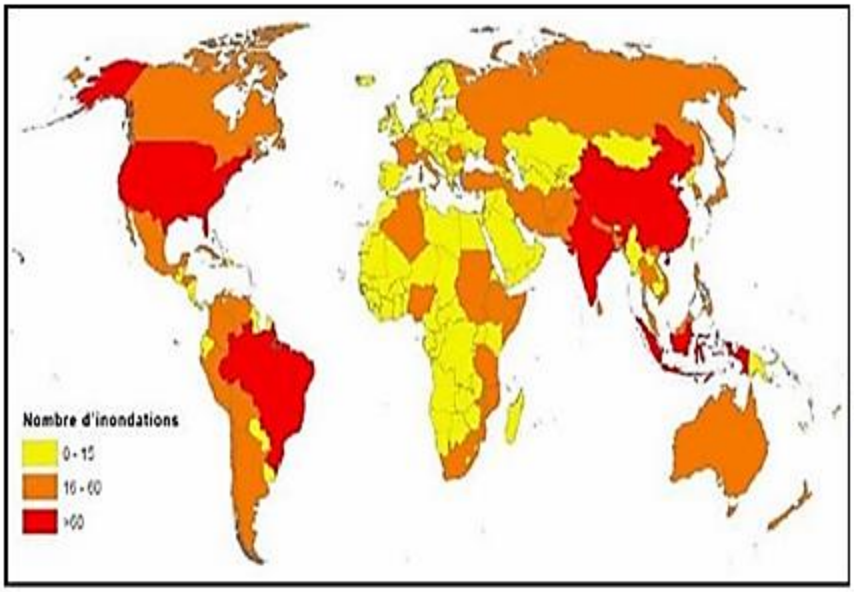


Figure 2.3: Flood Frequency and Impact by Country 1974 – 2003 (EM-DAT, 2015).

More than 2 billion people globally face the threat of flooding, with nearly 600 million of them living in poverty. Floods are the most frequent natural hazard, representing 44% of all hazard events between 2000 and 2019 (Eslamian et al., 2024). These events affected 1.65 billion people, caused approximately 100,000 deaths, and led to economic losses of up to 651 billion U.S. dollars. The frequency of global floods increased by 23% in 2020 compared to the previous 20 years, with 201 floods resulting in about 6,200 deaths. Regions like Western Europe, East Africa, South Asia, and China were particularly hard-hit, with the increase in flood frequency and severity attributed to climate change.

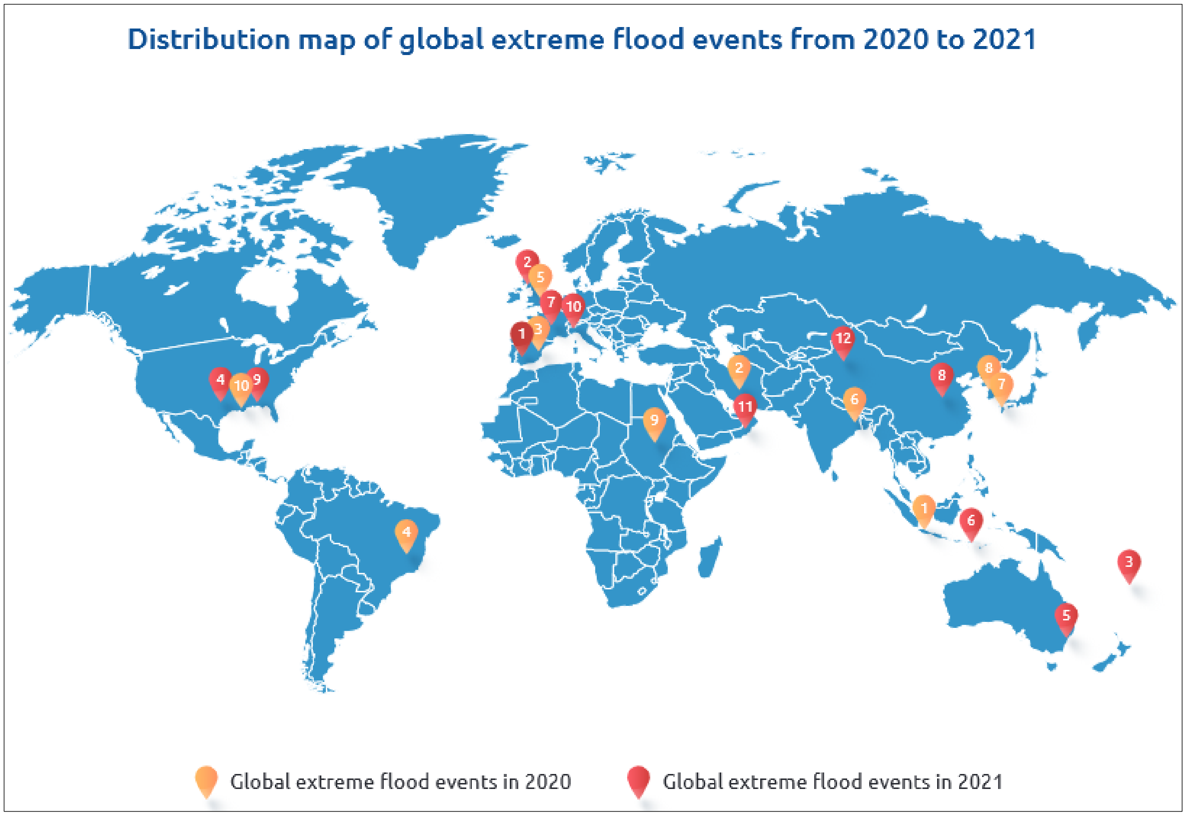


Figure 2.4: Distribution map of global extreme flood events from 2020 to 2021(Editorial Group, 2022)

Here are some of the countries that experienced significant flood events globally in 2020 and 2021, as illustrated in Figure 2.4, which represents countries in orange color for 2020 and red color for 2021 (Editorial Group, 2022).

- Indonesia: In January 2020, flooding caused by record-breaking rainfall in Jakarta left at least 66 people dead and displaced tens of thousands.
- Iran: In January 2020, severe floods in the Sistan-Baluchistan region led to significant damage, affecting over 200,000 people.
- Spain: Storm Gloria in January 2020 caused widespread damage, leaving at least 220,000 homes without power and resulting in four deaths.
- Brazil: In January 2020, torrential rains in southeastern Brazil led to catastrophic flooding, causing 59 deaths.
- United Kingdom: Storm Dennis in February 2020 triggered significant flooding across the UK, with tens of thousands of homes losing power.
- Bangladesh and India: In May 2020, Cyclone Amphan caused severe flooding, leading to 102 deaths and 4.9 million displacements.
- Japan: In July 2020, floods and landslides in Kyushu caused 77 deaths and destroyed over 15,000 buildings.
- South Korea: In August 2020, heavy rain caused flooding that resulted in 26 deaths and displaced over 1,000 people.
- Sudan: In September 2020, flooding in Khartoum caused nearly 100 deaths and damaged more than 100,000 homes.

- United States: Hurricane Delta in October 2020 led to significant flooding along the Gulf Coast, with insured losses reaching billions of dollars.
- China: In July 2021, extreme rains in Zhengzhou, Henan Province, caused severe flooding, resulting in substantial damage.
- Germany: In July 2021, a historic flood in western Germany caused massive damage, with 181 deaths and economic losses of 10 billion euros.
- Italy: In October 2021, a tropical cyclone in northwestern Italy led to severe flooding, setting a new European record for rainfall.

These flood events underscore the widespread impact of flooding across various regions globally, often exacerbated by extreme weather conditions

2.3.2 Local Overview of Floods Risk

Floods in Algeria have become a significant risk due to their profound impact on both the population and the environment. This growing risk has garnered considerable attention from both local and national authorities, leading to various studies aimed at mitigating the effects of flooding (Hafnaoui & Dabanlı, 2023).

a) **Urbanization and Flood Risk:** Menad et al. (2012) analyzed the impact of urban land use on flooding, with a particular focus on Bab El Oued. The study highlighted the exacerbation of flood risks due to urbanization in specific areas. Similarly, (Ketrouci et al., 2012) studied extreme floods in the Tafna catchment area, emphasizing the role of human activities, such as uncontrolled urbanization, in amplifying flood risks.

b) **Flood Frequency Analysis:** Hachemi and Benkhaled (2016) conducted a flood frequency analysis using data from the Foum El Gherza station, developing a flood-duration-frequency (QDF) curve and identifying the Pearson type III (P3) distribution as the most suitable for the data. Benameur et al. (2017) estimated flood events in the Abiod wadi using the generalized Pareto distribution (GPD), which better described the region's flood data.

c) **Flood Risk and Hazard Mapping:** (Yahiaoui, 2012) utilized HEC-RAS software to simulate flood-prone areas and produce flood hazard maps, essential tools for urban planning and risk mitigation. (SAMI et al., 2021) employed QGIS and HEC-RAS to create a flood hazard map for Chemora city, demonstrating the utility of GIS tools in flood risk assessment.

d) **Impact of Climate Change:** Benzater et al. (2021) analyzed extreme rain trends in the Macta watershed, noting that increasing rain intensity, particularly in August and September, has contributed to more frequent and severe flooding events. ZEGAIT and PIZZO (2023) proposed a flood control reservoir for the Idles basin, highlighting the need for infrastructure that can mitigate the effects of climate variability and extreme weather events.

e) **Historical Flood Data and Statistical Analysis:** An inventory of floods in Algeria from 1970 to 2000 was compiled, providing a foundational database for further studies. (Sardou et al., 2016) documented 127 flood events in northwestern Algeria from 1847 to 2014, offering insights into the historical patterns of flood occurrence.

f) **Recent Flood Trends (2010):** The period from 2010 to 2024 saw an increase in flood events, particularly in provinces such as Djelfa, M'sila, and Tamanrasset. Notably, on February 29th, 2024 (International Federation of Red Cross and Red Crescent Societies, 2024), the Wilaya of Jijel experienced significant flooding October 05-06, 2011 (Oued Ferrane and Hai El Fidayine El Bayadh) 10 deaths and dozens of missing people during 3 days of flooding. Although this period was marked by a high frequency of floods, it saw fewer fatalities compared to previous decades. (Abdelkebir, B. 2022)

Flood studies in Algeria reveal that flood risks are influenced by urbanization, climate change, and poor infrastructure. These studies improve understanding of flood dynamics and provide a basis for developing better mitigation strategies.

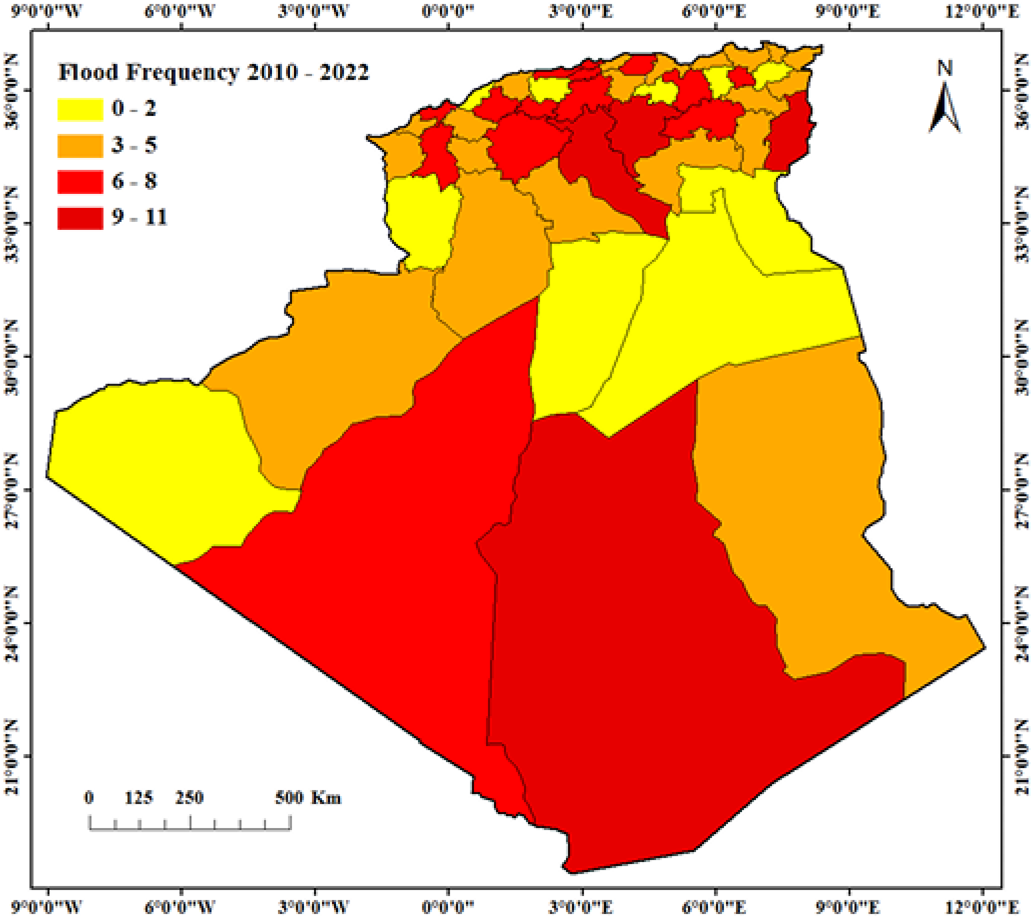


Figure 2.5: Flood Frequency Analysis in Algeria: Trends from 2010 to 2022 (Hafnaoui & Dabanlı, 2023)

The flood frequency analysis revealed that M'sila, Djelfa, and Tebessa experienced the highest number of floods, each with 11 events, while Tamanrasset followed with 9 floods. Batna, Tiaret, Alger, Medea, Oran, and Tissemsilt each recorded 8 floods, and Constantine and Adrar had 7 floods. The highest annual flood frequency was observed in Chlef, with four events occurring in 2021.
From 2010 to 2022, the analysis of flood frequency and death tolls indicated that M'sila, Djelfa, and Tamanrasset is particularly vulnerable to flooding. This frequent occurrence in these areas is likely due to inadequate vegetation cover, which exacerbates flood intensity. Based on the observations in

Figure 2.5, Djelfa, M’sila, Tamanrasset, Tiaret, Batna, Oran, and Medea are identified as the provinces most at risk for flooding (Hafnaoui & Dabanlı, 2023).

2.4 Conclusion

The increasing frequency and severity of extreme weather events driven by climate change have highlighted the critical need for global adaptation of urban drainage systems (Dharmarathne et al., 2024). In Africa, and especially in Algeria, this challenge is further intensified by varied climatic conditions and rapid urban growth. Cities along the northern coast, such as Algiers, are frequently exposed to heavy rainfall and flash floods, while the arid regions face their own challenges from sudden intense storms following prolonged dry periods. The vulnerability of these urban areas to both flooding and drought emphasizes the need for more resilient and adaptive stormwater management strategies.

Algeria's urban drainage systems, like many across the continent, were not designed to handle the increased variability in weather patterns brought about by climate change. As such, there is a critical need for modernized infrastructure, data-driven planning, and climate-resilient designs to mitigate flood risks and ensure the efficient management of stormwater. By incorporating both traditional hydrological models and emerging machine learning techniques, it is possible to enhance prediction capabilities and improve the decision-making process in managing these systems.

In conclusion, the future of urban drainage systems in Algeria and other rapidly urbanizing regions of Africa will depend on how effectively they can adapt to the evolving climate realities. This necessitates a comprehensive approach that combines engineering innovation with climate science to build resilient urban infrastructures capable of withstanding the stresses imposed by climate change.

Chapter 3 : Overview of Urban Drainage System Modelling: Hydrological, Hydraulic, and Machine Learning Approaches

3.1 Introduction

Urban drainage systems are essential for managing stormwater and preventing floods in urban areas. Traditional hydrological models and stormwater network software have played key roles in understanding and controlling these systems. However, their effectiveness is often limited by data reliance, model complexity, and computational challenges, making real-world application difficult.

Recent advancements in machine learning (ML) provide a promising alternative. ML algorithms, especially ensemble learning methods, can analyze large datasets, uncover complex patterns, and generate precise predictions, enhancing the efficiency and resilience of urban drainage systems.

This paper explores the changing landscape of urban drainage modeling, comparing the constraints of traditional methods with the benefits of ML techniques, particularly ensemble learning models. It discusses different ML approaches, such as supervised, unsupervised, deep learning, and reinforcement learning, and their relevance to urban drainage. The focus is on ensemble learning techniques, especially bagging and boosting, for improving the accuracy and reliability of flood prediction models. Through a review of recent studies, this paper highlights the potential of ML and ensemble learning to transform urban drainage management, leading to better flood prevention, resource optimization, and strengthened urban resilience.

3.2 Hydrological Models Overview

A model, according to Moradkhani and Sorooshian (2009), is a simplified representation of a real-world system. The optimal model is the one that produces results that closely align with reality while utilising the fewest parameters and minimising model complexity. Models are mostly utilised to forecast system behaviour and comprehend diverse hydrological phenomena. A model is composed of multiple parameters that determine the attributes of the model. A runoff model is a mathematical framework consisting of equations that quantify the amount of water runoff based on several parameters that describe the features of a watershed. Both rainfall data and drainage area are essential inputs for all models. In addition to these factors, water shed features such as soil qualities, vegetation cover, watershed topography, soil moisture content, and characteristics of the ground water aquifer are also taken into account. Hydrological models are currently regarded as a crucial and indispensable instrument for the management of water and environmental resources (Devia et al., 2015).

In the field of hydrological modeling, selecting the appropriate model is crucial for accurate simulation and prediction of water-related processes.

A deterministic model will consistently create the same result for a given collection of input values, while stochastic models can generate varied output values for the same set of inputs. Moradkhani and Sorooshian (2009) explain that in lumped models, the entire river basin is treated as a single unit, ignoring spatial variability. As a result, the outputs are generated without considering spatial processes. On the other hand, a distributed model divides the entire catchment into smaller units, such as square cells or a triangulated irregular network, allowing for spatially varying parameters, inputs, and outputs.

Another categorization is made between static and dynamic models, which are differentiated based on the consideration of the time factor. A static model does not take into account the element of time, whereas a dynamic model incorporates the concept of time. Moradkhani and Sorooshian (2009) categorized the models as either event-based or continuous models. The former one generates output just for certain time intervals, whereas the latter produces a continuous output.

The three major classifications in this field are empirical models, conceptual models, and physically-based models.

3.2.1 Conceptual Methods (Parametric Models)

This model includes all of the fundamental hydrological processes. The system has a series of interconnected reservoirs that serve as the physical components within a catchment area. These reservoirs are replenished through processes such as rainfall, infiltration, and percolation, and are depleted through evaporation, runoff, drainage, and other similar mechanisms. Semi-empirical equations are utilized for this purpose (H. V. Gupta et al., 1999).

The approach and model parameters are evaluated using both field data and calibration. A substantial quantity of meteorological and hydrological records is necessary for the process of calibration. The calibration process involves curve fitting, which adds complexity to the interpretation and can reduce the model's capacity to accurately predict the impact of land-use changes, as observed in the HBV model (Bai et al., 2019; Jehanzaib et al., 2020).

3.2.2 Physically Based Models

This is a mathematically abstract representation of the actual phenomenon. These models are referred to as mechanistic models, as they incorporate the fundamental principles of physical processes (Bai et al., 2019; Bin, 2015). The system employs state variables that can be measured and are dependent on both time and space. Finite difference equations are used to represent the hydrological processes of water transport. The calibration of the model does not necessitate substantial hydrological and meteorological data(Beven & Binley, 1992). However, it does involve the evaluation of a large number of factors that describe the physical characteristics of the watershed (Abbott et al., 1986)

A significant amount of data is required, including soil moisture levels, initial water depth, geographic features, topology, and the structure of the river network. The physical model holds an advantage over empirical and conceptual models because it includes parameters with physical interpretations, allowing it to address various limitations. It can produce detailed information beyond predefined boundaries and is suitable for diverse applications, as seen in models like SWMM and HEC-HMS (Chow, Maidment, and Mays 1988).

3.2.3 Empirical or Data-Driven Models

These models rely on observations and are entirely dependent on the data available, without consideration for the internal characteristics or processes of the hydrological system. Consequently, they are often referred to as data-driven models (Amatya et al., 2022; Bahrami et al., 2022). These models use mathematical equations developed from input-output time series data rather than the physical processes within the catchment, and they are applicable only within certain limits. A classic example of this approach is the unit hydrograph (UH).

Statistical approaches, such as regression and correlation models, are employed to define the functional relationship between inputs and outputs. Additionally, machine learning techniques such as artificial neural networks (ANN) and fuzzy regression are widely applied in hydroinformatics. Examples of data-driven models include the SCS-CN method, artificial neural networks, and unit hydrographs (Devia et al., 2015).

3.3 Overview of Stormwater Modeling Software

Stormwater models generally comprise both hydrologic and hydraulic components. Hydrologic models focus on examining systems that govern water movement and storage, aiming to clarify the interactions within the hydrological cycle (Solomatine & Wagener, 2011). In stormwater models, the hydrologic aspect primarily involves rainfall-runoff processes and associated calculations.(Lind, 2015) .

Hydrological modeling has seen significant advancements with the development and application of various software and tools. These tools facilitate the simulation and analysis of hydrological processes, enhancing our understanding and management of water resources.

Models for simulating stormwater quality and quantity appeared in the early 1970s, primarily developed by US government agencies such as the USEPA (V. G. Mitchell et al., 2001; Zoppou, 2001). Since then, many urban watershed models have been created. Most of these are 1D models based on the principles of mass, energy, and momentum conservation. These models may operate as either event-based or continuous simulations and are designed to model urban drainage systems until overflows from network inlets or manholes are no longer present. However, when overflows arise from limited capacity in downstream pipes or channels, accurately representing real flood extents with these 1D models becomes difficult.(Chatterjee et al., 2008) .

In stormwater modeling tools, flow can be represented in either 1-dimensional (1D) or 2-dimensional (2D) formats. In a 1D flow model, velocity depends on only one dimension, such as in pipe flow, where the flow rate is determined by the pipe's radius and a single dimension, xx (see Figure 3.1). Conversely, 2D flow relies on two dimensions, xx and yy, to define flow behavior. (Crowe, et al., 2010).

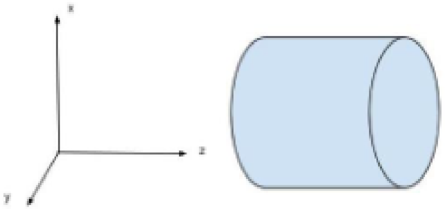


Figure 3.1: Dimensions in 1-dimensional and 2-dimensional flow within a pipe

To address the limitations of 1D models, 2D models or coupled 1D-2D flood inundation models, known as distributed flood models, were developed. Recently, the availability of distributed data, such as soil types, land use, and radar rainfall, has facilitated the development of simplified, physically meaningful distributed hydrological models (Todini, 2007) .The Table below provides an overview of some widely used hydrological modeling software, detailing their functionality, components, and 1D/2D capabilities.

Table 3.1: Overview of Widely Used some Hydrological Modeling Software selected

Software	Developed By	Components	1D/ 2D	Comments	Type	Access	Typical Applications
SWMM	U.S. EPA	Hydraulic and hydrologic modeling	1D	Latest version includes a GUI (SWMM 5.0)	Hydrologic /Hydraulic	Free	Urban drainage, stormwater management
InfoWorks ICM	Innovyze	Integrated 1D and 2D modeling	1D/ 2D	Advanced modeling for urban environments	Hydrologic /Hydraulic	Paid	Urban drainage, stormwater management, flood risk assessment
PCSWMM	Computational Hydraulics	Enhanced SWMM	1D	Includes GIS integration	Hydrologic /Hydraulic	Paid	Urban drainage,

	International (CHI)	with GIS and analytics		and advanced analytics			stormwater management
StormCAD	Bentley Systems	Stormwater drainage design	1D	User-friendly interface, focuses on design	Hydraulic	Paid	Stormwater design and analysis
SOBEK	Deltares	Integrated water management	1D/ 2D	Comprehensive modeling for flood risk and water system analysis	Hydrologic /Hydraulic	Paid	Flood risk assessment, integrated water management
SEWSYS	Lund University	Stormwater and wastewater modeling	1D	MATLAB/Simulink based, defines pollution sources within catchment	Hydrologic	Free	Stormwater and wastewater pollution analysis

3.3.1 Comprehensive Overview of MIKE+ for Urban Drainage and Flood Management

MIKE+ is a comprehensive platform developed by DHI that integrates various modules and tools designed for modeling, analyzing, and managing different aspects of water systems. This includes pipe networks, rivers, coastal areas, and more (Game et al., 2023). MIKE+ simplifies modeling by consolidating technology, allowing users to optimize water management strategies, make informed decisions, and enhance the sustainability and resilience of water infrastructure (DHI,2022). The platform supports the modeling of diverse system types, including rivers, stormwater collection systems, overland flows, and water distribution networks. It also incorporates the SWMM5 collection system format, allowing for effective urban drainage and stormwater management. SWMM 5: SWMM 5 is an advanced version of the original SWMM model that includes both 1D and 2D capabilities, enhancing its application for complex urban drainage and flood management scenarios (USEPA, 2019).

3.3.2 Key Features of MIKE+

Usability and Flexibility: MIKE+ emphasizes user-friendly interfaces and flexible workflows, allowing users to model complex water systems efficiently. It supports both 1D and 2D simulations, enabling integrated modeling of stormwater and river networks (DHI,2022).

- ✓ GIS Integration: The software offers robust GIS capabilities, allowing users to leverage spatial data for improved analysis and visualization of water systems. This integration facilitates better decision-making in urban drainage and flood management (DHI,2022).
- ✓ Data Management: MIKE+ utilizes advanced database systems (SQLite and PostgreSQL) for efficient data handling, which is crucial for managing large datasets and complex models (DHI 2024)
- ✓ Enhanced Modeling Capabilities: The software includes modules for water quality modeling and dynamic flood simulations, aiding in the development of climate adaptation strategies and flood risk assessments (DHI 2024).

3.3.3 Benefits of MIKE+ for Urban Drainage Systems

Choosing MIKE+ allows you to streamline your workflow by modeling all water systems on one platform. With MIKE+, there is no need to switch between different platforms when modeling urban water systems, rivers, and flooding scenarios. Instead, you can integrate, model, and manage all your water systems, including water distribution networks, collection systems, rivers, and flooding, on a single comprehensive platform(Boughandjioua et al. 2024).

Moreover, MIKE+ enables you to tailor your water management strategy thanks to its comprehensive array of integrated software solutions designed for water systems. The software is built on a versatile and scalable model manager core, allowing you to pick and choose different modules to create a custom solution that meets your specific water management needs.

Finally, MIKE+ maximizes efficiency and sustainability through data-driven decision support. It offers a unified suite of software solutions for water systems, centered around a flexible and scalable model manager core. With the MIKE+ platform, you can select and pay only for the modules you need, including options for modeling pollutant transport, sediment transport, and water quality. This modular approach helps you make the most of your resources and improve water management efficiently and sustainably.

3.3.4 Applications and Effectiveness of MIKE+ in Flood Management

The MIKE model suite by DHI demonstrates exceptional capabilities in flood management. Tina Moni Boruah (2021)focuses on developing 1D and 2D hydrodynamic flood models using the MIKE software suite by DHI. The paper emphasizes the practical application and contribution of these models to flood management and disaster mitigation efforts. Similarly, Game et al. (2023) utilized using DHI MIKE Plus 2023 model to develop a 1D-2D coupled model for real-time flood management in Nice, France. The model accurately simulates and predicts flood events, demonstrating a strong correlation between observed and predicted values. It also supports urban development by providing real-time data.Vidyapriya and Ramalingam (2012) demonstrates the application of MIKE URBAN software for flood modeling in an urban watershed. Their work helps understand flood dynamics in rapidly urbanizing areas and provides valuable tools for flood management and mitigation, enhancing the resilience of cities like Chennai to flooding events.

These studies underline the practical benefits of the MIKE model, such as improved flood management, better understanding of urban flood dynamics, and enhanced disaster mitigation strategies. Including explicit references and quantitative data from these studies could provide a clearer demonstration of the software's accuracy and effectiveness in practical applications (Saheed & Misra, 2024).

3.3.5 Limitations of Traditional Hydrological and Stormwater Models

Hydrological models and stormwater network software are essential tools for managing water resources, but they come with various limitations that can affect their effectiveness in real-world applications.

a) Limitations of Hydrological Models

1. **Data Dependency:** Hydrological models require extensive input data, including climatic, soil, and land use data (Ramovha et al., 2024; Shahed Behrouz et al., 2020) .

2. Simplification of Natural Processes: These models are approximations of natural processes, relying on simplified equations to represent complex hydrological dynamics. This simplification can lead to inaccuracies, particularly when predicting responses to extreme weather events or changes in land use.

3. Calibration Challenges: Model calibration is crucial for ensuring accuracy, but it can be complex and resource-intensive. Models may require adjustments to input variables to align outputs with observed data, which can be difficult to achieve without sufficient data points. Over-calibration can also limit the model's flexibility to predict under varying conditions(Shahed Behrouz et al., 2020) .

4. Computational Limitations: Some models may be computationally intensive, requiring significant processing power and time, especially for high-resolution simulations. This can restrict their use in scenarios where quick decision-making is necessary.

5. Uncertainty and Error: There is often a high degree of uncertainty associated with model outputs due to inherent errors in input data and the modeling process itself. This uncertainty can complicate the interpretation of results and the decision-making process based on those results (Hossain et al., 2019; Shahed Behrouz et al., 2020).

b) Limitations of Stormwater Network Software

1. User Expertise: The effectiveness of stormwater network software can be compromised by the user's understanding of the model. Sophisticated models may yield poor results if the user lacks the necessary expertise to operate them effectively, leading to reliance on these tools as "black boxes" without a proper understanding of their limitations (Ramovha et al., 2024).

2. Software Support and Updates: Not all software comes with adequate support or regular updates. This can lead to outdated methodologies being used, which may not reflect the latest research or best practices in stormwater management (Ramovha et al., 2024).

3. Model Specificity: Different software tools may be tailored for specific applications, which can limit their versatility. For instance, while some models focus on urban runoff, others might be better suited for rural settings or specific types of stormwater control measures, making it challenging to find a one-size-fits-all solution(Ramovha et al., 2024).

4. Cost and Accessibility: Some advanced modeling tools can be expensive, both in terms of initial purchase and ongoing maintenance. This can limit access for smaller municipalities or organizations with constrained budgets (Ramovha et al., 2024; Shahed Behrouz et al., 2020).

3.4 Machine Learning Methods for Urban Drainage Systems

Machine learning is a branch of artificial intelligence that concentrates on creating algorithms and statistical models that allow computers to learn from data and make predictions or decisions autonomously, without the need for explicit programming (T. M. . Mitchell, 1997) .It involves training models on historical data to identify patterns and make inferences or predictions about new data.

Recent research has highlighted the growing interest in applying machine learning to urban drainage system modeling. Several studies have explored the use of machine learning methods to improve the operation and control of drainage systems(Meijer et al., 2023). Machine learning techniques have been increasingly used in urban drainage systems as an alternative to classical models due to their numerous benefits (Barua et al., 2023). These benefits include improved accuracy and predictive capabilities (Khan et al., 2022), the ability to handle complex and non-linear relationships (Vanpoucke et al., 2020), and the potential for real-time monitoring and control (Okoro, 2023). Machine learning algorithms are highly effective in analyzing extensive and varied datasets, enabling more detailed and comprehensive modeling of urban drainage systems (S. Gupta et al., 2024).

Furthermore, machine learning methods can continuously adapt and improve with new data, making them well-suited for dynamic and changing urban settings (Jiang et al., 2021). Overall, the use of machine learning in urban drainage systems offers the potential for more efficient and effective management of stormwater and flood control(Jagaba et al., 2021).

In our study, we provide a comprehensive overview of various machine learning techniques employed to model urban drainage systems, including supervised learning, unsupervised learning, deep learning, reinforcement learning, and hybrid methods (UDSs) (Boughandjioua et al. 2023), as depicted in Figure 3.2.

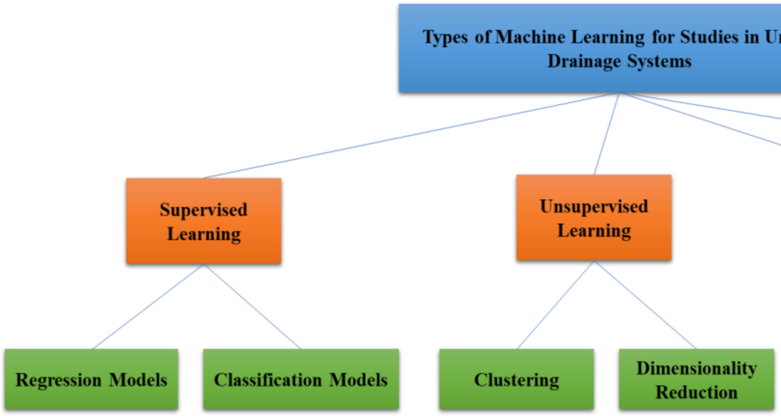


Figure 3.2: Types of ML for UDSs

In the field of urban drainage systems, the application of machine learning techniques has grown significantly to enhance various functionalities. These methods are employed in areas such as stormwater management and overflow forecasting for early warning systems, demonstrating their ability to improve the efficiency and effectiveness of urban drainage systems by utilizing the extensive data collected from sensors installed within these systems (Di Nardo et al., 2021). Machine learning approaches can address several important challenges faced by urban drainage systems. To categorize the different methodologies in machine learning studies, the following sections can be considered (Boughandjioua et al. 2023).

3.4.1 Supervised Learning

Supervised learning is a type of machine learning algorithm that learns from labeled data, as illustrated in Figure 3.3. Labeled data consists of input data and the corresponding output labels. The algorithm learns from the labeled data to make predictions about new, unseen data (T. M. Mitchell, 2010).

Supervised learning can be categorized into two distinct algorithmic categories:

- Regression Models: Predict continuous values, such as overflow rates or water levels. Examples include Linear Regression (Han et al., 2011), Support Vector Regression, and Decision Trees (Breiman et al., 2017).
- Classification Models: Categorize data into classes, such as predicting the likelihood of system failure or overflow. Examples include Logistic Regression (Cox, 1958), Random Forest Classifiers (Breiman, 2001), and Gradient Boosting Machines(Friedman, 2001).

The applications of supervised learning in Urban Drainage System Modeling include:

a) Predicting rainfall-runoff relationships: By analyzing historical rainfall and runoff data, a supervised learning model can be trained to predict future runoff based on forecasted rainfall (Gauch et al., 2020).

b) Improving the accuracy of flood inundation mapping: By combining high-resolution topographic data and historical flood events, a supervised learning model can be trained to predict the extent and depth of flooding for different rainfall scenarios(Prakash et al., 2021) .

c) Optimizing the design of drainage systems: By simulating different drainage system designs under various rainfall scenarios, a supervised learning model can be used to identify the most efficient and cost-effective design (Fu et al., 2022).

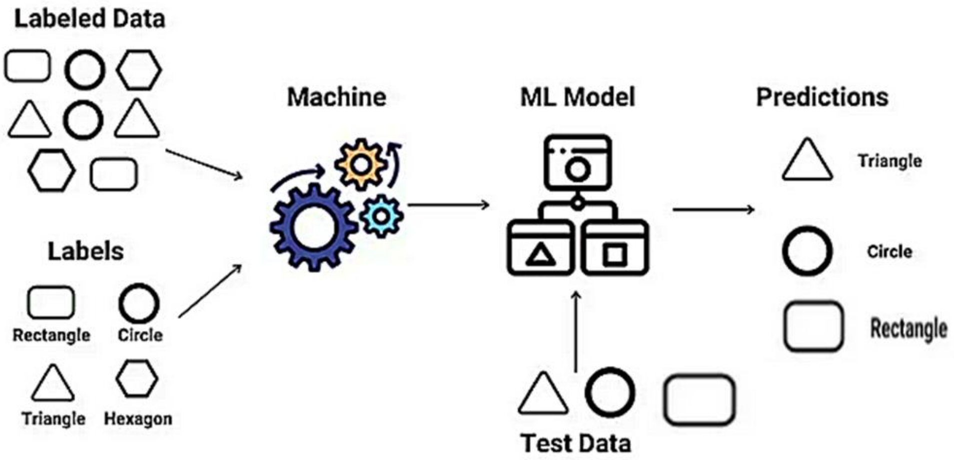


Figure 3.3: Supervised Learning Process

3.4.2 Unsupervised Learning

This is a type of machine learning algorithm that learns from unlabeled data, as illustrated in Figure 3.4. Unlabeled data consists of input data without any corresponding labels. The algorithm learns from the unlabeled data to identify patterns and structure in the data(T. M. Mitchell, 2010).

- Clustering: Group similar data points, which (MacQueen, 1967) can be useful for identifying patterns or anomalies in drainage systems. Examples include K-Means Clustering and Hierarchical Clustering(Johnson, 1967).
- Dimensionality Reduction: Simplify models by reducing the number of features while retaining essential information. Techniques like Principal Component Analysis (PCA) (Greenacre et al., 2022) and t-Distributed Stochastic Neighbor Embedding (t-SNE) (Van Der Maaten & Hinton, 2008)are commonly used.

The applications of unsupervised learning in Urban Drainage System Modeling include:

a) Classifying land use types: By analyzing satellite imagery or other geospatial data, an unsupervised learning model can be used to classify different land use types based on their impact on runoff (Saini & Rawat, 2023)

b) Identifying potential sources of pollution: By analyzing water quality data, an unsupervised learning model can be used to identify potential sources of pollution in a drainage system (Aslam et al., 2022).

c) Detecting anomalies in drainage system performance: By analyzing sensor data from a drainage system, an unsupervised learning model can be used to detect leaks, blockages, or other anomalies(Yan & Tao, 2022) .

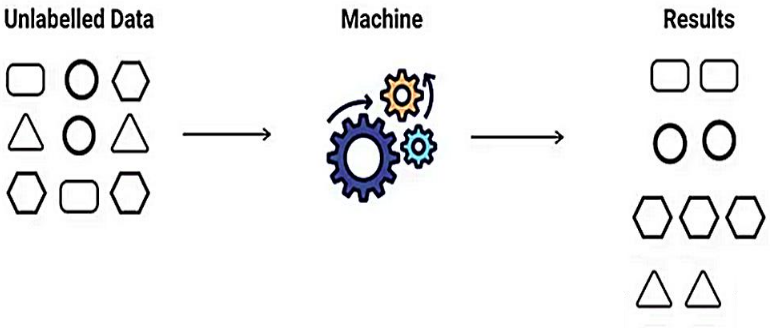


Figure 3.4: Unsupervised Learning Process

3.4.3 Deep Learning

It is a type of machine learning that uses artificial neural networks (ANNs) with multiple layers to learn from data, as illustrated in Figure 3.5. ANNs are inspired by the structure of the human brain, and they can learn complex patterns from data (Heaton, 2018). Examples include Convolutional Neural Networks (CNNs) (Lecun et al., 2015) for spatial data and Recurrent Neural Networks (RNNs) (Rumelhart et al., 1986) for sequential data.

The applications of deep learning in Urban Drainage System Modeling include:

a) Enhancing Rainfall-Runoff Prediction Accuracy: The models can effectively capture the intricate physical processes governing rainfall-runoff relationships, leading to more accurate runoff predictions(Fu et al., 2022) .
b) Developing Real-time Control Systems for Drainage Systems: Deep reinforcement learning (DRL) models can optimize drainage system operations in real-time, considering factors like rainfall intensity and drainage capacity (Mullapudi et al., 2020a)
c) Automating Drainage System Inspection and Maintenance: The models can analyze sensor data and inspection images to identify potential issues and prioritize maintenance needs (Uludağ 2022).

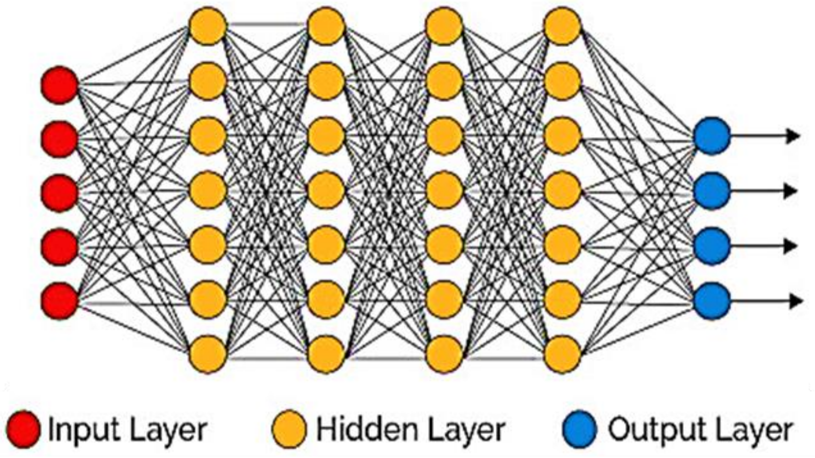


Figure 3.5: Deep Learning Neural Network

3.4.4 Reinforcement Learning

This is a type of machine learning algorithm that learns by interacting with its environment, as illustrated Figure 3.6. The algorithm learns to take actions that maximize a reward signal (Sutton & Barto,2018.). The applications of Reinforcement Learning in Urban Drainage System Modeling include:

a) Optimizing Pump and Valve Operation: RL models can determine the optimal operation of pumps and valves in a drainage system, balancing stormwater removal with energy conservation (Zhang et al., 2022)
b) Developing Adaptive Control Systems for Drainage Systems: RL models can adapt drainage system operations to changing conditions, such as rainfall patterns or land use changes(Mullapudi et al., 2020b).

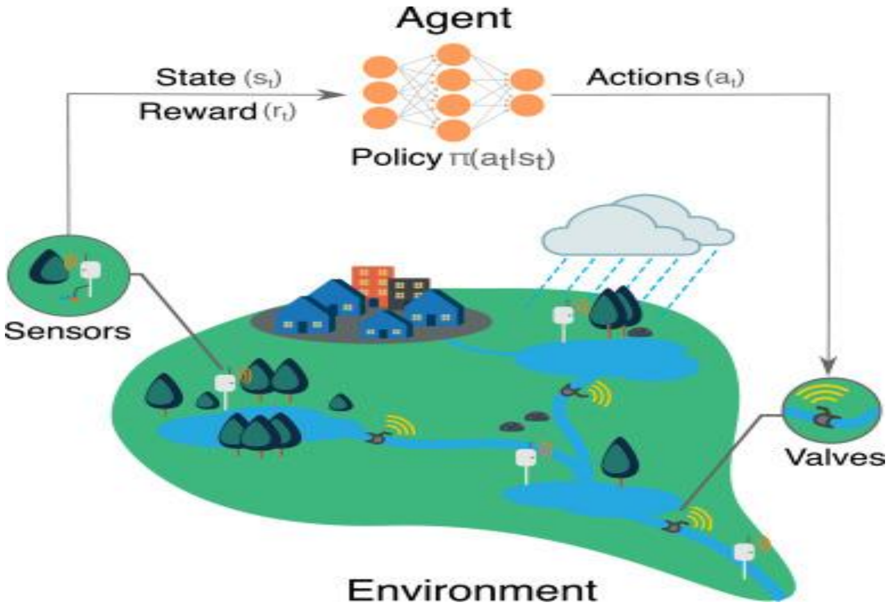


Figure 3.6: Enhancing Stormwater Control through Real-Time Deep Reinforcement Learning Approaches (Mullapudi et al. 2020).

3.4.5 Hybrid Approaches

A hybrid approach in urban drainage systems combines decentralized and centralized strategies to manage excess stormwater and enhance urban flooding resilience1. This integrated approach aims to address the challenges of urban water management by combining various techniques and technologies to optimize water resources and reduce the risk of flooding(D'ambrosio et al., 2021).The hybrid approach can be applied in urban drainage systems through the following components:

a) Decentralized Strategies: These strategies focus on capturing and storing stormwater at the source, thereby reducing the flow of water to the centralized stormwater drainage system. Examples of decentralized strategies include green roofs, infiltration trenches, and permeable pavements (Bakhshipour et al., 2019).
b) Centralized Strategies: These strategies involve the use of traditional stormwater drainage systems, such as storage and infiltration, located downstream of subcatchments1. These systems help attenuate the peak flow of stormwater and reduce the risk of flooding(D'ambrosio et al., 2021).

c) Real-time flood forecasting: Hybrid models can be used to combine real-time rainfall data with historical data and physical models to make more accurate predictions of flood inundation and potential damage(Noymanee & Theeramunkong, 2019).

Table 3.2 presents a comparison of various machine learning techniques applied to stormwater management and flood forecasting. It summarizes key studies, the datasets used, the techniques employed, and the findings for each approach.

Table 3.2: Comparison of Machine Learning Techniques for Stormwater Management and Flood Forecasting(Samira BOUGHANDJIOUA, 2023)

Type of Machine Learning	**Authors**	**Data Set Utilized**	**Techniques Employed**	**Findings**
Supervised Learning	Noymanee & Theeramunkong, 2019	Water-level records from 2012-2016	Linear regression, neural network regression, Bayesian linear regression, boosted decision trees	Enhanced modeling reduces runoff forecasting error.
	H. Wang & Song, 2020	Water levels	SVM-based machine learning method	Highlights accuracy and speed; traditional models can be complex.
	Ke et al., 2020	Rainfall intensity	Machine Learning methods	Models classify flooding with 96.5% accuracy; false alert rate reduced to 25%.
	Dtissibe et al., 2020	Discharge	Physical-based flood forecasting; multilayer perceptron neural network	Model demonstrates effectiveness and good forecasting capabilities.
	Kwon & Kim, 2021	Hydrological data	Various ML techniques (ANN, CNN, LTM)	Extensive ML use enhances performance and efficiency in urban drainage systems.
Deep Learning	Rjeily et al., 2017	Rainfall intensity and water depth variations in manholes	Nonlinear Auto Regressive with eXogenous	Effective in predicting changes in manhole water depth during storms.
	Cruz et al., 2018	Real-time monitoring sensors (rain gauges, water levels, soil moisture)	Multi-layer artificial neural network using MATLAB	High goodness-of-fit metrics and low RMSE.
	Le et al., 2019	Daily discharge data	Data-based methods	Achieved forecasting accuracies of 94.5% and 86.2%.

	Mousavi et al., 2021	Data from four gauging stations in Pennsylvania (2013-2019)	IoT with LoRaWAN technology; Deep learning models	Outperform traditional models in flood detection.
	J. Chen et al., 2023	Rainfall scenarios	Combination of LSTM neural networks with hydrological and hydrodynamic models	High accuracy with minimal deviation from measured data.
Reinforcement Learning	Mullapudi et al., 2020c	Simulated storm scenarios	Real-time control using Reinforcement Learning and Deep Neural Networks	Effectively manages stormwater control; performance relies on reward formulation.
	Bowes et al., 2020	Sensor and forecast data	Deep Deterministic Policy Gradient (DDPG) algorithm	Reduces flood volume by an average of 70.5%; comparable to passive systems.
Hybrid Approaches	Y. Wang et al., 2019	Historical hydrological data	BP neural network combined with a genetic algorithm	Improved flood forecasting precision, integrating local insights.
	Kan et al., 2020	Rainfall and antecedent runoff	ANN and K-nearest neighbor method	Successfully forecasts peak flow with high accuracy.
	Keum et al., 2020	Flood databases	Combining hydraulic theory with machine learning; uses Latin hypercube sampling	Achieves 85% goodness-of-fit in a short runtime.
	Motta et al., 2021	Local weather measurements and fire department data	Random Forest classifiers; GIS Hot Spot analysis	High accuracy and developed flood risk index.
	Kumar & Biradar, 2022	Historical rainfall data, divers duration, and water flow	Neural network model utilizing fuzzy and sigmoid functions	Delivers optimal results, outpacing existing techniques.

The main components of the various machine learning methodologies employed in urban drainage systems can be succinctly summarized as follows:

a) The Supervised Learning techniques utilize conventional ML algorithms to perform diverse tasks, demonstrating their adaptability in tackling different aspects of flood prediction.
b) Deep Learning Techniques are highly effective in managing intricate jobs and live monitoring, offering impressive predictive capabilities in flood forecasting situations.

c) Reinforcement Learning Techniques are specifically designed to achieve real-time control and have proven to be highly effective in adapting stormwater systems to prevent flooding.
d) Hybrid approaches utilize the advantages of several machine learning techniques by integrating neural networks with genetic algorithms, expert systems, and GIS techniques. This combination enhances the accuracy and robustness of flood prediction systems.

These studies emphasize the numerous methodologies and their efficacy in predicting floods and managing urban drainage systems. The selection of techniques typically relies on the particular dataset and research goals, with deep learning and hybrid methodologies commonly attaining notable precision and efficiency.

3.4.6 Ensemble Learning Models

Ensemble learning merges the predictions of several models to improve accuracy, making it a robust method in machine learning. This approach has two main motivations: (Cheriguene et al., 2018).

- Performance: Ensemble methods often provide better predictions and higher overall performance than the individual models that contribute to the ensemble. (Cheriguene et al., 2018).
- Improved Accuracy and Reliability: Different models capture various patterns within data, and combining them creates a robust, reliable predictor. This integration is particularly effective when achieving high accuracy is crucial, as it minimizes individual model weaknesses and amplifies strengths (Cheriguene et al., 2018).
- Reduction of Overfitting: Using an ensemble reduces the tendency of individual models to memorize training data, especially high-variance models like decision trees. By aggregating predictions, ensemble models offer smoother and more generalized predictions, making them less susceptible to noise or outliers (Cheriguene et al., 2018).

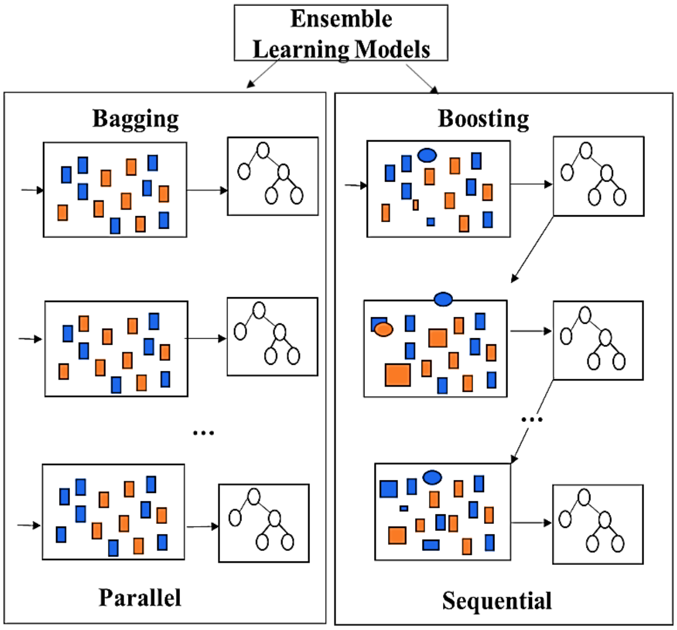


Figure 3.7: Ensemble Learning Models (ELM) - Bagging and Boosting

Methods like Bagging and Boosting achieve this through different strategies, like re-weighting training data or creating diverse subsets of data for model training. Stacking, a third ensemble approach, combines predictions of varied models through a meta-learner, leveraging each model's unique contributions for a more comprehensive result, as shown in Figure 3.7 (Cheriguene et al., 2019).

3.4.6.1 Bagging and Boosting Techniques

a) Bagging

The Bagging algorithm short for Bootstrap Aggregating was introduced by Breiman in 1996 as an ensemble learning technique to reduce model variance ((Zounemat-Kermani 2017; Breiman in 1996). This method involves training multiple models on different subsets of the training data and then combining their predictions for a final, more stable result. Bagging achieves this by randomly selecting subsets with replacement from the original training data (Zounemat-Kermani et al., 2021). This means some data points may appear multiple times within a subset, while others may be excluded entirely. Each model is trained on these unique subsets, which helps to reduce prediction variability, as each model captures slightly different aspects of the data. This diversity is particularly effective in mitigating overfitting, especially for high-variance models that might otherwise adapt too closely to the training data. Bagging is widely used across domains such as classification, regression, and time series forecasting due to its ability to generate a variety of predictions and enhance model stability (Polikar 2009; Wu et al. 2020).

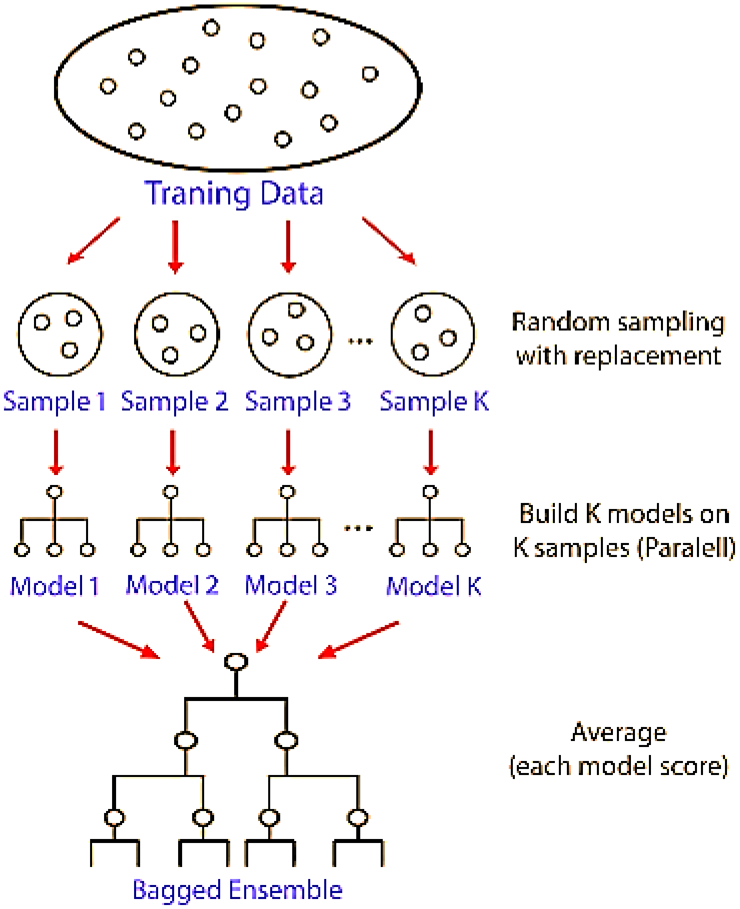


Figure 3.8: Ensemble Learning Models (ELM) - Bagging (Zounemat-Kermani et al., 2021)

b) Boosting

Boosting is a well-established ensemble learning technique designed to improve the accuracy and performance of machine learning models by building a series of models sequentially. The core concept in boosting is to enhance "weak learners" (models that perform slightly better than random guessing) into strong learners by focusing on errors made by previous models in the sequence. Each new model attempts to correct errors from the previous one, which makes the ensemble progressively more accurate (Zounemat-Kermani et al., 2021).

Unlike bagging, where models are trained independently, boosting creates a dependency between models by adjusting each subsequent model based on the performance of the last. Initially, the first model is trained on the full dataset. Afterward, a new model is added and trained to correct the errors of the previous one, adjusting its focus to the more challenging cases. This process continues, with each model receiving a weighted vote depending on its accuracy. In the end, these weak learners are combined often through a weighted majority voting mechanism to form a robust, highly accurate model (Zounemat-Kermani, 2017).

This technique helps reduce both variance and bias and is particularly effective in preventing overfitting in high-variance models like decision trees. Boosting is widely applied across tasks requiring precise predictions, including classification, regression, and time series forecasting due to its ability to manage complex patterns within datasets. (Schapire 2002; Alfaro et al. 2013). Figure 3.9 depicts the process of the Boosting algorithm.

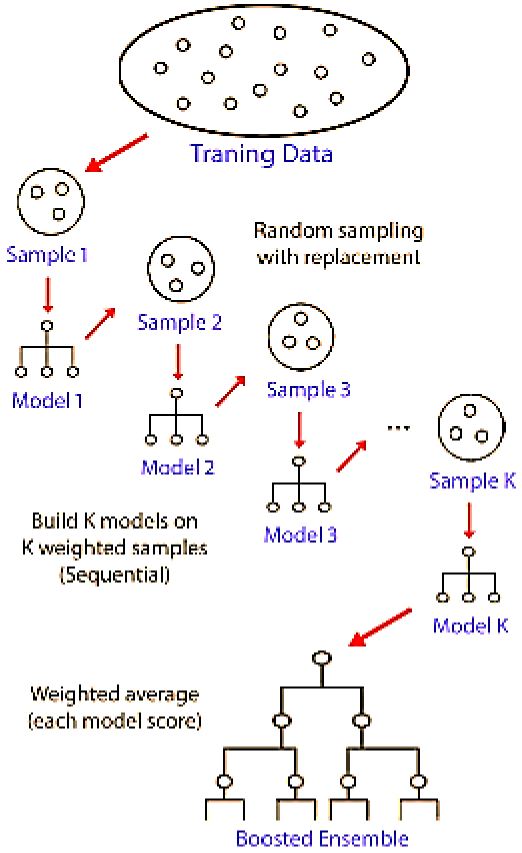


Figure 3.9: Ensemble Learning Models (ELM) – Boosting (Zounemat-Kermani et al., 2021)

In this study, predicting the average overflow rate in the stormwater network across varying rainfall durations and return periods involves applying ensemble learning techniques. Specifically, the bagging approach is implemented using the Random Forest algorithm, which improves prediction accuracy through multiple decision trees. On the other hand, the boosting technique uses Gradient Boosting, which refines model predictions by sequentially correcting errors in previous iterations, resulting in high-performance outcomes. Table 3.3 provides a detailed comparison of the key characteristics of both Random Forest and Gradient Boosting techniques.(Salim & Bhattacharyya, 2023):

Table 3.3: Key Features of Random Forest and Gradient Boosting

Feature	Random Forest	Gradient Boosting
Overfitting Risk	Reduced by averaging multiple decision trees, lowering overall variance and prediction error.	Bias is reduced by sequentially building and correcting weak prediction models.
Accuracy	High accuracy in both regression and classification tasks.	High accuracy, especially in cases like logistic regression and shallow decision trees.
Robustness	Robust to noise and missing values, suitable for incomplete datasets.	Can be sensitive to noise but generally robust through iterative correction of errors.
Computational Efficiency	Moderate; depends on the number of trees and features.	Enhanced by selectively choosing features and reducing dimensionality during training.

3.5 Review of Ensemble Machine Learning Applications in Hydrological Systems and Flood Modeling

Ensemble learning has emerged as a transformative tool in hydrological modeling, offering notable improvements over traditional single-model approaches. This review examines the application of ensemble learning in hydrology, highlighting its advantages and discussing key studies that showcase its effectiveness. Hydrological systems are inherently complex, influenced by various factors such as meteorological conditions, environmental changes, and human activities. Traditional hydrological models often struggle to capture this complexity, which can lead to uncertainties in simulation results. To address these challenges, researchers have developed diverse modeling frameworks including physical, statistical, stochastic, mathematical, and numerical approaches that each offer unique strengths in simulating the interactions within hydrological systems. These frameworks are essential for accurately predicting water flow, distribution, and other hydrological processes (Yuan et al., 2019). Recently, machine learning (ML) techniques have gained prominence in hydrological modeling due to their capacity to handle complex, nonlinear relationships. ML techniques are particularly valuable in hydrology for their ability to detect patterns, trends, and underlying rules from large, intricate datasets, enhancing prediction accuracy in hydrological forecasts (Rajaee et al., 2020; Zounemat-Kermani & Scholz, 2013). Within supervised learning in hydrology, two main strategies stand out: standard learning and ensemble learning.

Ensemble learning techniques combine multiple models to improve predictive accuracy, making them particularly promising for hydrological applications. These techniques mitigate uncertainties by enhancing the performance of individual machine learning models. Ensemble methods have proven to be effective in simulating surface hydrological processes, increasing the robustness and precision of predictions (Adnan et al., 2021). A number of studies highlight the applications and benefits of ensemble learning in hydrology. For instance, Zounemat-Kermani et al. (2021) provided a comprehensive review of ensemble methodologies in hydrological modeling, illustrating how ensemble techniques outperform traditional single-model approaches. Their review details the ensemble process from generating base models to combining them for robust predictions and covers diverse applications, including river and streamflow modeling, rainfall-runoff processes, groundwater modeling, flood susceptibility mapping, and drought forecasting. This study emphasizes the importance of continued research into novel ensemble learning strategies to further enhance model reliability. Recent studies reaffirm the effectiveness of ensemble learning in hydrology. For example, Hukkeri et al. (2023) combined ensemble learning with advanced ML models to improve meteorological drought predictions, while Dang et al. (2024) used Random Forest (RF) to capture interactions between input variables and flood depth, resulting in accurate predictions. The performance of RF made it an ideal candidate for urban flood depth modeling. Similarly, Chen et al. (2018) and Hukkeri et al. (2023) showed that ensemble learning models deliver enhanced accuracy in predicting stormwater overflow rates, underscoring the strengths of ensemble methods in flood forecasting and management.

In conclusion, ensemble learning has proven to be a valuable approach in hydrological modeling, enhancing predictive accuracy and robustness compared to traditional single-model methods. The reviewed studies illustrate the diverse applications and benefits of ensemble learning in hydrology, including flood prediction and management, groundwater modeling, and drought forecasting. Continued research into innovative ensemble learning strategies and the integration of multiple approaches will further improve the reliability and accuracy of hydrological models, making a significant contribution to water resource management and disaster risk reduction. The table below summarizes some selected studies that utilize ensemble learning in flood prediction.

Table 3.4: Review of Selected Studies Utilizing Ensemble Learning in Flood Research

Authors (Year)	Year	Research Motivation	Ensemble Methods Applied
Shu and Burn	2004	Index flood estimation	Bagging, Boosting
Tiwari and Chatterjee	2010	Flood forecasting	Bagging
Li et al.	2016	Flood forecasting	Boosting
Lee et al.	2017	Flood susceptibility prediction	Bagging, Boosting
Shafizadeh-Moghadam et al.	2018	Flood susceptibility mapping	Boosting
Venkatesan and Mahindrakar	2019	Flood forecasting	Boosting
Arabameri et al.	2020	Flash flood susceptibility modeling	Bagging
Pham et al.	2021	Flood susceptibility mapping	Bagging

The results emphasize the critical importance of ensemble learning methods in improving flood management and mitigation strategies, showcasing their effectiveness as an alternative to traditional models for tackling the complexities and limitations of hydrological systems.

3.6 Conclusion

This research emphasizes the transformative potential of machine learning (ML) and ensemble modeling for managing urban drainage systems. It shows how data-driven methods can overcome traditional hydrological model limitations, leading to more efficient and resilient hydraulic systems. The study's findings indicate that ML can refine flood prediction, streamline hydraulic operations, and adapt to dynamic hydraulic conditions, also enabling automated maintenance. Integrating both decentralized and centralized stormwater strategies enhances urban hydraulic resilience. Further exploration of ML and ensemble techniques is essential for creating safer, sustainable urban hydraulic net.

Chapter 4 : Study Area and Methodology

4.1 Introduction

This chapter explores the intricate hydrological patterns of Bir Farina, Algeria, focusing on the interactions between environmental factors and urban land use. It highlights the importance of efficient stormwater management in a region with a Mediterranean climate, where fluctuations in annual rainfall and temperature significantly affect water movement and storage. The study aims to identify the factors contributing to flooding in the area, using advanced modeling techniques and machine learning to improve prediction accuracy.

A central aspect of the research is the integration of hydrological and hydraulic models with data from storm events, using MIKE+ software to simulate urban flooding scenarios. This comprehensive analysis not only assesses the performance of the stormwater network model but also stresses the importance of well-planned infrastructure in reducing flood risks. Additionally, the chapter presents a new methodology, SWN-ML (Stormwater Network-Machine Learning), which utilizes machine learning techniques to enhance overflow rate predictions.

In conclusion, the chapter offers a detailed analysis of stormwater management in Bir Farina, providing valuable insights for urban resilience planning and flood risk management in similar settings.

4.2 Description of the Geographical Location and Characteristics of the Study Area

The Bir Farina study region is located in the northeastern part of Algeria, within the Azzaba city department of Skikda. It is geographically bounded by latitudes 36°44’48.91” to 36°45’9.15” North and longitudes 7°6’52.10” to 7°7’48.69” East, as illustrated in Figure 4.1. Covering an area of 42 hectares, Bir Farina experiences a Mediterranean climate characterized by notably hot and dry summers as well as mild and wet winters. (Zaidi et al., 2021). In this study, the Bir Farina area in Azzaba city has altitudes ranging from 85 m to 72 m. It receives about 727 mm of rainfall annually, with an average temperature of around 20 °C. The land slopes gently between 0 and 1%. Dominated by clayey loam, clay, and saline soil types, the soils fall under categories C and D based on FAO's Land Use and Soil Maps (FAO-soilGrids250m). Land use is divided into seven types: residential, administrative, roadways, sidewalks, green spaces, agricultural areas, and forested land. (Laouacheria and Mansouri 2015).

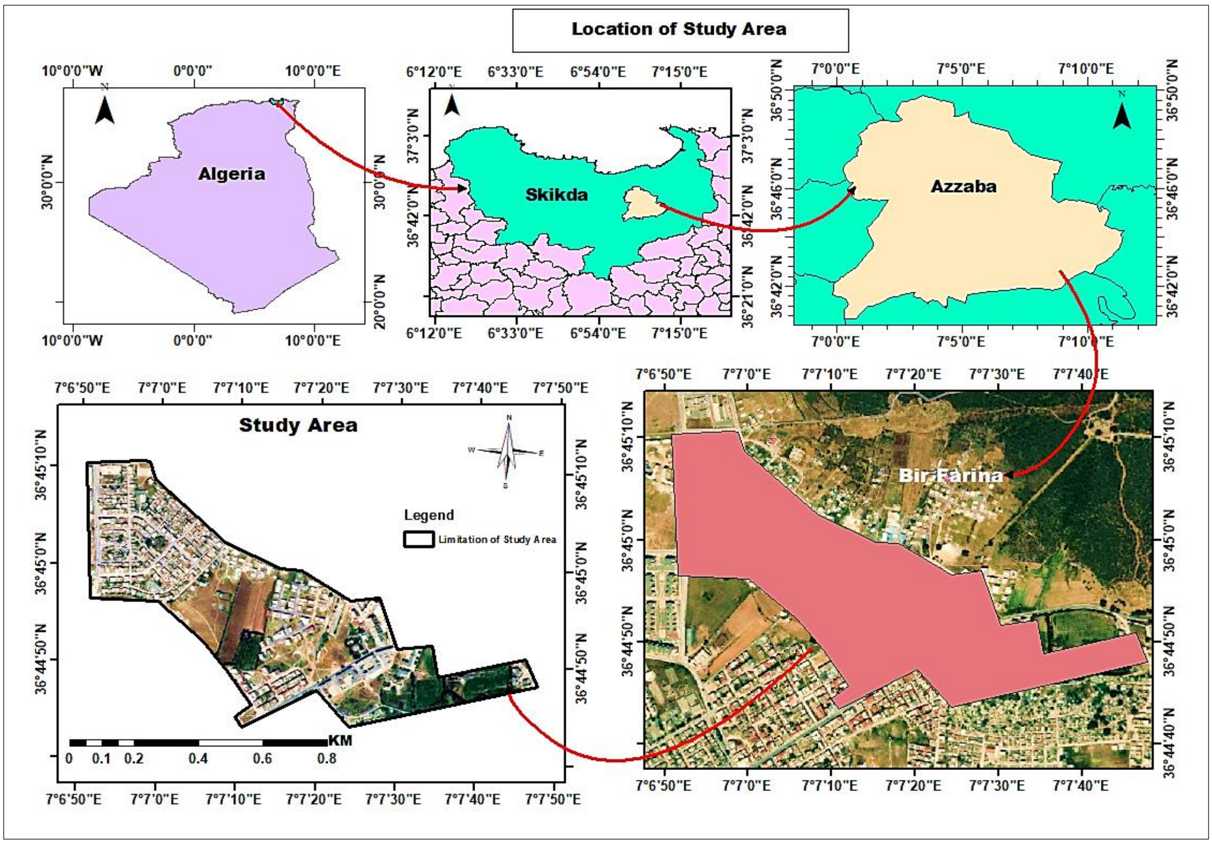


Figure 4.1: Study Area Geographic Location

4.2.1 Stormwater Network Model

Accurate stormwater network data is essential for establishing the 1D-2D coupled stormwater network model. This model is used to simulate pluvial floods and analyze the interactions between stormwater network manholes and overland flow. The study area, covering approximately 42 hectares, was partitioned into 97 sub-catchments using the Thiessen polygons method. This method directs water flow towards manholes, as seen in Figure 4.2. The average percentage of impervious area in the sub-catchments is 85%. The width of the sub catchments ranges from 6.01 m to 118 m. The curve number ranges from 77 to 94.

The stormwater network consists of 306 manholes, each with a single outlet, and 306 pipes, with a total length of 8.33 kilometers. All pipelines are constructed using reinforced concrete round pipes. The pipe diameters range from 300 mm to 1400 mm, while the pipe slopes vary between 0.5% and 5%, as illustrated in Figure 4.3. The Manning roughness coefficient is 0.012. Figure 4.2 shows the existing stormwater network model in MIKE+.

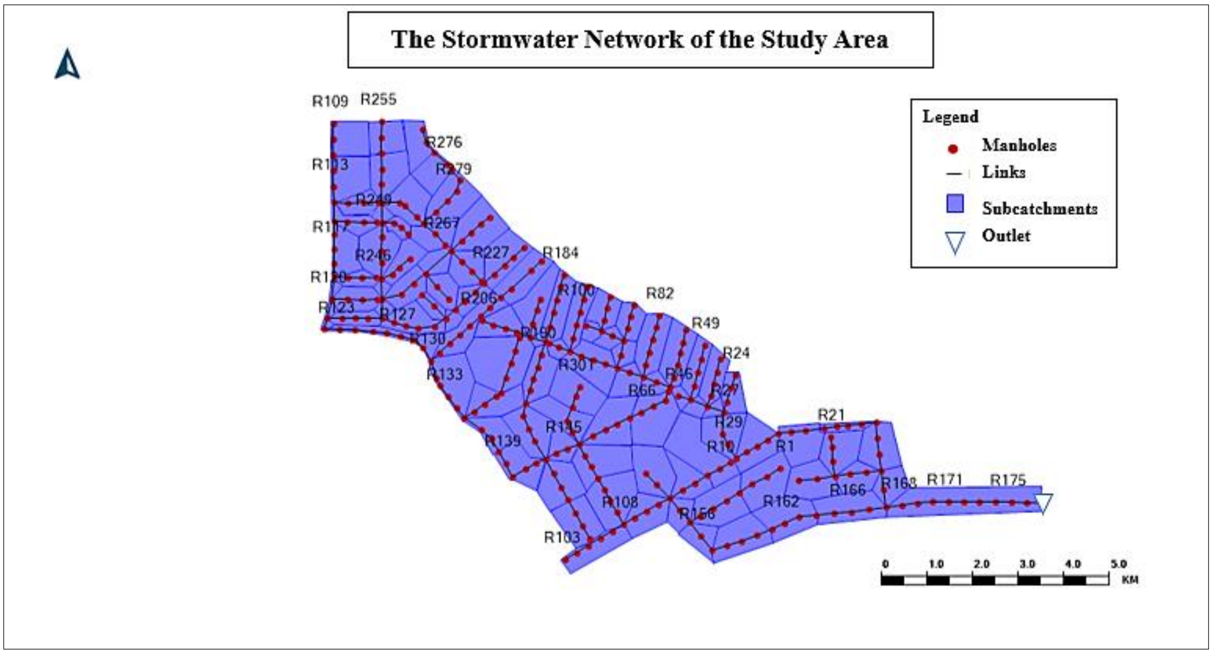
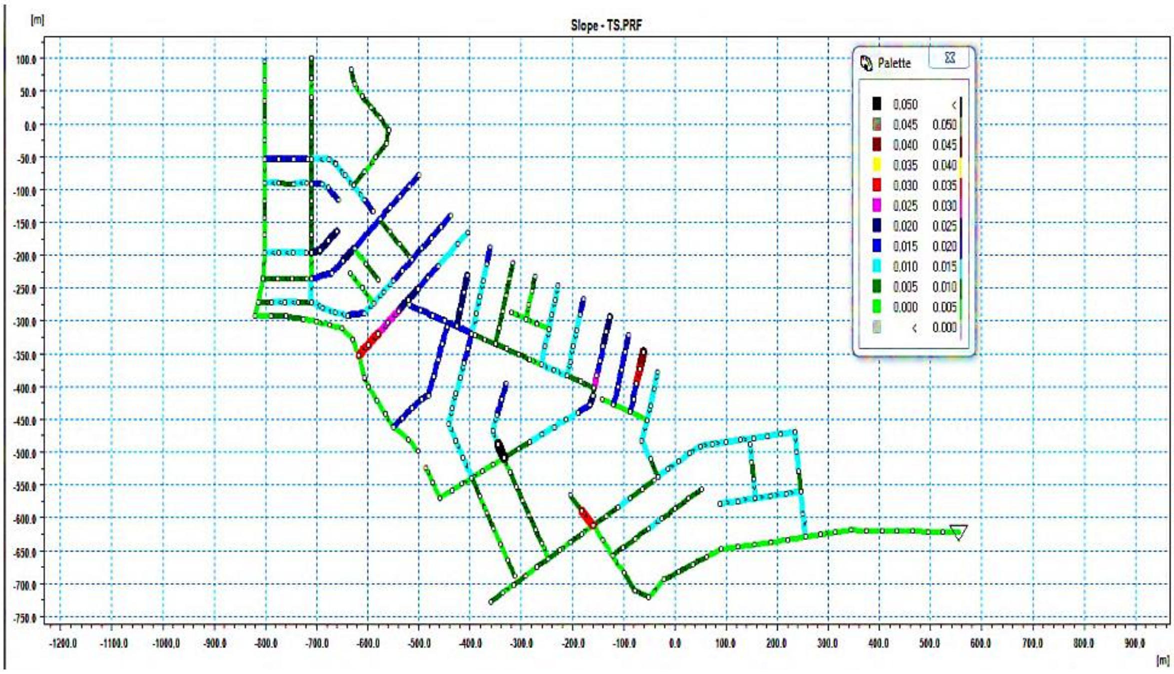
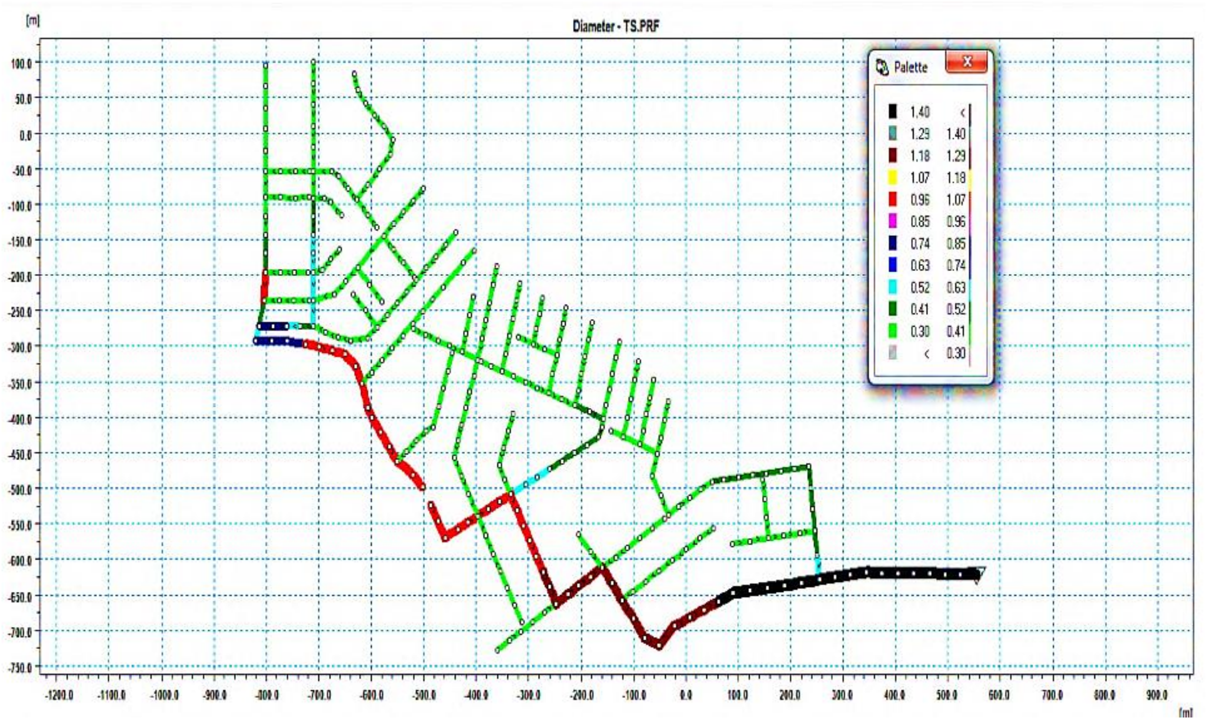


Figure 4.2: Stormwater Network Model Constructed in MIKE+ Software



■	0.050	<
■	0.045	0.050
■	0.040	0.045
■	0.035	0.040
■	0.030	0.035
■	0.025	0.030
■	0.020	0.025
■	0.015	0.020
■	0.010	0.015
■	0.005	0.010
■	0.000	0.005
■	<	0.000

a) Slop



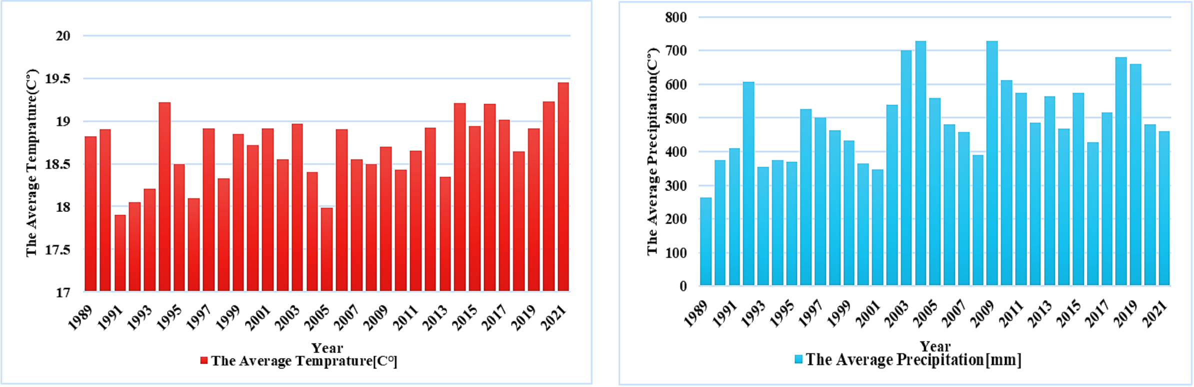
b) diameter

Figure 4.3: Slope and Diameter Characteristics of the Stormwater Network

4.2.2 Climate Characteristics and Hydrological Modeling for the Study Area

4.2.2.1 Climate Classification and Precipitation Data

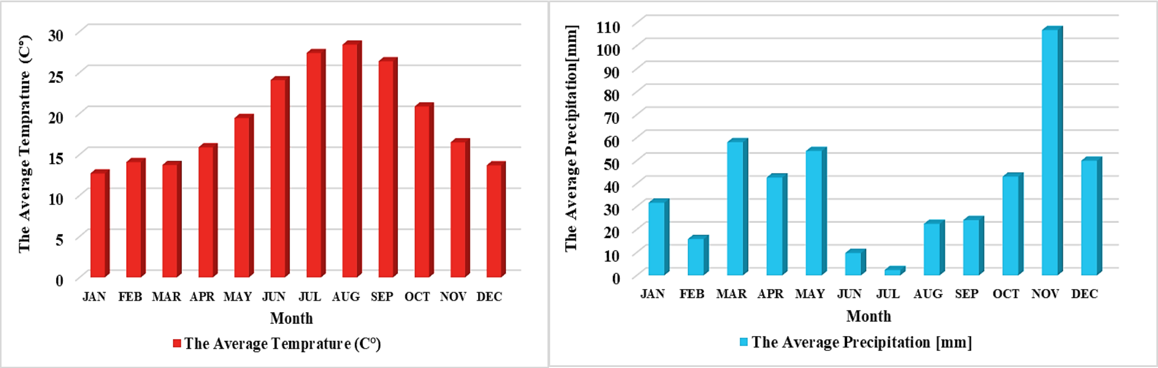
The study area of Azzaba, located in the department of Skikda in Algeria, is characterized by a Mediterranean climate. This climate features two distinct seasons: a mild, rainy winter and a hot, dry summer(Zaidi et al., 2021).Climatological data from NASA Power Data, covering the period from January 1, 1989, to December 2021, was utilized to generate Intensity-Duration-Frequency (IDF) curves for Azzaba, situated in the Skikda province.



a)

b)

Figure 4.4: Annual Climate Data: a) Average Temperature, b) Average Precipitation



a) **b)**

Figure 4.5: Climate Data by Month: a) Temperature Averages, b) Precipitation Averages

The average annual precipitation in the region varies between a maximum of 727 mm/year and a minimum of 26 mm/year, with an annual temperature range from a maximum of 20°C to a minimum of 18°C, as shown in Figure 4.4. The average monthly precipitation in 2021 ranged from a maximum of 107 mm to a minimum of 2 mm, with temperatures ranging from a maximum of 30°C to a minimum of 13°C, as shown in Figure 4.5.

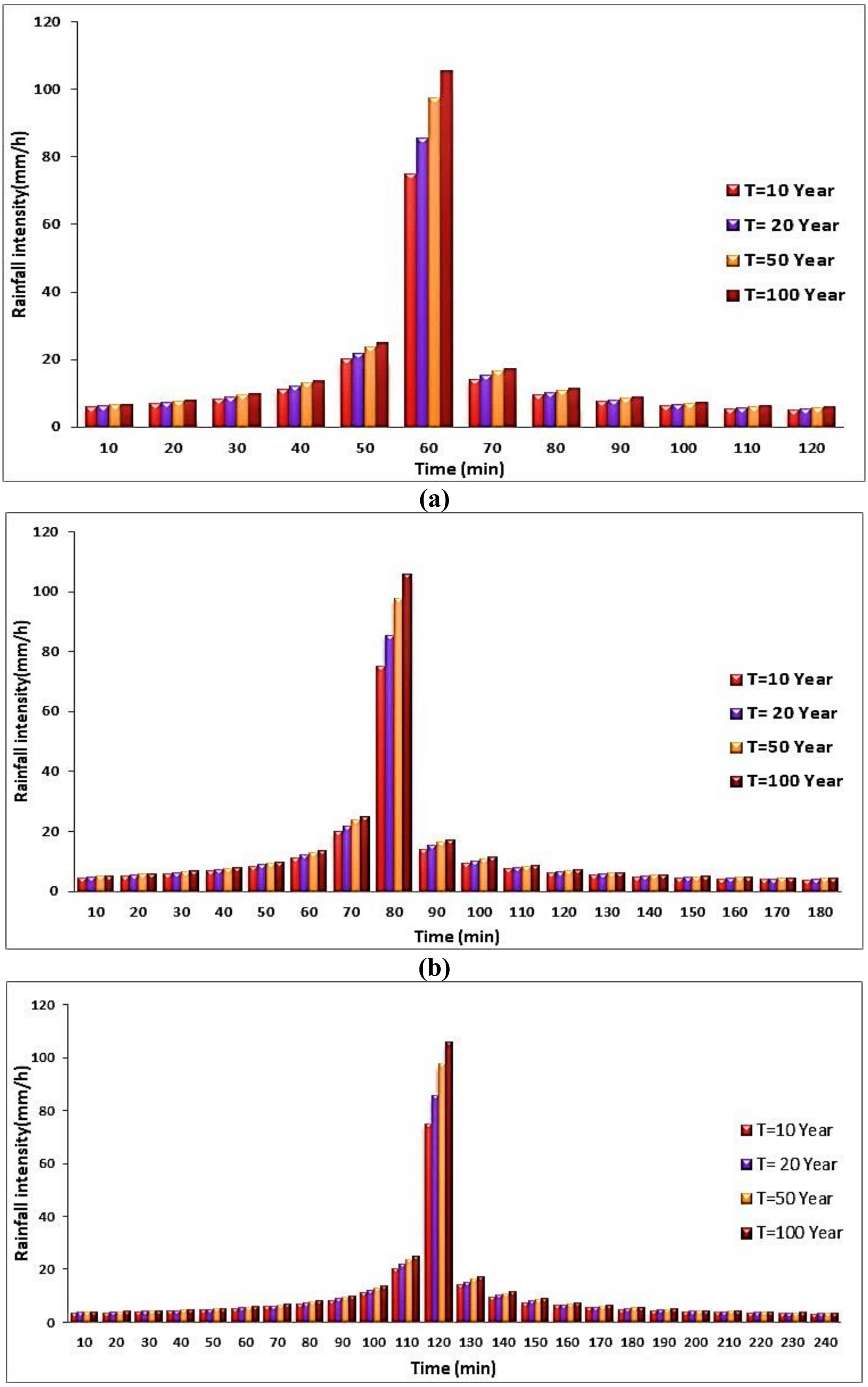
4.2.2.2 Hydrological Modeling and Design Rainfall

Hydrological modeling frequently requires the determination of design rainfall or rainfall hyetographs. Design rainfall is essential as it establishes a relationship between rainfall intensities and their likelihood of occurrence over a specific duration. In this research, design rainfall was derived using Intensity-Duration-Frequency (IDF) curves, which provide a statistical basis for understanding the frequency of various rainfall intensities over different durations. This approach is crucial for accurately predicting stormwater runoff and managing flood risks effectively (Maidment, 1993).

✓ **Composite Hyetograph:**

In this study, a composite hyetograph was employed to evaluate the effectiveness of the stormwater network by examining how the duration of rainfall affects hydraulic modeling at the outlet. This approach is deemed more precise than a uniform distribution because it derives maximum intensities from Intensity-Duration-Frequency (IDF) curves (Laouacheria et al., 2019). The rainfall plot (Figure 4.6) represents return periods of 10, 20, 50, and 100 years, with durations ranging from 2 to 4 hours. This refined methodology enhances the understanding of hydrological behavior and supports more accurate simulations, ultimately leading to improved stormwater management strategies (Vu et al., 2018)

.



(a)

(b)

(c)

Figure 4.6: Rainfall Hyetograph Composites for 2-Hour (a), 3-Hour (b), and 4-Hour Durations

4.3. MIKE+ Modeling Approach

MIKE+ is an advanced water decision-support software used for modeling various water systems, including rivers, collection systems, and overland flows. The software supports

SWMM5 for collection systems and water distribution, and provides a comprehensive Model Type Editor to manage and customize features and modules. The key features activated for SWMM5 collection system models include:

- **Collection System Network**: Models hydrodynamics of urban storm drainage and wastewater collection networks.
- **2D Overland Flow:** Simulates surface flooding resulting from surcharging pipe networks.

The MIKE+ model demonstrates considerable effectiveness in flood forecasting, especially when integrated with machine learning techniques (Dang et al., 2024). n this research, MIKE+ was employed to analyze urban flooding and to simulate the depth of inundation caused by manhole overflows within the stormwater network of Bir Farina, Azzaba City.

4.3.1 Selection of Models for Urban Stormwater Overflow Prediction

The SWMM5 Collection System model and the 2D Overland Flow model within MIKE+ were selected for their ability to handle complex interactions between stormwater systems and surface flows (DHI, 2023). These models are adept at simulating both hydrological and hydraulic processes, essential for understanding and predicting urban flood behavior.

a) Hydrologic and Hydraulic Modeling

Hydrologic models study the movement and storage of water within the hydrological cycle and focus on rainfall-runoff processes. For this study, the SCS-CN (Soil Conservation Service Curve Number) method was used. This method estimates runoff based on the water balance equation (Eslamian et al., 2024)

, with the primary formula:

$$\mathbf{Q = \frac{(P - Ia)^2}{P - Ia + S} \qquad (4.1)}$$

Where,
Q = actual runoff in mm
P = Average Rainfall in mm
Ia = Initial abstraction representing all the losses before the runoff begins and is given by the empirical equation.
S = Potential infiltration after the runoff given by following equation (Sifa et al., 2023)

$$\mathbf{S = \frac{25400}{CN} - 254 \qquad (4.2)}$$

After substituting $\mathrm{Ia} = 0.2\mathrm{S}$ into the primary equation:

$$\mathbf{Q = \frac{(P - 0.2S)^2}{P - 0.8S} \qquad (4.3)}$$

For, P > Ia (0.2S) which is Otherwise Q = 0CN represents the runoff potential of land cover-soil complex characteristics, impacted by antecedent moisture conditions (AMC) of soil and land use management.

Hydraulic models calculate water flow and pressures in pipes and channels. They are critical for the design and analysis of sewer systems. In this study, the dynamic wave model was employed, which uses the full Saint-Venant equations to account for inertial, pressure, gravity, and friction forces. The model consists of:

✓ **Continuity Equation:**

$$\mathbf{\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0 \qquad (4.4)}$$

Where:

- A = cross-sectional area of flow (m^2)

- Q = flow rate (m^3/s)

- x = longitudinal distance along the channel (m)

- t = time (s)

- ✓ **Momentum Equation:**

The momentum equation, which describes the conservation of momentum, includes the effects of inertia, pressure, gravity, and friction (Thi, 2008)

. It is expressed as:

$$\frac{\partial A}{\partial t} + \frac{\partial}{\partial x}(\frac{Q^2}{A} + gA(S_0 + S_f) = 0 \qquad (4.5)$$

Where:

- g = acceleration due to gravity (m/s^2)

- S_0= slope of the channel bed

- S_f= slope of the energy grade line

Together, the continuity and momentum equations form the Saint-Venant equations, which describe unsteady open channel flow. They are essential for modeling flood waves, dam breaks, tidal bores, and other rapidly varying flows (Mujumdar 2001).

b) Calibration

The MIKE+ model was calibrated using data on water depth from the storm event of February 4, 2019. Calibration was performed based on the coefficient of determination (R^2) to ensure that the model predictions accurately represented observed data, providing reliable insights into the performance of the stormwater network under extreme weather conditions.

4.4 Machine Learning Approaches

In our research for predicting overflow rate averages simulated by MIKE+ models, we used ensemble learning models, specifically Random Forest and Gradient Boosting, as well as classical machine learning models, including Linear Regression and Neural Networks.

4.4.1 Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) are machine learning models inspired by the human brain, designed to recognize patterns and make predictions from complex data (Guresen & Kayakutlu, 2011; Tian et al., 2020; Yang & Yang, 2014). An ANN consists of interconnected artificial neurons that process data and generate outputs (Walczak & Cerpa, 2003; Tian et al., 2020). Key components of an ANN include:

- **Input Layer**: Collects raw data or features to be processed by the network (Tian et al., 2020).

- **Hidden Layers**: Perform data transformations and capture complex, non-linear patterns by applying activation functions to the weighted sum of inputs (Walczak & Cerpa, 2003; Tian et al., 2020).
- **Output Layer**: Produces the final output based on the network's calculations (Batina et al., 2019).

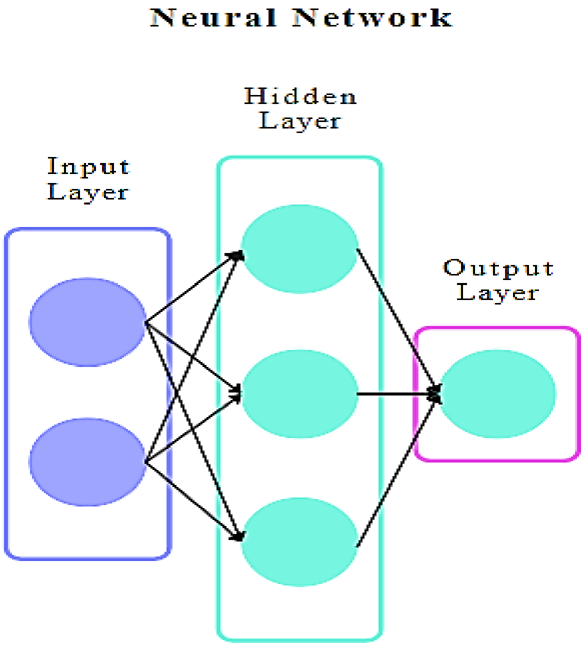


Figure 4.7: Neural network diagram

4.4.2 Linear Regression (LR)

Linear regression is a supervised machine learning algorithm that establishes a linear connection between a dependent variable and one or more independent variables by fitting a linear equation to the data observed. This technique helps in predicting outcomes based on the relationship defined by the fitted equation, making it a fundamental method in predictive modeling and statistical analysis **(**GeeksforGeeks,2024**)**.

Simple Linear Regression refers to a scenario where there is a single independent variable. This represents the most basic form of linear regression, which involves only one dependent variable and one independent variable.

Multiple Linear Regression involves the use of two or more independent variables to predict the outcome of one dependent variable. This method allows for more complexity as it can account for multiple factors influencing the dependent variable **(**GeeksforGeeks,2024**)**.

4.4.3 Random Forests

Random Forests (RF), introduced by Breiman (2001) , are popular ensemble learning algorithms that use multiple decision trees as their base learners. Each tree in a Random Forest model is trained on a different bootstrap sample, meaning that samples are randomly selected with replacement from the main dataset. This method enhances the diversity among the trees by introducing randomness in the selection of features used at each split, allowing each tree to be trained on a unique subset of the data.

Random Forests can handle both classification tasks, where they predict categories, and regression tasks, where they predict continuous values. By aggregating the outputs from multiple trees, Random Forests reduce the risk of overfitting, a common issue with single decision trees. This approach generates a more stable and generalizable model as it averages or votes on individual predictions. The RF algorithm proceeds through several key steps: first, it draws random bootstrap samples from the dataset. For each sample, a decision tree is built, and predictions are generated. Finally, the model aggregates these predictions through majority voting (for classification) or averaging (for regression), producing the final result based on the most popular outcome. This ensemble technique is thus recognized for improving predictive accuracy and robustness (Breiman, 2001; Cutler et al., 2012; Salim & Bhattacharyya, 2023).

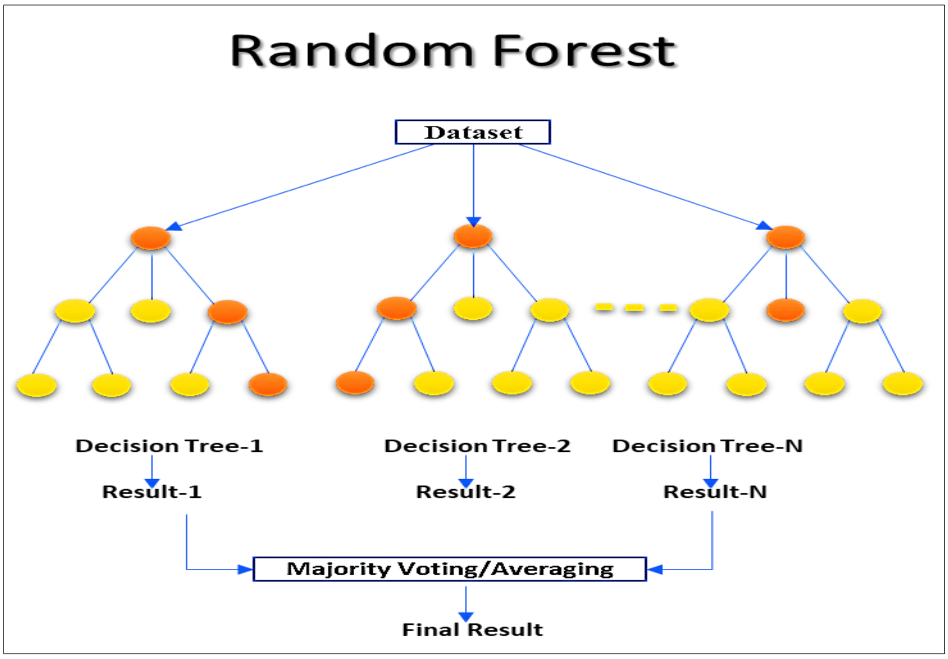


Figure 4.8: Random Forest algorithms

4.4.4 Gradient Boosting

Gradient boosting is a machine learning technique introduced by Leo Breiman and Jerome H. Friedman (Friedman, 1999) .It is a method for improving the accuracy of predictions made by models. The core idea behind gradient boosting is to iteratively add new models (typically decision trees) to correct the errors made by the previous model(s) in the sequence. This ensemble learning approach enables gradient boosting to deliver robust predictive performance across a wide range of tasks. The key components of gradient boosting algorithms include weak learners (typically decision trees), a loss function that guides the learning process, and an additive model structure. The input to a gradient boosting algorithm is a labeled dataset, and the output is a predictive model that can be used to make inferences on new, unseen data (Friedman, 2001).

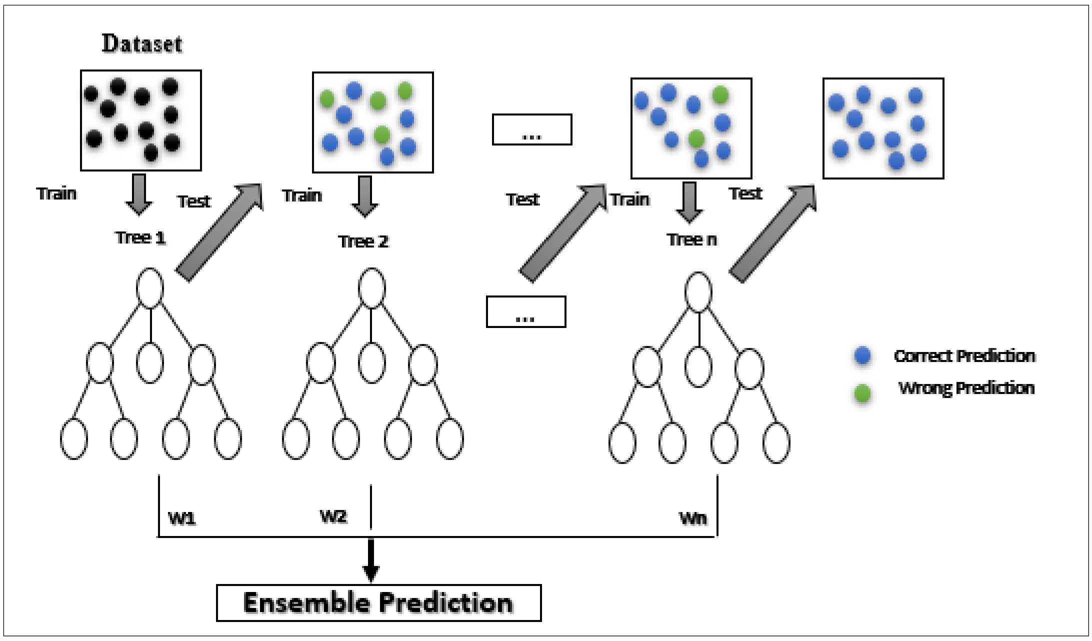
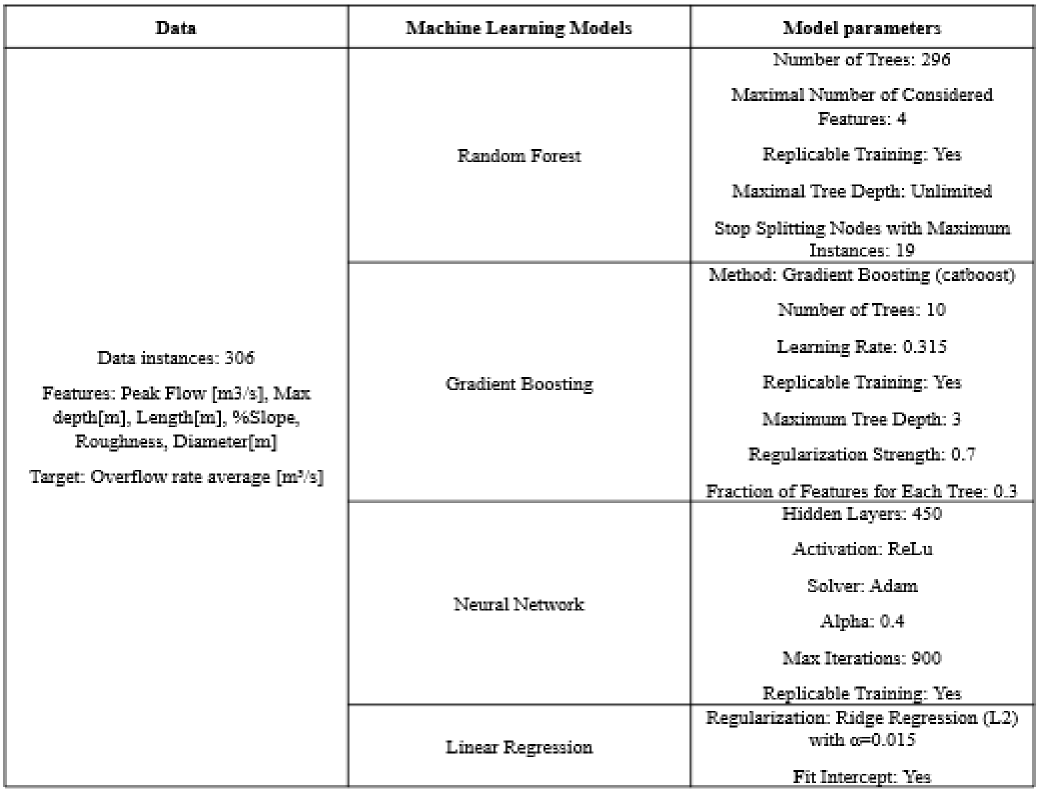


Figure 4.9: Gradient Boosting Model Workflow

4.5 Hyperparameters of Applied Machine Learning Models

Table 4.1 provides the hyperparameters for the machine learning models used in this study, including ensemble methods and traditional models like Linear Regression and Neural Networks. Each model, configured with specific parameters, is applied to a dataset with 306 records and feature variables such as Peak Flow, Max Depth, Length, Slope percentage, Roughness, and Diameter, targeting the average Overflow Rate. These configurations, optimized through iterative testing, significantly impact the predictive performance of each model, showcasing their efficiency in the current analytical environment.

Table 4.1: Hyperparameters of Applied Machine Learning Models



Data	Machine Learning Models	Model parameters
Data instances: 306 Features: Peak Flow [m3/s], Max depth[m], Length[m], %Slope, Roughness, Diameter[m] Target: Overflow rate average [m³/s]	Random Forest	Number of Trees: 296 Maximal Number of Considered Features: 4 Replicable Training: Yes Maximal Tree Depth: Unlimited Stop Splitting Nodes with Maximum Instances: 19
	Gradient Boosting	Method: Gradient Boosting (catboost) Number of Trees: 10 Learning Rate: 0.315 Replicable Training: Yes Maximum Tree Depth: 3 Regularization Strength: 0.7 Fraction of Features for Each Tree: 0.3
	Neural Network	Hidden Layers: 450 Activation: ReLu Solver: Adam Alpha: 0.4 Max Iterations: 900 Replicable Training: Yes
	Linear Regression	Regularization: Ridge Regression (L2) with α=0.015 Fit Intercept: Yes

4.6 Performance Metrics

In the realm of statistical analysis and machine learning, evaluating predictive models is essential for understanding their performance and reliability. Various metrics are employed to assess how well a model predicts outcomes based on given input data. Understanding these metrics enables researchers and practitioners to make informed decisions when selecting and fine-tuning their models. Among the most commonly used evaluation criteria are Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R-squared (R^2), and the Pearson Correlation Coefficient (r). Each of these metrics provides unique insights into different aspects of model accuracy and predictive power. Mean Squared Error (MSE) measures model accuracy by calculating the average of the squares of the differences between actual and predicted values. The formula for MSE is given by (Lu et al., 2023):

$$\mathbf{MSE} = \frac{\sum(\mathbf{Y_i} - \bar{\mathbf{Y}})^2}{\mathbf{N}} \qquad (\mathbf{4.8})$$

Lower MSE values indicate better model performance. Mean Absolute Error (MAE) offers a similar assessment by averaging the absolute differences between actual and predicted values, providing insight into overall prediction error without regard to direction:

$$\mathbf{MAE} = \frac{\sum(\mathbf{Y_i} - \bar{\mathbf{Y}})}{\mathbf{N}} \qquad (\mathbf{4.9})$$

Root Mean Squared Error (RMSE) refines this by taking the square root of MSE, allowing for interpretation in the same units as the original data and representing the standard deviation of prediction errors:

$$\mathbf{RMSE} = \sqrt{\frac{\sum(\mathbf{Y_i} - \bar{\mathbf{Y}})^2}{\mathbf{N}}} \qquad (\mathbf{4.10})$$

R-squared (R^2) measures the proportion of variance in the dependent variable explained by the independent variables in a regression model. It indicates the model's goodness of fit, with values ranging from 0 to 1; higher values suggest better explanatory power (Chicco et al., 2021):

$$\mathbf{R^2} = \mathbf{1} - \frac{\sum(\mathbf{X_i} - \bar{\mathbf{Y}})^2}{\sum(\bar{\mathbf{Y}} - \boldsymbol{Y_i})^2} \qquad (\mathbf{4.11})$$

Finally, the Pearson Correlation Coefficient (r) quantifies the linear relationship between two variables, with values ranging from -1 to +1, where +1 indicates perfect positive correlation, -1 indicates perfect negative correlation, and 0 signifies no correlation (Mukaka, 2012):

$$\mathbf{r} = \frac{\sum((\mathbf{X_I} - \bar{\mathbf{X}}) \times (\mathbf{Y_I} - \bar{\mathbf{Y}}))}{\sqrt{\sum_i(\mathbf{X_i} - \bar{\mathbf{X}})^2 \times \sum_i(\mathbf{Y_i} - \bar{\mathbf{Y}})^2}} \ (\mathbf{4.12})$$

Where:

N: Total number of observations.

Yi: Observed value.

$\bar{Y}$: Predicted value.

X_I and Y_I :Individual data points.

$\bar{X}$ and $\bar{Y}$: means of X_I and means of Y_I respectively

Together, these metrics provide a comprehensive framework for evaluating model performance and understanding relationships between variables, ultimately leading to improved decision-making and more robust predictions.

4.7 Method for Extracting Feature Importance (MSE Method)

Feature importance is a method used in machine learning to assign numerical values to input features based on their effectiveness in predicting a target variable, as shown in the Figure 4.10. These values establish a hierarchy, enabling you to determine which features have the greatest impact on your model's predictions (Kolena, 2024). Various techniques are used to calculate these importance scores, including decision trees, linear models, neural networks (Azaria, 2022), and ensemble methods such as Random Forest and Gradient Boosting (Huang et al., 2024). Additionally, methods like Mean Squared Error (MSE) are employed to evaluate feature importance (Shin, 2023). These scores help rank features based on their contribution to the final prediction and are valuable for:

- Feature Selection: Identifying and selecting the most relevant features to reduce dimensionality, noise, and enhance model interpretability.
- Model Interpretability: Understanding which features influence the model's predictions and gaining insights into data relationships.
- Model Debugging: Identifying features that may be causing issues and affecting model performance.
- Improving Model Performance: Removing less important features to reduce overfitting and training time (Kolena, 2024).

The Mean Squared Error (MSE) method for extracting feature importance is widely used in the context of ensemble learning models, particularly when dealing with regression tasks (Azaria, 2022). Here's a structured explanation of the methodology used to extract feature importance from ensemble learning models using the MSE method.

a) **Model Training:** Train the ensemble learning models (Random Forest and Gradient Boosting) on the dataset containing six input features. This process involves fitting the models to the training data to learn the relationships between the features and the target variable (Islam & Rony, 2024).
b) **Model Evaluation:** Evaluate the model's performance on a validation dataset to calculate the baseline Mean Squared Error (MSE). This value serves as a reference for assessing the impact of individual features.
c) **Feature Permutation:** For each of the six features, permute (shuffle) its values across all instances in the validation set. This randomization breaks the relationship between the feature and the target variable while keeping other features intact.
d) **Recalculate MSE:** Use the trained model to make predictions on the modified dataset (with the permuted feature) and compute the new MSE.
e) **Importance Score Calculation:** Calculate the importance score for each feature based on the increase in MSE due to the permutation. The formula is as follows:

$$\mathbf{Importance(Xi) = MSEpermuted - MSE\ baseline}$$

A higher increase in MSE indicates greater importance of the feature (Shin, 2023).

f) **Statistical Analysis:** Repeat the permutation process for each feature multiple times (in this case, 10 permutations) to obtain a mean and standard deviation as of the importance scores for each feature. This helps quantify the reliability of the importance estimates.

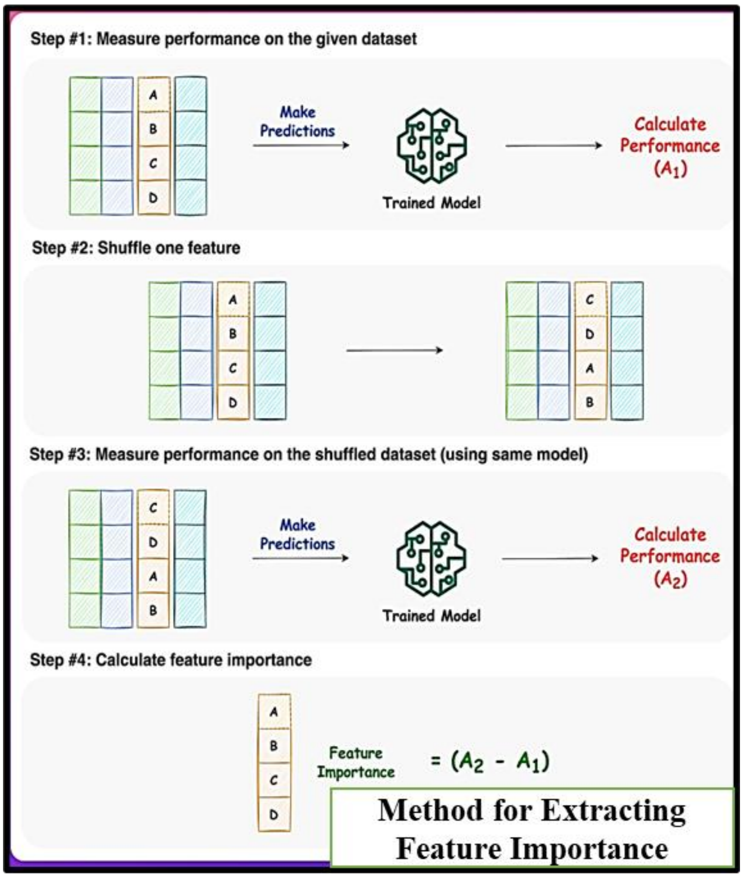


Figure 4.10: Method for Extracting Feature Importance

In this work, we extract feature importance using the MSE method for a 2-hour rainfall duration across return periods, as shown in Figures 5.15 to 5.18. We also provide the mean and standard deviation for the top three features, as shown in Figures 5.19 to 5.21.

4.8 Proposed Methods

In this study, we introduce a novel methodology called SWN-ML (**S**tormwater **W**ater **N**etwork-**M**achine **L**earning), which is illustrated in the flowchart depicted in Figure 4.11.

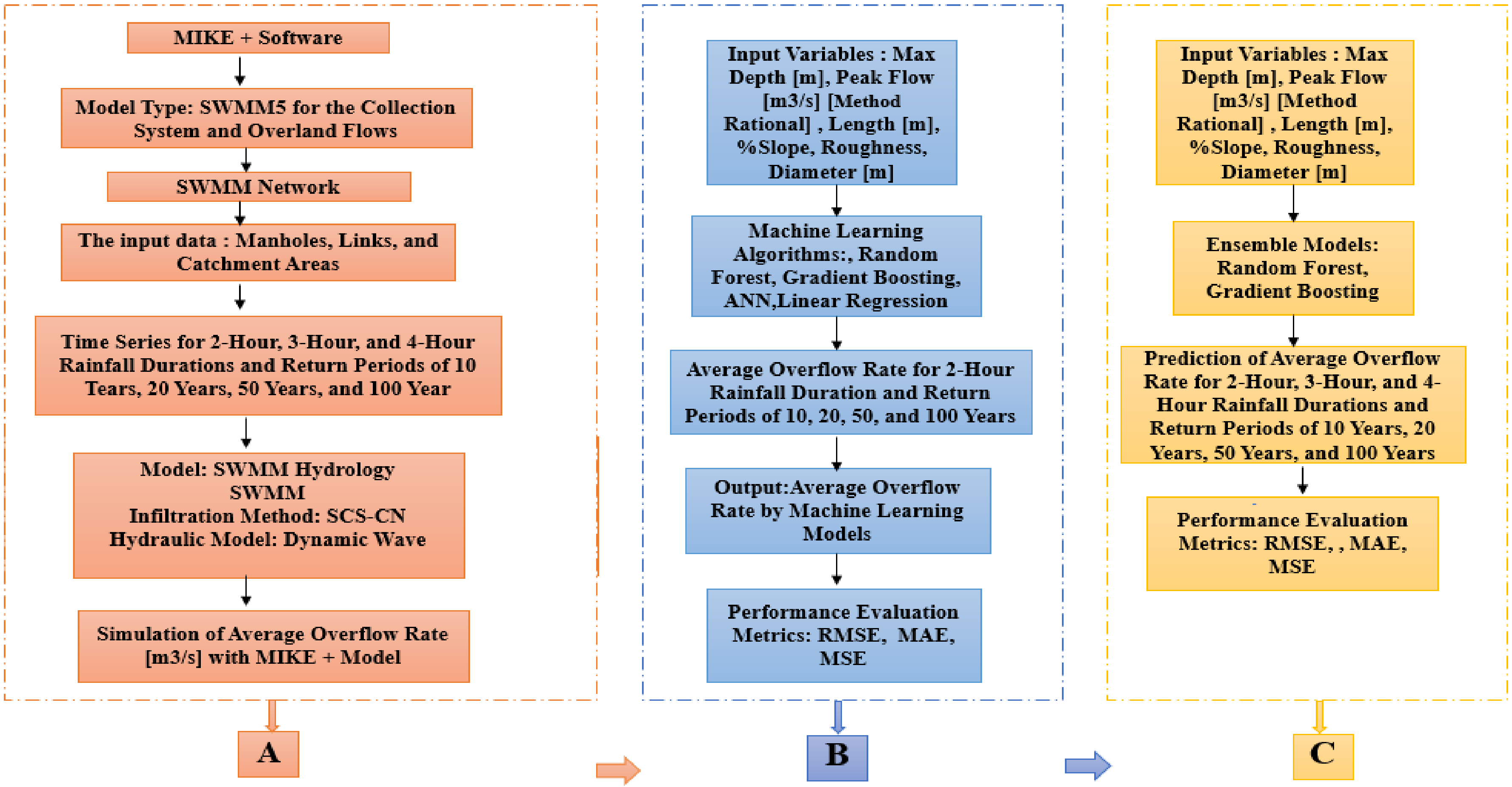


Figure 4.11: Flowchart of the SWN-ML Approach

Sections A, B, C, and D outline the different phases or components of a study or research methodology that incorporates classical modeling along with ensemble method-based machine learning. Below is a brief summary of each section as illustrated in Figure 4.11.

4.8.1 Section A: Simulation Using MIKE+ Software

This section outlines a simulation conducted using MIKE+ software, which incorporates the SWMM 5 model for managing collection systems and overland flows. The SWMM network included 306 manholes, 306 links, 97 subcatchments, and one outlet, mirroring the network design utilized in the study by Laouacheria et al. (2019). Calibration of the model was performed using only the observed data for water depth at the outlet during the storm event on February 4, 2019. The effectiveness of the calibration was evaluated using the Coefficient of Determination (R^2), as illustrated in Figure 5.1. The simulation integrates time series rainfall data with durations of 2, 3, and 4 hours, represented as a composite hyetograph. For each rainfall duration, return periods of 10, 20, 50, and 100 years were considered. The SWMM hydrological model, combined with the SCS CN infiltration method and the dynamic wave hydraulic model, was utilized to simulate the average overflow rate using the MIKE+ software. These simulations account for different rainfall durations and return periods.

4.8.2 Section B: Investigation of Machine Learning Algorithms

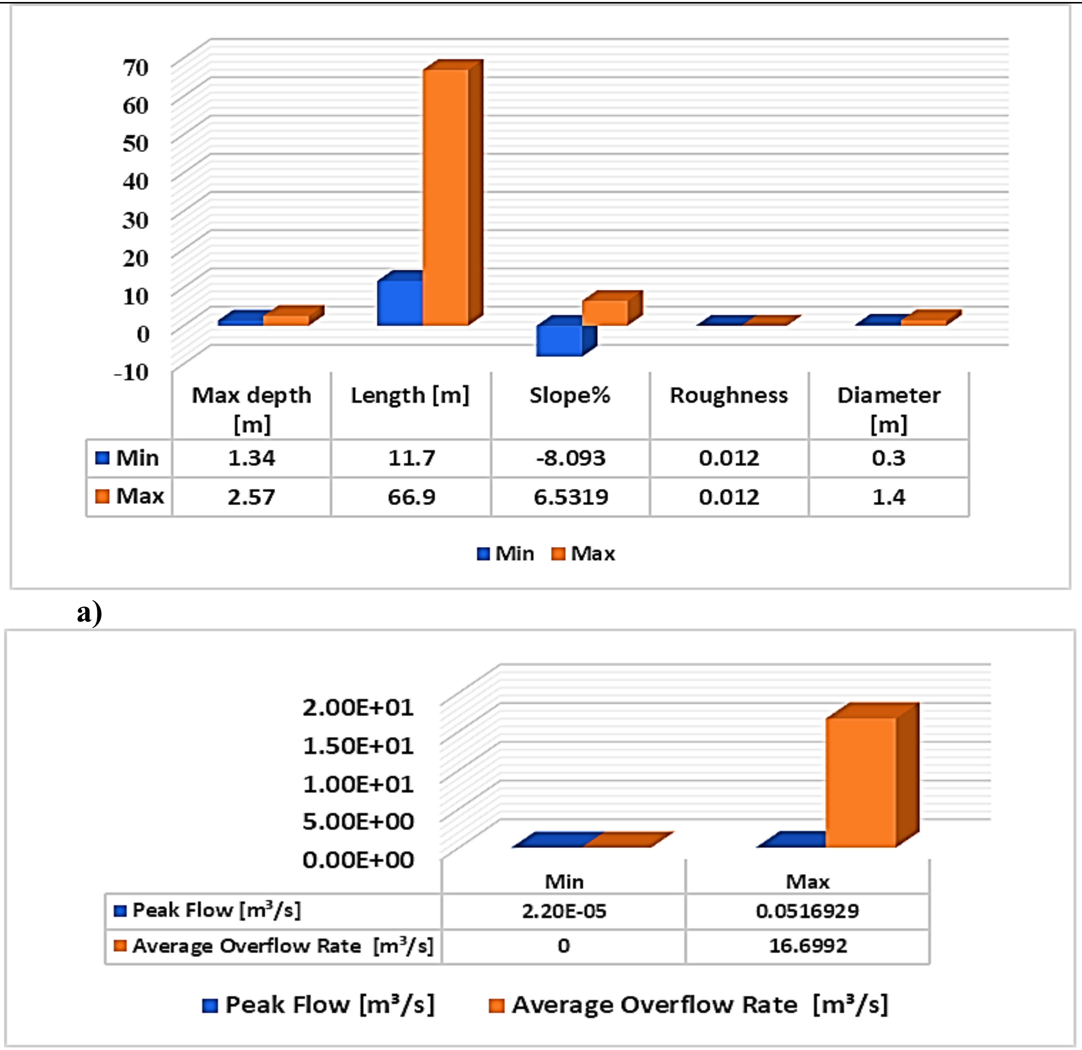
This section primarily aims to explore various machine learning methods, including ensemble learning models such as Random Forest and Gradient Boosting, alongside traditional techniques like Artificial Neural Networks (ANN) and Linear Regression. The objective is to accurately predict the average overflow rate.

a) Database Construction for Machine Learning Models

In this study, we established a comprehensive numerical database tailored for ensemble learning models applied to stormwater networks. This database encompasses data from 306 manholes and links, with six key features. The input variables include Peak Flow (calculated using the Rational Method) in m^3/s, Maximum Depth in meters, and link attributes such as Length (m), Slope (%), Roughness, and Diameter (m). The target variable, Average Overflow Rate, was derived from simulations using MIKE+ models.

b) Data Analysis and Insights:

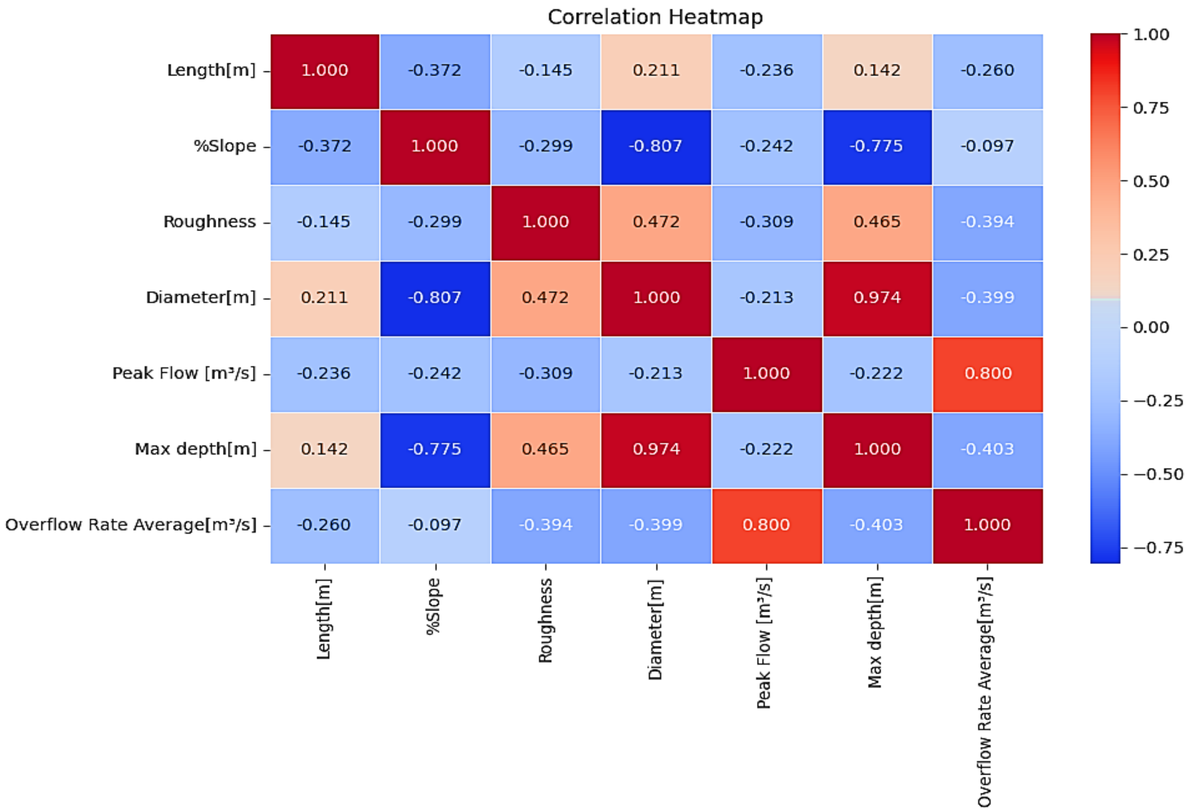
In the first phase, the data is examined to gain a comprehensive understanding of the various attributes of the stormwater network and the potential for overflow issues. The Minimum and Maximum Values Overview for the dataset, as shown in Figure 4.12, along with the Pearson Correlation Heatmap of hydrological parameters and average overflow rate, depicted in Figure 4.13, presents an analysis of key metrics such as peak flow, depth, slope, and diameter. These metrics are utilized to evaluate the complexities of the system and assess the factors contributing to overflow risks.



a)

b)

Figure 4.12: Minimum and Maximum Values Overview for Dataset



	Length[m]	%Slope	Roughness	Diameter[m]	Peak Flow [m³/s]	Max depth[m]	Overflow Rate Average[m³/s]
Length[m]	1.000	-0.372	-0.145	0.211	-0.236	0.142	-0.260
%Slope	-0.372	1.000	-0.299	-0.807	-0.242	-0.775	-0.097
Roughness	-0.145	-0.299	1.000	0.472	-0.309	0.465	-0.394
Diameter[m]	0.211	-0.807	0.472	1.000	-0.213	0.974	-0.399
Peak Flow [m³/s]	-0.236	-0.242	-0.309	-0.213	1.000	-0.222	0.800
Max depth[m]	0.142	-0.775	0.465	0.974	-0.222	1.000	-0.403
Overflow Rate Average[m³/s]	-0.260	-0.097	-0.394	-0.399	0.800	-0.403	1.000

Figure 4.13: Pearson Correlation Heatmap of Hydrological Parameters and Overflow Rate Average

The data presented in Figure 4.12 reveals the minimum and maximum values, which are essential for understanding the range, variability, and extremes of the dataset. This knowledge is crucial for analyses, such as assessing the risk of overflow in a stormwater network. These values provide a quick overview of the data's distribution and can be instrumental in identifying outliers or anomalies. Additionally, the Pearson Correlation Heatmap highlights key relationships, such as the strong positive correlation between diameter and peak flow, as well as the negative correlation between roughness and flow capacity. These correlations are valuable for identifying the main factors influencing overflow and can assist in optimizing the system.

The decreased peak flow rates in the study area are affected by various factors, such as the design and capacity of the stormwater network, changes in land use resulting from urbanization, and the implementation of stormwater management practices.(Diogo & do Carmo, 2019; Laouacheria & Mansouri, 2015). The approach to urbanization in the study area has incorporated green infrastructure, which plays a crucial role in alleviating the impacts of increased impervious surfaces. Parks, open spaces, and other green areas function as natural sponges, soaking up rainfall and diminishing the volume of runoff that flows into the stormwater system.(Laouacheria et al., 2019). Furthermore, the fluctuations in annual precipitation are a key factor in influencing peak flow rates. A decrease in the intensity of rainfall or a reduction in the occurrence of heavy storms in the area would naturally lead to lower peak flows. Additionally, the duration of storms is important; shorter and less intense storms produce less runoff, resulting in smaller peak flows. The wide range of simulated overflow rates highlights the necessity of effective modeling techniques for accurately predicting and managing stormwater overflow events.

c) Data Normalization and Preprocessing:

At this stage, the dataset is scaled to a range of [0-1] in order to provide equal representation of variables. Normalization is a technique that prevents certain variables from having an excessive effect on a model because of their bigger scale. It achieves this by taking into account the wide range of values that these variables can have. In the provided dataset, the variable 'Peak Flow [m^3/s]' has a range of values from 2.20E-05 to 0.0516929, whereas the variable 'Length[m]' has a range of values from 11.7 to 66.9. This technique provides that each variable contributes proportionally to the ultimate forecast, taking into account the underlying patterns in the data.

d) Model Training and Validation:

The study used data from 2-hour rainfall durations across different return periods. In machine learning, cross-validation is a method to evaluate a model’s performance on unseen data by dividing the dataset into several parts, or folds. For this study, five-fold cross-validation was applied, where each fold took turns serving as the validation set while the model was trained on the remaining folds. This process was repeated, with each fold acting as the validation set once, and the results were averaged to yield a more reliable measure of the model’s performance.(GeeksforGeeks, 2023).

The main goal of cross-validation is to prevent overfitting, which occurs when a model is overly tailored to the training data, resulting in poor performance on new, unseen data. By evaluating the model across multiple validation sets, cross-validation provides a more precise assessment of the model’s ability to generalize, ensuring it performs effectively on new data. (GeeksforGeeks, 2023).

e) Model Evaluation and Selection:

In this study, we used MAE, MSE, and RMSE to evaluate performance. The best-performing algorithms were selected to predict overflow rates for different rainfall durations and return periods.

4.8.3 Section C: Predictive Modeling Using Ensemble Learning Models

In this section, we forecast outcomes for varying rainfall durations (2, 3, and 4 hours) and return periods (10, 20, 50, and 100 years) using the top-performing model identified in Section B, which includes ensemble techniques like GB and RF. The models' effectiveness is evaluated using standard metrics: RMSE, MAE, MSE, and r.

4.9 Conclusion

In conclusion, the study of Bir Farina's stormwater network and its characteristics has led to the development of a comprehensive database and the application of machine learning algorithms. By utilizing ensemble learning models and traditional approaches, the research aims to accurately predict average overflow rates. The proposed SWN-ML methodology, which integrates MIKE+ software and machine learning techniques, offers a novel approach for simulating and analyzing stormwater network behavior. The study's findings contribute to improved flood risk management and urban resilience planning, providing valuable insights for future research and practical applications in the field of hydrological forecasting.

Chapter 5 : Results and Discussion

5.1 Introduction

Urban stormwater management has become increasingly critical due to rapid urbanization and climate change, which have intensified the frequency and severity of flooding events (Wang et al. 2022). Effective management of stormwater systems requires accurate predictions of overflow rates, particularly under extreme weather conditions. Traditional physical-based models, such as MIKE+, have been widely used to simulate hydrological and hydraulic processes within urban drainage networks. However, these models often face limitations due to their reliance on extensive observed data, which may not always be available, and their complexity in capturing the nonlinear interactions between various factors influencing stormwater overflow.

In recent years, ensemble learning approaches have emerged as powerful tools for predicting urban stormwater dynamics and offer a viable alternative to classical hydrological models (Zounemat-Kermani et al., 2021). Techniques such as Random Forest and Gradient Boosting have demonstrated their ability to handle complex datasets and provide accurate predictions even in data-scarce environments. The integration of ensemble learning models with traditional physical-based simulations offers a promising avenue for enhancing the accuracy and reliability of flood predictions, ultimately contributing to more effective flood risk management strategies.

An important aspect of improving model performance is understanding the relative importance of different features in influencing stormwater overflow predictions. This study employs the Mean Squared Error (MSE) method to assess feature importance within ensemble learning models. By evaluating the impact of various factors on the accuracy of predictions, this approach helps identify key variables that significantly affect overflow rates. The findings from this analysis provide valuable insights into the sensitivity of hydraulic systems to different rainfall durations and impact factors on stormwater networks.

This chapter presents an in-depth analysis of the experimental results obtained from both physical-based and ensemble learning models. The performance of MIKE+ simulations is evaluated based on their ability to replicate observed water depths, while the effectiveness of ensemble learning models is assessed through various metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). The findings underscore the importance of integrating ensemble learning approaches as an alternative to traditional hydrological models to improve the prediction and management of stormwater overflow in urban areas.

5.2 Calibration of Physical-Based Models

Due to the absence of observed flow data at the outlet, the only available data for water depth at the outlet, obtained from the rainfall event on February 4, 2019 (47.50 mm) at the Azzaba gauge station, was used to calibrate the model. This rainfall event is estimated to have a recurrence period of around 50 years. To evaluate the model's accuracy, the observed water depth was compared with the simulated water depth at the outlet point. The results demonstrated a strong correlation between the observed and simulated water depths, with an R^2 value of 0.904 (Figure 5.1).

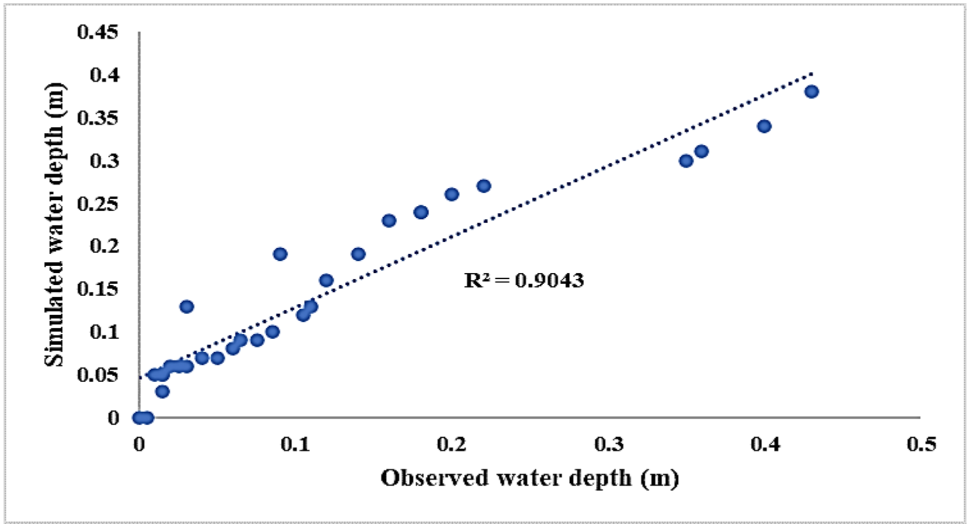


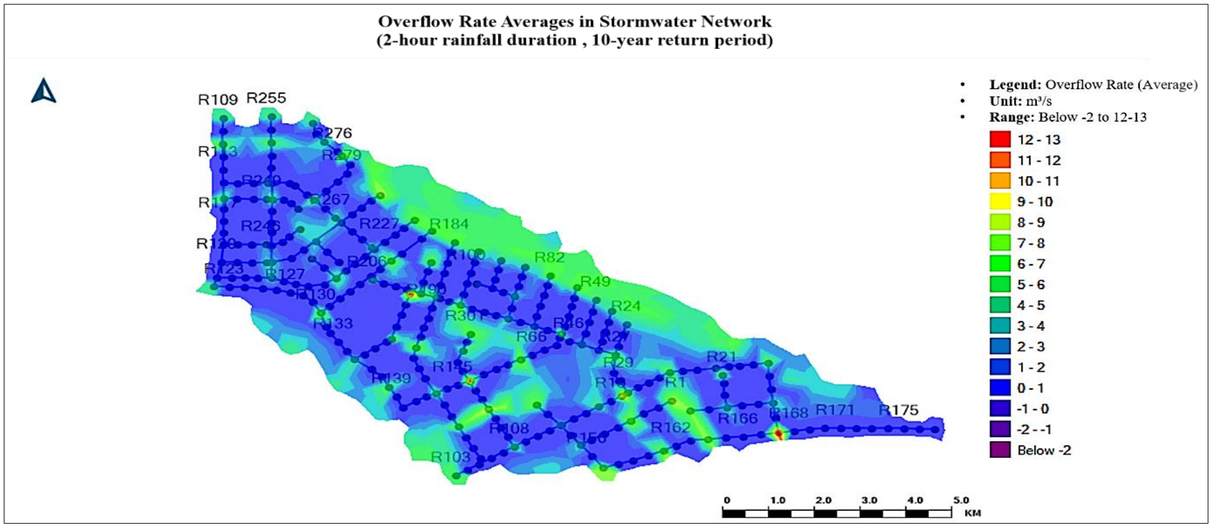
Figure 5.1: Scatter Plot Representation of Observed and Simulated Water Depth Measurements

5.3 MIKE+ Model Scenario Simulations

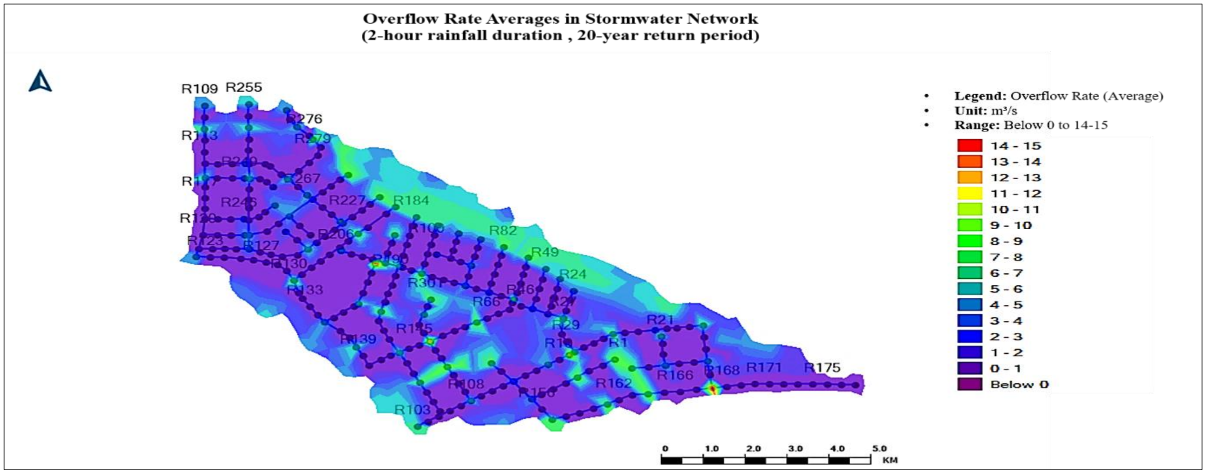
5.3.1 Scenario Simulation for a 2-Hour Rainfall Duration

This section focuses on the analysis and evaluation of the average overflow rates that were simulated by MIKE+ models throughout a 1-day simulation period. The overflow rate represents the volume of water exceeding the capacity of the stormwater network within a specified time period. This analysis covers rainfall durations of 2 hours, 3 hours, and 4 hours, as illustrated in Figures 5.2, 5.3, and 5.4. The results will inform the development of a new model using ensemble machine learning techniques to predict average overflow rates. Understanding the dynamics of capacity flow and detecting overflow spots within the stormwater network during flood events is essential. These valuable observations are crucial for improving flood risk management by identifying locations that are prone to flooding.

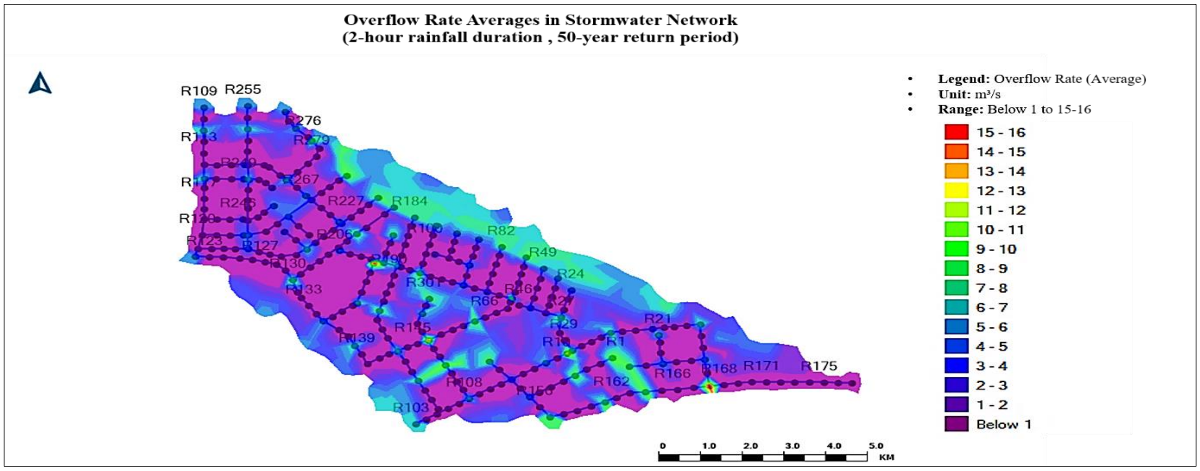
In this work, we have summarized the findings in Tables 5.1 to 5.3 for each rainfall duration across different return periods in all simulation scenarios, based on Figures 5.2 – 5.4. The average overflow rate represents the amount of water exceeding the network's capacity, which could potentially lead to flooding. An overflow rate of 0 m^3/s indicates no flooding. Overflow rates ranging from 0 to 10 m^3/s are considered moderate, while an average overflow rate of 10 m^3/s within the stormwater network indicates severe flooding.


Overflow Rate Averages in Stormwater Network
(2-hour rainfall duration , 10-year return period)
Legend: Overflow Rate (Average)
Unit: m³/s
Range: Below -2 to 12-13
12 - 13
11 - 12
10 - 11
9 - 10
8 - 9
7 - 8
6 - 7
5 - 6
4 - 5
3 - 4
2 - 3
1 - 2
0 - 1
-1 - 0
-2 - -1
Below -2
R109
R255
R276
R279
R267
R227
R184
R206
R127
R130
R133
R82
R49
R24
R301
R66
R46
R29
R21
R1
R139
R166
R168
R171
R175
R108
R162
R103
R150
0 1.0 2.0 3.0 4.0 5.0 KM

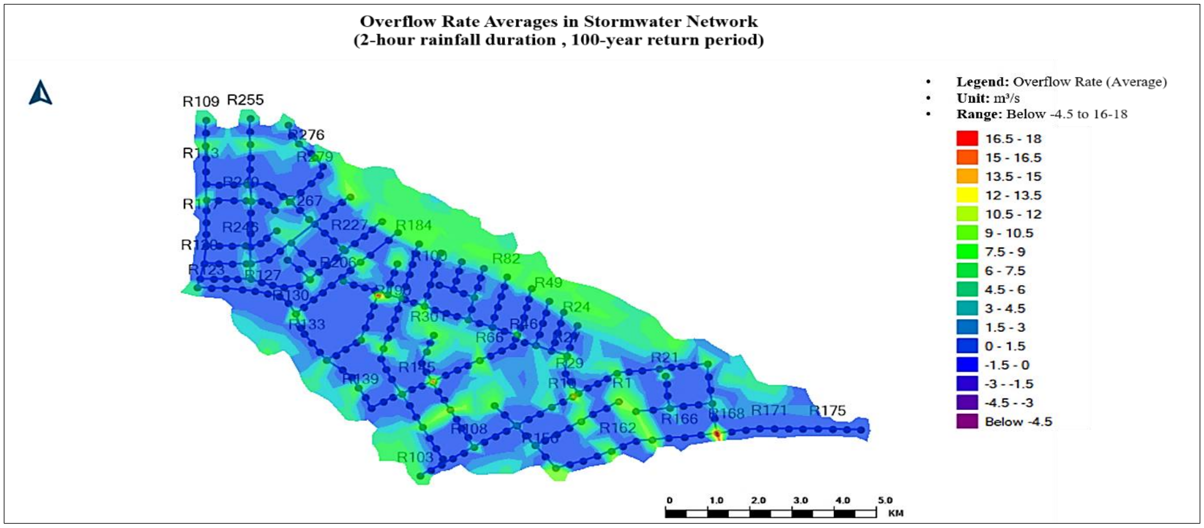
(a)


Overflow Rate Averages in Stormwater Network
(2-hour rainfall duration , 20-year return period)
Legend: Overflow Rate (Average)
Unit: m³/s
Range: Below 0 to 14-15
14 - 15
13 - 14
12 - 13
11 - 12
10 - 11
9 - 10
8 - 9
7 - 8
6 - 7
5 - 6
4 - 5
3 - 4
2 - 3
1 - 2
0 - 1
Below 0
R109
R255
R276
R279
R267
R227
R184
R206
R127
R130
R133
R82
R49
R24
R301
R66
R46
R29
R21
R1
R139
R166
R168
R171
R175
R108
R162
R103
R150
0 1.0 2.0 3.0 4.0 5.0 KM

(b)


Overflow Rate Averages in Stormwater Network
(2-hour rainfall duration , 50-year return period)
Legend: Overflow Rate (Average)
Unit: m³/s
Range: Below 1 to 15-16
15 - 16
14 - 15
13 - 14
12 - 13
11 - 12
10 - 11
9 - 10
8 - 9
7 - 8
6 - 7
5 - 6
4 - 5
3 - 4
2 - 3
1 - 2
Below 1
R109
R255
R276
R279
R267
R227
R184
R206
R127
R130
R133
R82
R49
R24
R301
R66
R46
R29
R21
R1
R139
R166
R168
R171
R175
R108
R162
R103
0 1.0 2.0 3.0 4.0 5.0 KM

(c)



(d)

Figure 5.2: MIKE+ Simulation of Average Overflow Rates (m³/s) in Stormwater Network for 2-Hour Rainfall Durations and Different Return Periods :(a) 10 Years, (b) 20 Years, (c) 50 Years, (d) 100 Years

Table 5.1: 2-Hour Rainfall Duration: Average Overflow Rates Across Multiple Return Periods

Category	Return Period (Year)	Color Representation	Nodes	Overflow Rate Range (m³/s)
Low Overflow Rates (≤ 1 m³/s)	**10**	Violet, Blue Dark	R274, R6, R81, etc.	Below -2 to 0-1
	20	Violet		Below 0 to 0-1
	50	Violet		Below 1-1-2
	100	Blue Dark, Violet		Below -4.5 to 0-1.5
Moderate Overflow Rates (1 m³/s - 10 m³/s)	**10**	Blue Dark, Green, Yellow	R130, R246, R247, R227, etc.	1-2 to 9-10
	20	Blue Dark, Green		1-2 to 9-10
	50	Blue Dark, Green		1-2 to 9-10
	100	Blue Dark, Green		0-1.5 to 9-10.5
Maximum Overflow Rates (> 10 m³/s)	**10**	Orange, Red, Green, Yellow	R45, R198, R71, R9	10-11 to 12-13
	20	Green, Yellow, Red, Orange	R9, R71, R198, R45	10-11 to 14-15
	50	Green, Yellow, Red, Orange	R285, R1, R62, R148, R9, R71, R198, R45	10-11 to 15-16
	100	Green, Yellow, Red, Orange	R164, R4, R285, R1, R62, R148, R9, R71, R198, R45	10.5-12 to 16.5-18

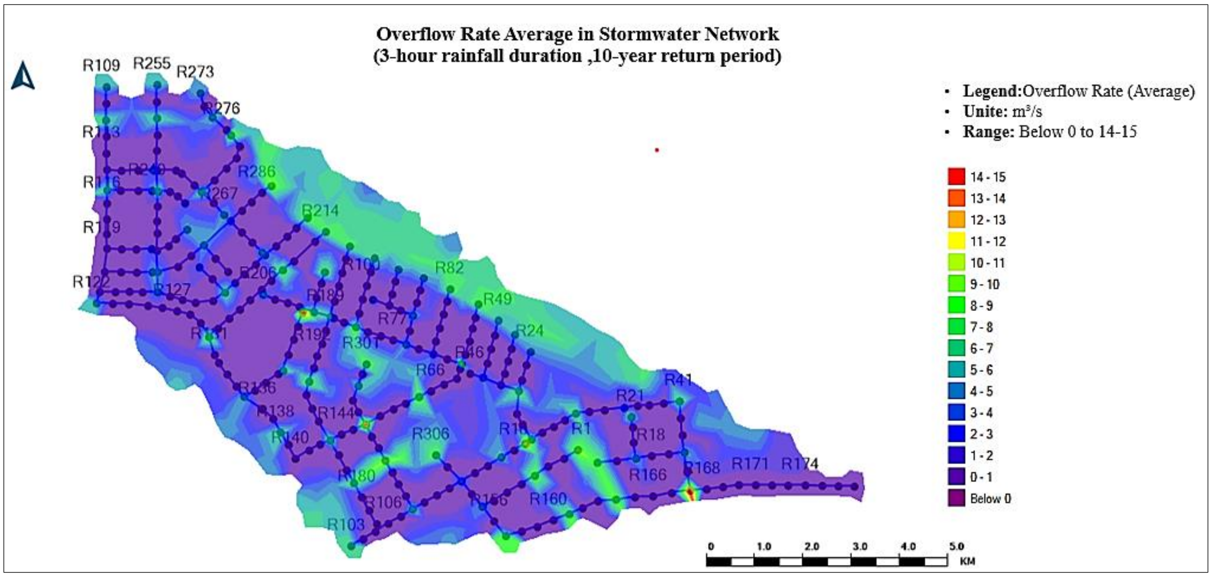
The Table 5.1 illustrates the scenario for a 2-hour rainfall duration across various return periods and the corresponding ranges of average overflow rates, as shown in the figure. Below are some observations based on the provided data:

- **Low Overflow Rates (≤ 1 m³/s):** For a 10-year return period, overflow rates range from below -2 to 0-1 m³/s, suggesting minimal overflow with violet and blue-dark colors. For longer return periods, overflow rates increase slightly, indicating a gradual rise in overflow risk.

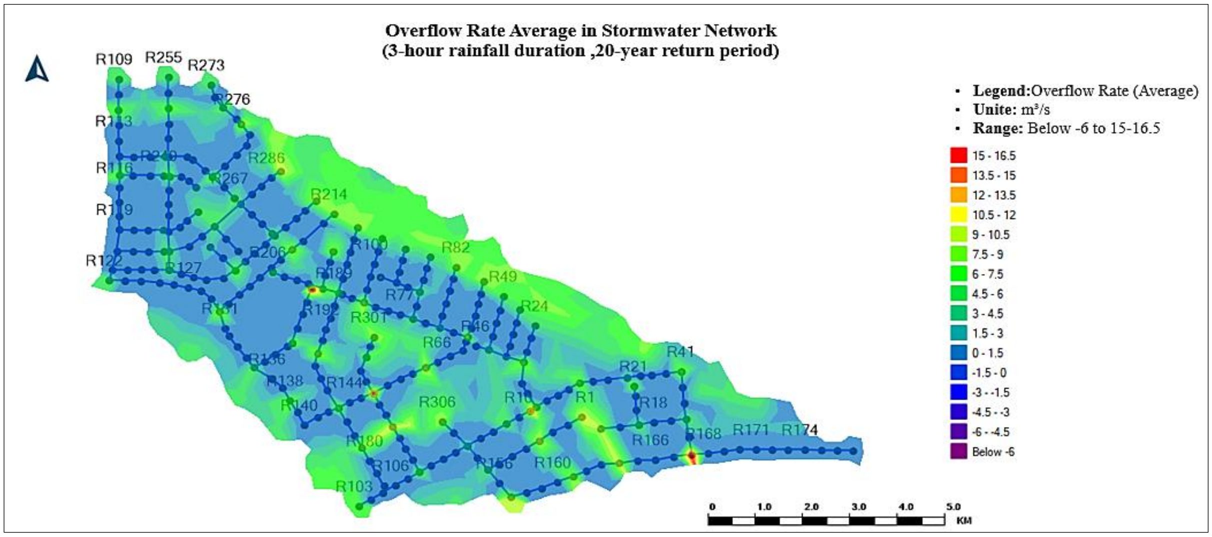
- **Moderate Overflow Rates (1 m³/s - 10 m³/s):** Rates range from 1-2 to 9-10 m³/s for the 10-year return period, remaining consistent across different return periods, with slight increases in extreme scenarios.
- **Maximum Overflow Rates (> 10 m³/s):** Overflow rates vary from 10-11 to 12-13 m³/s for a 10-year return period, increasing with longer return periods, indicating heightened overflow risk during extreme events.

5.3.2 Scenario Simulation for 3-Hour Rainfall Duration

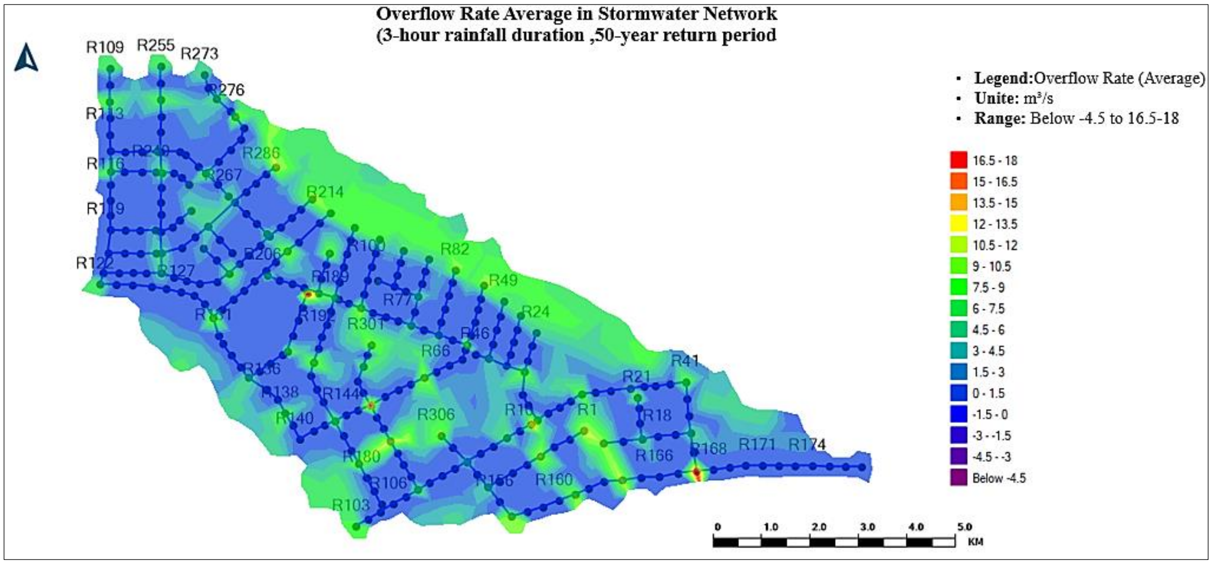
In this simulation, 3-hour rainfall durations with return periods of 10, 20, 50, and 100 years were considered (see Figures 5.3(a), 5.3(b), 5.3(c), and 5.3(d) respectively). We have also summarized the key findings in the table.



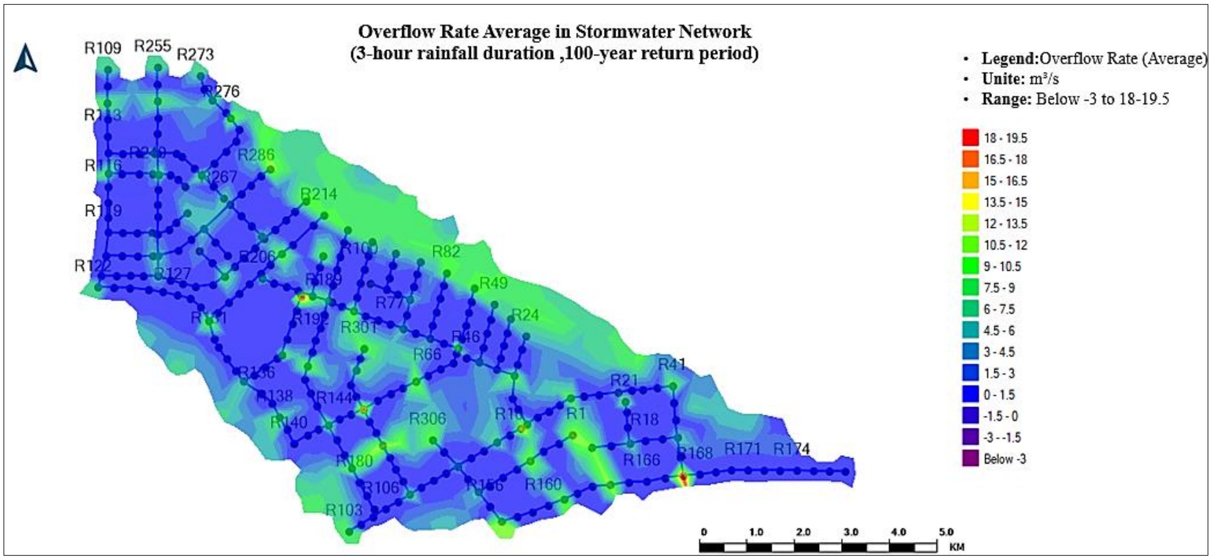
(a)



(b)



(c)



(d)

Figure 5.3: MIKE+ Simulation of Average Overflow Rates (m³/s) in Stormwater Network for 3-Hour Rainfall Durations and Different Return Periods : (a) 10 Years, (b) 20 Years, (c) 50 Years, (d) 100 Years

Table 5.2: 3-Hour Rainfall Duration: Average Overflow Rates Across Multiple Return Periods

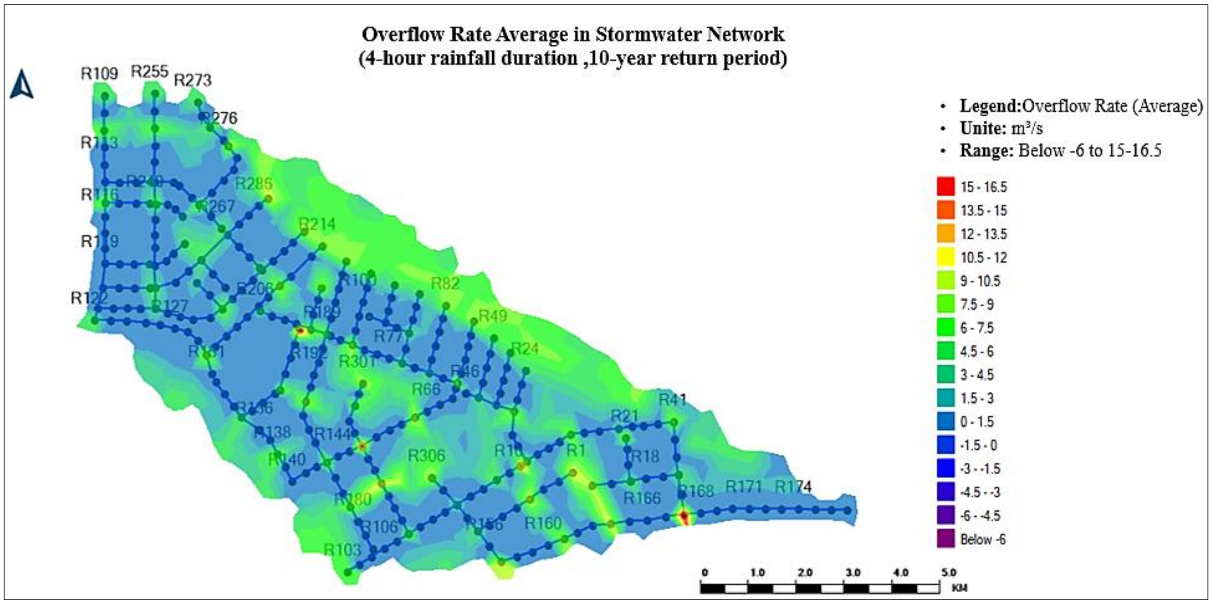
Category	**Return Period (Year)**	**Color Representation**	**Nodes**	**Overflow Rate Average Range (m³/s)**
Low Overflow Rates (≤ 1 m³/s)	**10**	Violet	R274, R6, R81, etc.	Below 0 to 0-1
	20	Blue Dark, Violet		Below -4 to 0-1.5
	50	Blue Dark, Violet		Below -4 to 0-1.5
	100	Blue Dark, Violet		Below -3 to 0-1.5
	10	Blue Dark, Green	R130, R246, R247, R227, etc.	1-2 to 9-10
	20	Blue Dark, Green		0-1.5 to 7.5-9

Moderate Overflow Rates (1 m³/s - 10 m³/s)	**50**	Blue Dark, Green		1.5-3 to 9-10.5
	100	Blue Dark, Green		0-1.5 to 9-10.5
Maximum Overflow Rates (> 10 m³/s)	**10**	Orange, Red, Green, Yellow	R45, R198, R71, R9	10-11 to 14-15
	20	Green, Yellow, Red, Orange	R9, R71, R198, R45	9-10.5 to 15-16.5
	50	Yellow, Red, Orange	R285, R1, R62, R148, R9, R71, R198, R45	10.5-12 to 16.5-15
	100	Green, Yellow, Red, Orange	R164, R4, R285, R1, R62, R148, R9, R71, R198, R45	10.5-12 to 18-19.5

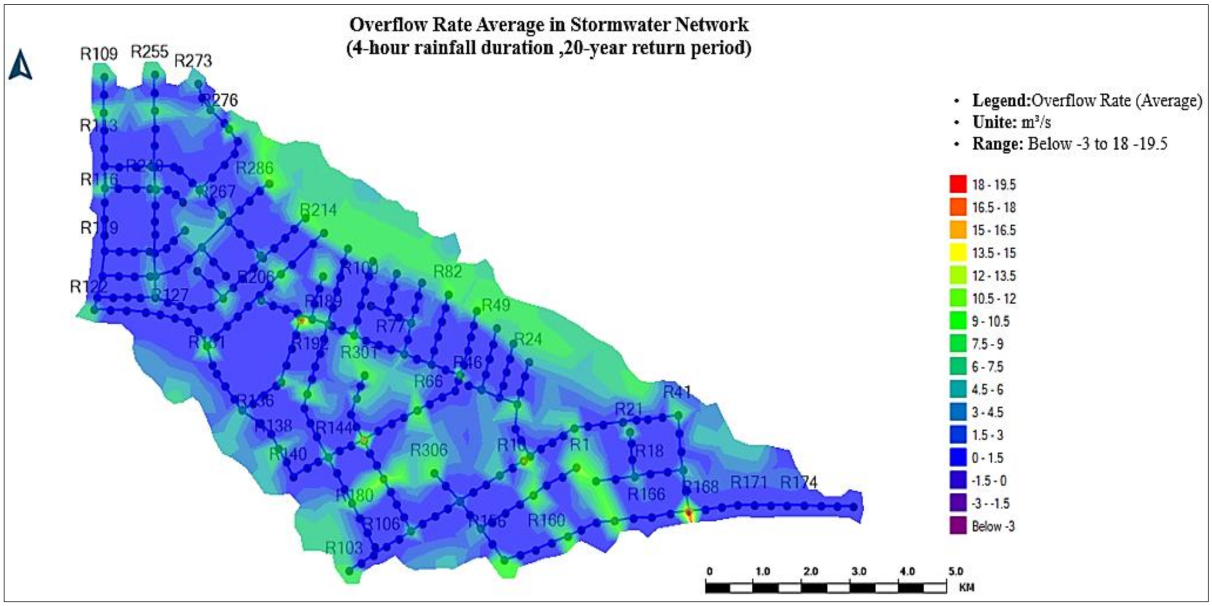
The Table 5.2 also outlines the scenario for a 3-hour rainfall duration. Key observations are:

- **Low Overflow Rates (≤ 1 m³/s):** For the 10-year return period, rates range from below 0 to 0-1 m³/s, indicating minimal overflow. As the return period increases, the range widens slightly, reflecting a higher potential for overflow in less frequent events.
- **Moderate Overflow Rates (1 m³/s - 10 m³/s):** Rates are between 1-2 and 9-10 m³/s for a 10-year return period, showing consistent risk with slight adjustments in longer return periods.
- **Maximum Overflow Rates (> 10 m³/s):** Overflow rates range from 10-11 to 14-15 m³/s for a 10-year return period, increasing significantly with longer return periods, highlighting substantial overflow potential in severe events.

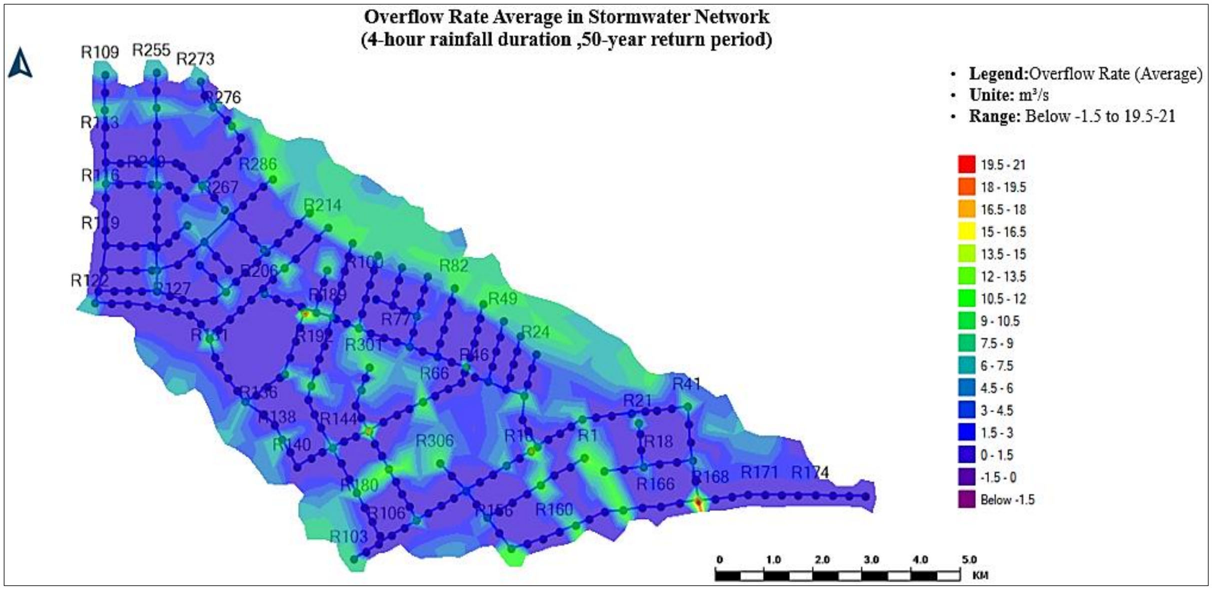
5.3.3 Scenario Simulation for a 4-Hour Rainfall Duration



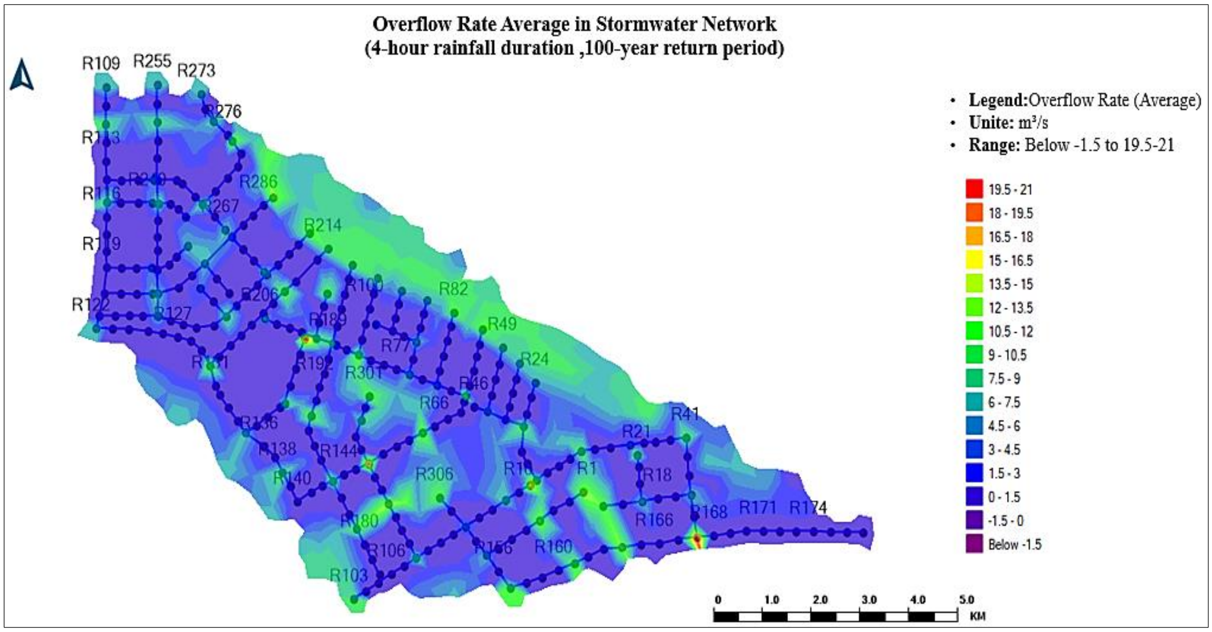
(a)


Overflow Rate Average in Stormwater Network
(4-hour rainfall duration ,20-year return period)
Legend: Overflow Rate (Average)
Unite: m³/s
Range: Below -3 to 18 -19.5
18 - 19.5
16.5 - 18
15 - 16.5
13.5 - 15
12 - 13.5
10.5 - 12
9 - 10.5
7.5 - 9
6 - 7.5
4.5 - 6
3 - 4.5
1.5 - 3
0 - 1.5
-1.5 - 0
-3 - -1.5
Below -3
R109
R255
R273
R276
R113
R249
R116
R286
R267
R119
R214
R122
R127
R206
R100
R82
R189
R49
R131
R77
R192
R301
R24
R46
R66
R136
R41
R21
R138
R144
R140
R306
R19
R1
R18
R168
R171
R174
R180
R166
R106
R156
R160
R103
0 1.0 2.0 3.0 4.0 5.0 KM

(b)


Overflow Rate Average in Stormwater Network
(4-hour rainfall duration ,50-year return period)
Legend: Overflow Rate (Average)
Unite: m³/s
Range: Below -1.5 to 19.5-21
19.5 - 21
18 - 19.5
16.5 - 18
15 - 16.5
13.5 - 15
12 - 13.5
10.5 - 12
9 - 10.5
7.5 - 9
6 - 7.5
4.5 - 6
3 - 4.5
1.5 - 3
0 - 1.5
-1.5 - 0
Below -1.5
R109
R255
R273
R276
R113
R249
R116
R286
R267
R119
R214
R122
R127
R206
R100
R82
R189
R49
R131
R77
R192
R301
R24
R46
R66
R136
R41
R21
R138
R144
R140
R306
R19
R1
R18
R168
R171
R174
R180
R166
R106
R156
R160
R103
0 1.0 2.0 3.0 4.0 5.0 KM

(c)



(d)

Figure 5.4: MIKE+ Simulation of Average Overflow Rates (m³/s) in Stormwater Network for 4-Hour Rainfall Durations and Different Return Periods: (a) 10 Years, (b) 20 Years, (c) 50 Years, (d) 100 Years

Table 5.3: 4-Hour Rainfall Duration: Average Overflow Rates Across Multiple Return Periods

Category	Return Period (Year)	Color Representation	Nodes	Overflow Rate Average Range (m³/s)
Low Overflow Rates (≤ 1 m³/s)	**10**	Blue Dark, Violet	R274, R6, R81, etc.	Below -6 to 0-1.5
	20	Blue Dark, Violet		Below -3 to 0-1.5
	50	Blue Dark, Violet		Below 1.5 to 0-1.5
	100	Blue Dark, Violet		Below 1.5 to 0-1.5
Moderate Overflow Rates (1 m³/s - 10 m³/s)	**10**	Blue Dark, Green	R130, R246, R247, R227, etc.	0-1 to 10.5-12
	20	Blue Dark, Green		0-1.5 to 9-10.5
	50	Blue Dark, Green		0-1.5 to 9-10.5
	100	Blue Dark, Green		0-1.5 to 9-10.5
Maximum Overflow Rates (> 10 m³/s)	**10**	Orange, Red, Green, Yellow	R45, R198, R71, R9	9-10.5 to 15-16.5
	20	Green, Yellow, Red, Orange	R9, R71, R198, R45	10.5-12 to 18-19.5
	50	Green, Yellow, Red, Orange	R285, R1, R62, R148, R9, R71, R198, R45	9-10.5 to 19.5-21
	100	Green, Yellow, Red, Orange	R164, R4, R285, R1, R62, R148, R9, R71, R198, R45	9-10.5 to 19.5-22

For a 4-hour rainfall duration, the Table 5.3 details the following:

- **Low Overflow Rates (≤ 1 m^3/s):** Overflow rates range from below -6 to 0-1.5 m^3/s for a 10-year return period, indicating minimal overflow. The range for longer return periods shows a slight increase in overflow risk.
- **Moderate Overflow Rates (1 m^3/s - 10 m^3/s):** Rates vary from 0-1 to 10.5-12 m^3/s for a 10-year return period, with consistent values across different return periods.
- **Maximum Overflow Rates (> 10 m^3/s):** Overflow rates increase from 9-10.5 to 15-16.5 m^3/s for a 10-year return period, reaching up to 19.5-22 m^3/s for a 100-year return period, reflecting the highest potential for overflow.

The analysis shows that as the return period increases, the potential for overflow rates grows across all rainfall durations. This is evident in both moderate and maximum overflow categories. Understanding these trends is essential for evaluating the capacity and resilience of the stormwater network and designing effective flood management strategies.

The sensitivity analyses reveal key parameters affecting hydraulic sensitivity in the stormwater network. Factors such as rainfall duration, return period, and network capacity significantly influence overflow rates. Variations in these parameters can lead to substantial differences in overflow risk, which underscores the need for robust models to account for different scenarios.

Key parameters impacting hydraulic sensitivity include:

- **Rainfall Duration:** Longer durations increase the volume of runoff, leading to higher overflow rates. This highlights the importance of considering extended rainfall events in flood risk assessments.
- **Return Period:** Longer return periods correspond to more severe rainfall events, which elevate the overflow rates. Understanding these patterns helps in designing infrastructure to withstand extreme conditions.
- **Network Capacity:** The capacity of the stormwater network to handle overflow is crucial. Identifying bottlenecks and areas with insufficient capacity enables targeted improvements to reduce flood risk.

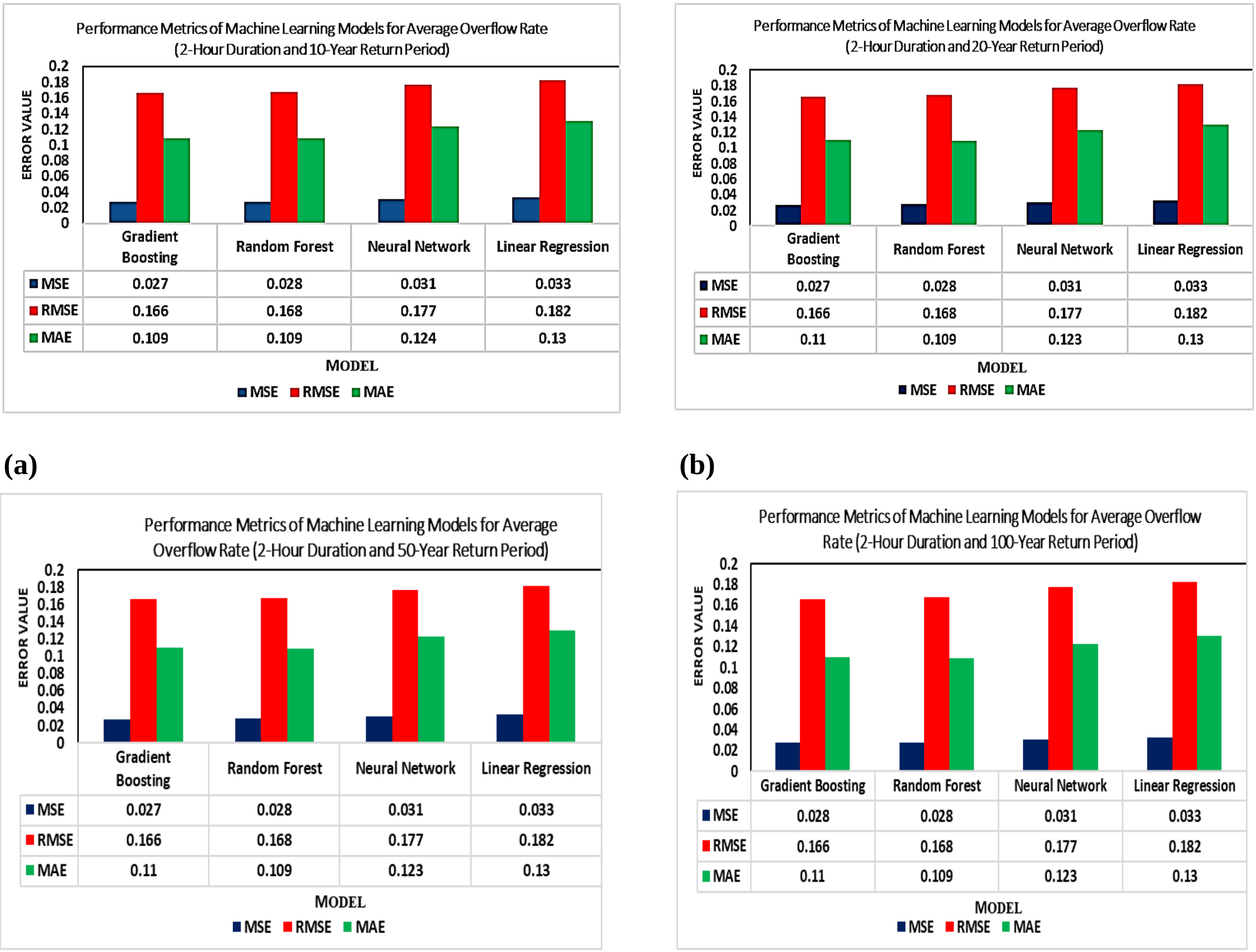
The results emphasize the need for adaptive flood management strategies that incorporate the impacts of varying rainfall durations and return periods. Enhanced monitoring and targeted interventions in areas with high overflow potential are essential for improving resilience and mitigating flood risks.

5.4 Machine Learning Results

5.4.1 Performance of Machine Learning Models

In Figure 5.5, the performance of classical models and ensemble machine learning models in predicting rainfall for various return periods and durations is depicted, along with cross-validation using 5 folds. The results from the training stage for each model, based on a 2-hour rainfall duration, are shown under different conditions, including 10-year, 20-year, 50-year, and 100-year return periods

.



(a) (b)

(c) (d)

Figure 5.5: Performance Evaluation of ML Models for Predicting Average Overflow Rate Across Various Return Periods (2-Hour Rainfall Duration)

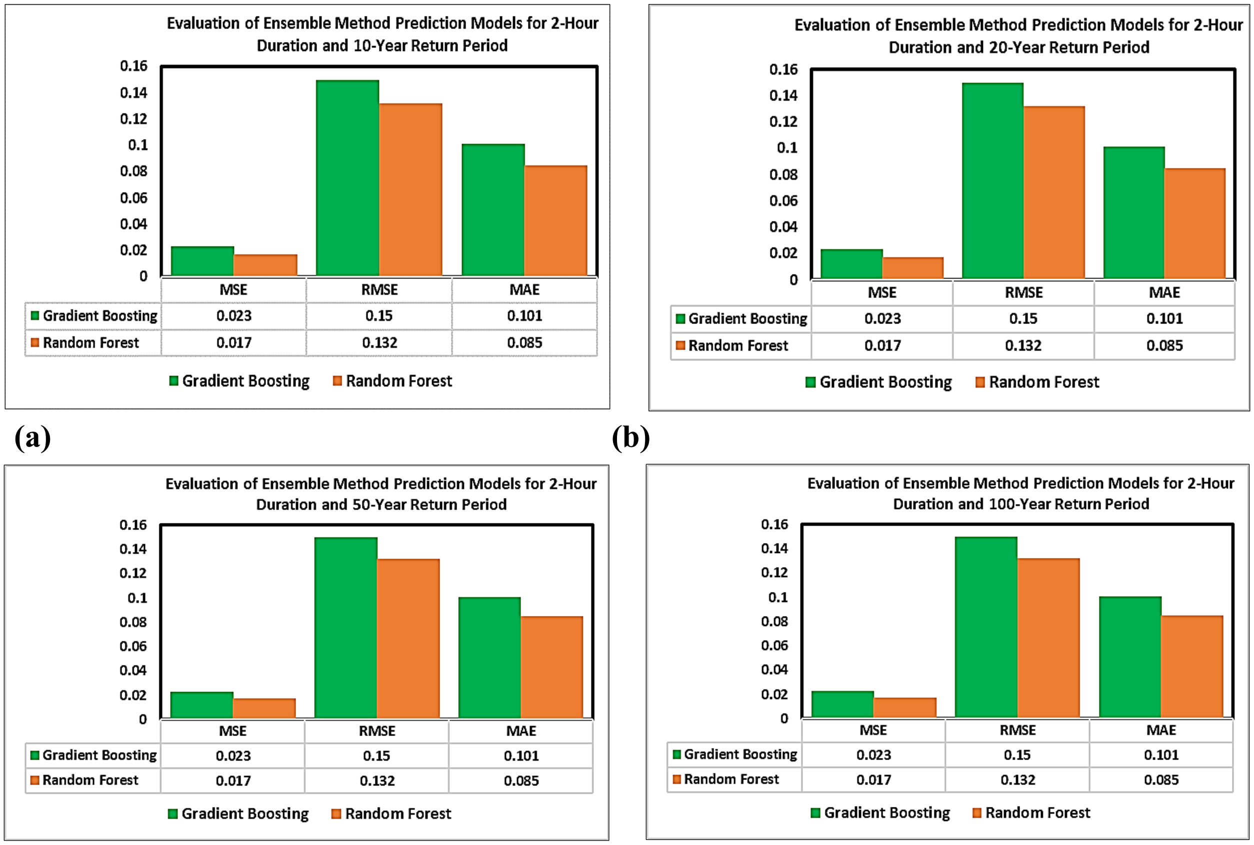
The results provide a detailed assessment of various machine learning models for predicting overflow incidents in stormwater manholes across different return periods—specifically, for a 2-hour rainfall event with return periods of 10, 20, 50, and 100 years, as depicted in Figures 5.5(a) through 5.5(d). These metrics reveal the models' effectiveness in forecasting overflow in stormwater networks under varying rainfall durations and intensities.

Summarizing the findings, for a 2-hour rainfall event, all models demonstrated comparable performance, with MSE values ranging between 0.027 and 0.033, RMSE between 0.166 and 0.182, and MAE between 0.109 and 0.130. As the return periods increased to 20, 50, and 100 years, the models maintained stable performance, showing minimal variation in these metrics. Notably, the Random Forest and Gradient Boosting models performed slightly better than the Neural Network and Linear Regression models, as evidenced by their lower MSE, RMSE, and MAE values, suggesting a more robust capacity to capture the complexities of rainfall-overflow relationships within stormwater networks.

This slight performance advantage of Random Forest and Gradient Boosting models underscores their efficacy in interpreting complex interactions between rainfall events and overflow dynamics an essential factor in addressing urban flooding challenges exacerbated by urbanization and climate change (Lu et al., 2023; Mukaka, 2012). The findings offer essential insights for future refinement of these ensemble models, providing guidance for practical flood prediction applications in urban stormwater management systems. These insights support the practical implementation and further optimization of machine learning approaches in flood forecasting contexts.

5.4.2 Predictions Using ELM

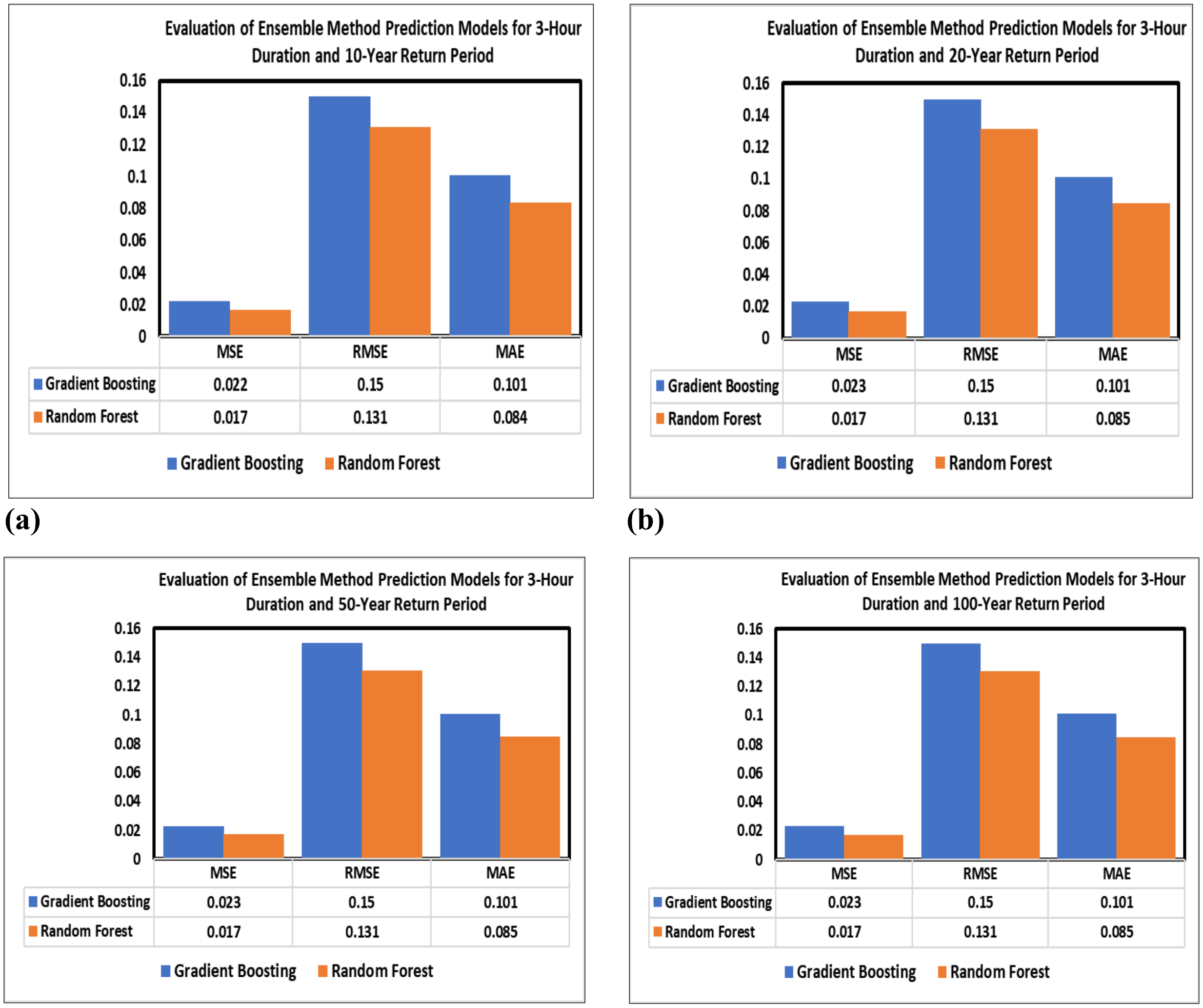
This section presents the forecast results for the best-performing algorithms, Random Forest and Gradient Boosting, in predicting the average overflow rate for a stormwater network. The predictions were made for different rainfall durations and return periods. The performance measures employed comprise (MSE), (RMSE), and (MAE).



(a) **(b)**

(c) **(d)**

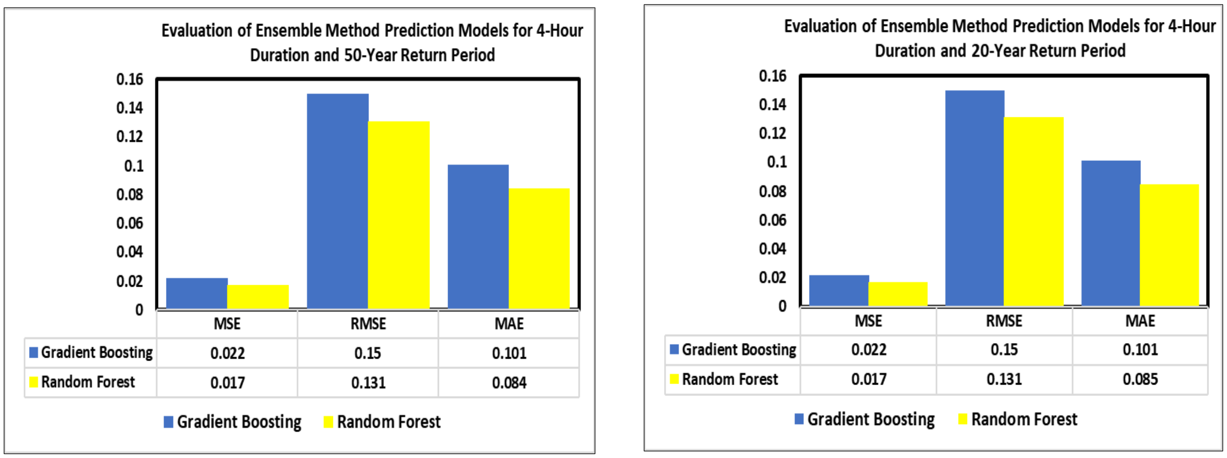
Figure 5.6: Performance Metrics of Ensemble Prediction Models for 2-Hour Rainfall Duration Across 10-Year (a), 20-Year (b), 50-Year (c), and 100-Year (d) Return Period



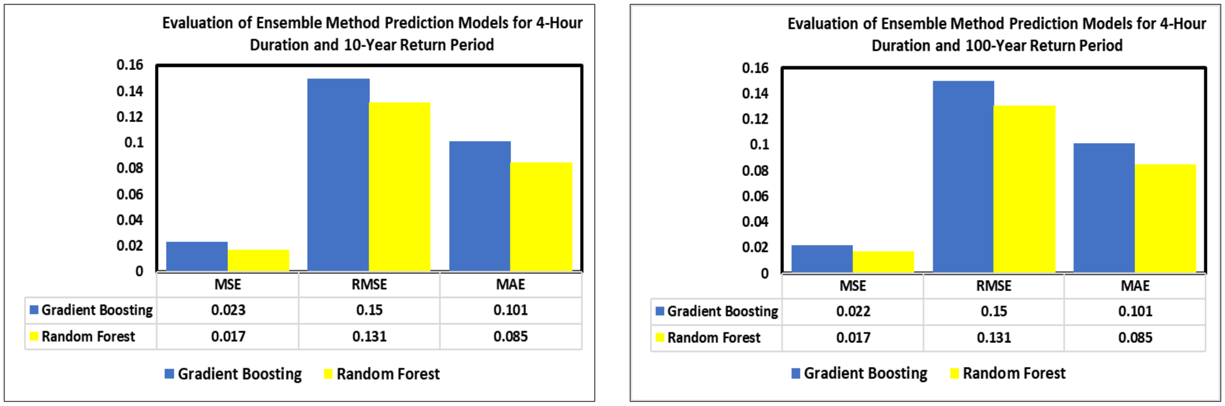
(a) **(b)**

(c) **(d)**

Figure 5.7: Performance Metrics of ELM for 3-Hour Rainfall Duration Across 10-Year (a), 20-Year (b), 50-Year (c), and 100-Year (d) Return Periods



(a) **(b)**



(c) **(d)**

Figure 5.8: Performance Metrics of ELM for 4-Hour Rainfall Duration Across 10-Year (a), 20-Year (b), 50-Year (c), and 100-Year (d) Return Periods

Figures 5.6 to 5.8 evaluate the performance of Gradient Boosting and Random Forest models in predicting overflow rates within stormwater networks across different rainfall durations (2, 3, and 4 hours) and return periods (10, 20, 50, and 100 years). Both models show consistent and reliable performance across all scenarios, with Random Forest slightly outperforming Gradient Boosting in predictive accuracy (RMSE and MAE). This consistency across durations and return periods highlights the models' robustness and capability to generalize effectively under varying environmental conditions.

In summary, Figures 5.5 to 5.8 illustrate the effectiveness of various machine learning models, particularly ensemble techniques, in predicting floods across differing return periods and rainfall durations. Specifically, Gradient Boosting and Random Forest consistently outperform other models, such as Neural Networks and Linear Regression, in key performance metrics like MSE, RMSE, and MAE. This superior performance is maintained across all tested rainfall durations and return periods, underscoring the benefits of ensemble methods in leveraging the strengths of multiple models while addressing their individual limitations (Dietterich, 2000).

The exceptional performance of the Gradient Boosting and Random Forest models can be attributed to their ability to uncover intricate, nonlinear relationships within the data. Gradient Boosting, with its iterative approach of building and aggregating weak learners into a robust predictor, is particularly effective at handling diverse and imbalanced dataset(Natekin & Knoll, 2013). Random Forest enhances model stability and reduces overfitting by averaging the predictions from a collection of decision trees, leading to improved generalization across various datasets(Breiman 2001).

These findings are consistent with previous research that emphasizes the success of ensemble learning methods in hydrological modeling. For example, a study by Hukkeri et al. (2023) showcased the use of ensemble learning methods combined with advanced machine learning models to achieve precise meteorological drought predictions at a district level. In another study, (Zounemat-Kermani et al. 2021) showcased the use of ensemble learning methods combined with advanced machine learning models to achieve precise meteorological drought predictions at a district level Dang et al. (2024) showcased that the Random Forest (RF) model effectively captured the complex interactions between input variables and flood depth, resulting in precise predictions. The RF algorithm's outstanding performance made it an ideal candidate for replacing the numerical model in this urban flood depth modeling research.

This research demonstrated enhanced accuracy in predicting the average stormwater overflow rate, outperforming previous studies that employed ensemble learning models. (W. Chen et al., 2018; Hukkeri et al., 2023) . It also emphasizes and confirms the effectiveness of machine learning techniques. (Dang et al., 2024), particularly ensemble learning, in enhancing flood forecasting and management.

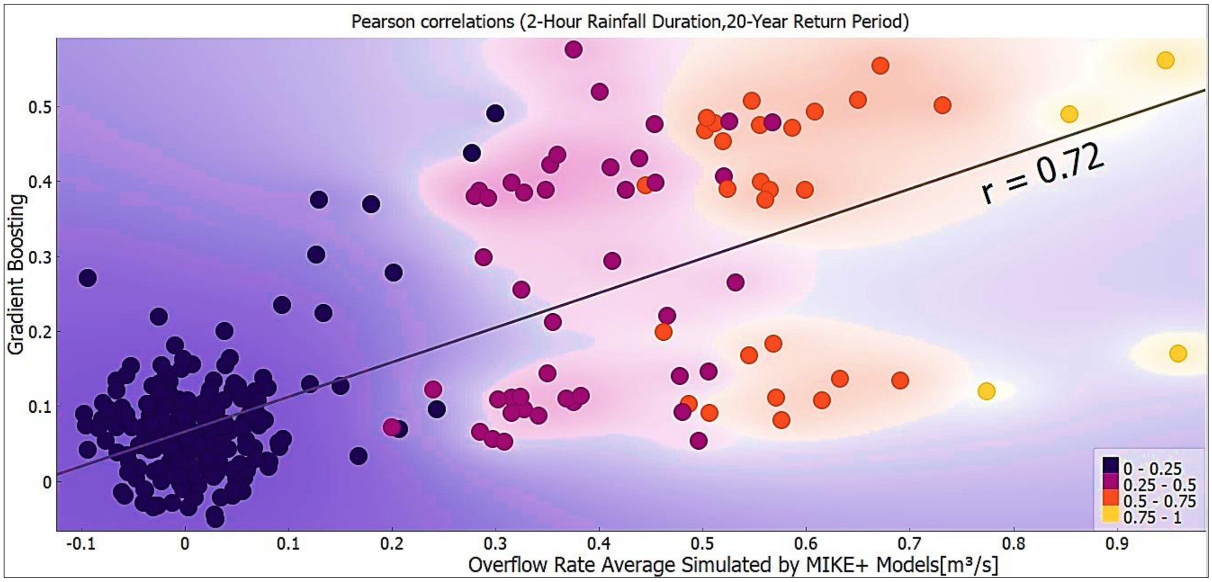
While the results are promising, several limitations warrant attention. The accuracy of the model largely depends on the quality and quantity of input data; insufficient or biased datasets can lead to suboptimal model performance. Therefore, future research should focus on incorporating more extensive datasets that encompass a broader range of hydrological conditions. Moreover, although ensemble approaches typically help reduce overfitting, there remains a potential risk, especially when dealing with limited datasets. To mitigate this risk, it is essential to employ cross-validation techniques and regularization methods to decrease the likelihood of overfitting.

In conclusion, ensemble models demonstrate reliable and consistent performance in forecasting overflow rates in stormwater networks across various rainfall durations and return periods. While Random Forest slightly edges out in terms of RMSE and MAE values, both models are effective for stormwater management and infrastructure planning. Their stability and robustness indicate their suitability for practical applications, where accurate overflow rate predictions are crucial for informed decision-making. Future research and validation will be necessary to evaluate the models' performance under different environmental conditions and to further enhance their predictive capabilities.

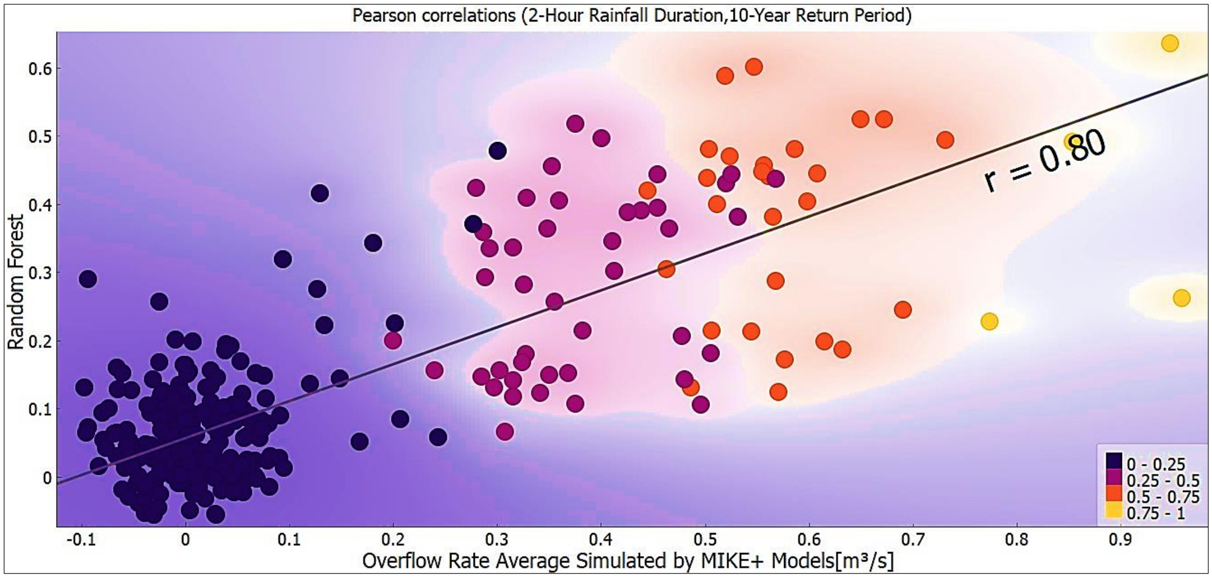
5.5 Comparative Analysis

5.5.1 Comparison for 2-Hour Rainfall Duration

Figures 5.9 to 5.12 illustrate the association between MIKE+ results and ensemble model predictions across various return periods using Pearson correlation coefficients. This analysis emphasizes the top-performing machine learning models, providing a comparative assessment of the ensemble model's effectiveness, computational efficiency, and accuracy. The Pearson correlation coefficients shown in these figures play a key role in evaluating the performance of the leading prediction models.

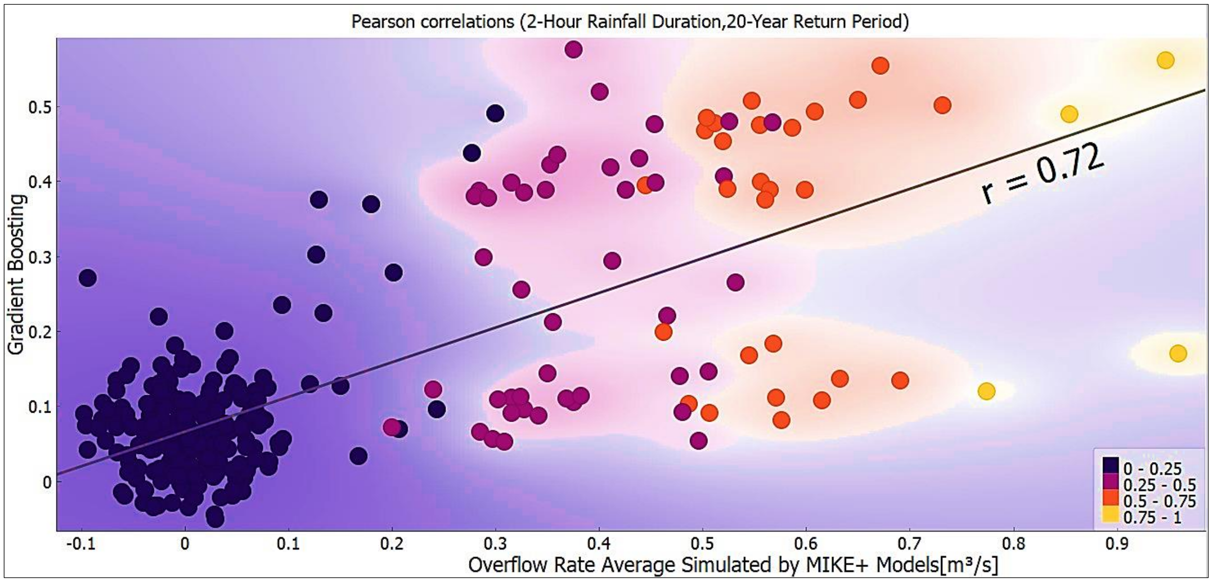


(a)

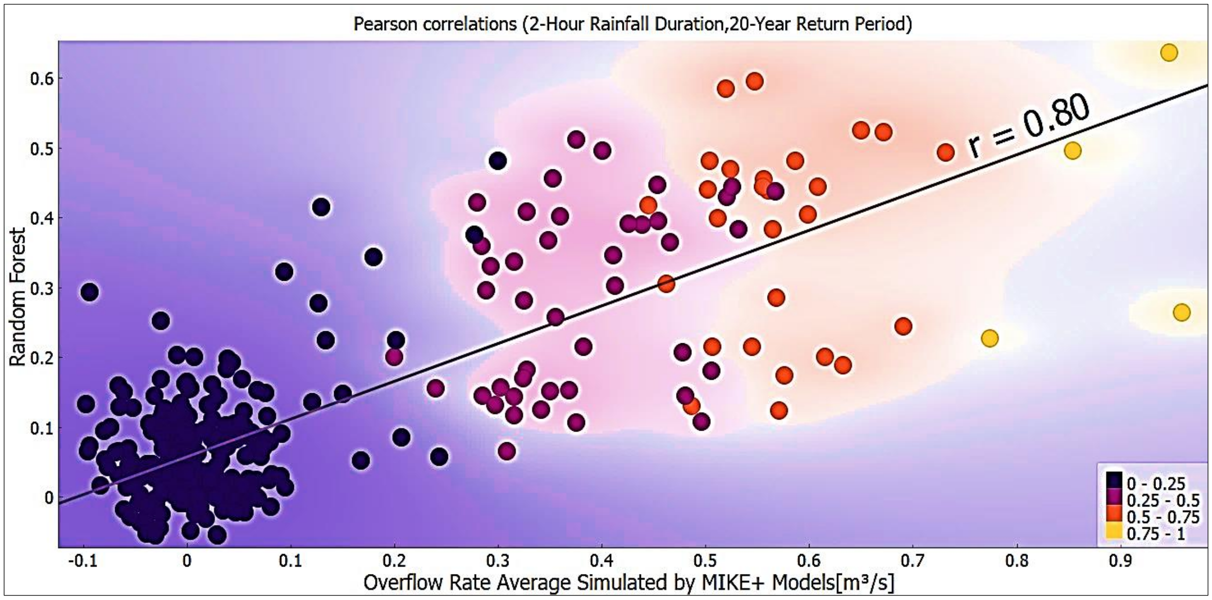


(b)

Figure 5.9: S Scatter Plot Analysis of Pearson Correlations for Ensemble Models and MIKE+ Simulations: 2-Hour Rainfall Duration at a 10-Year Return Period: (a) Gradient Boosting, (b) Random Forest

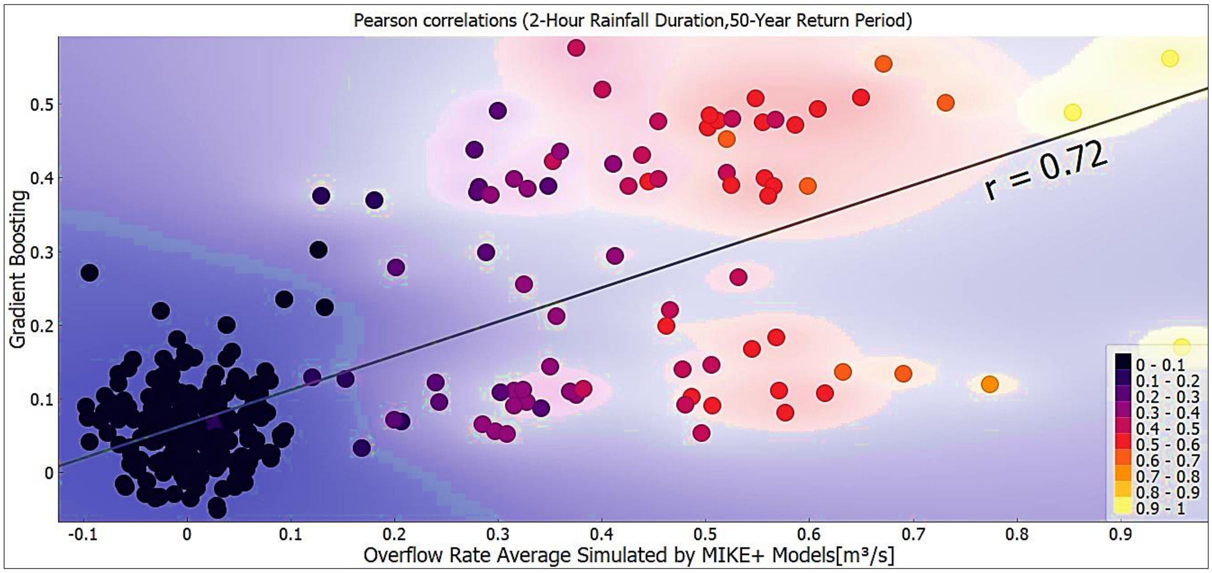


(a)

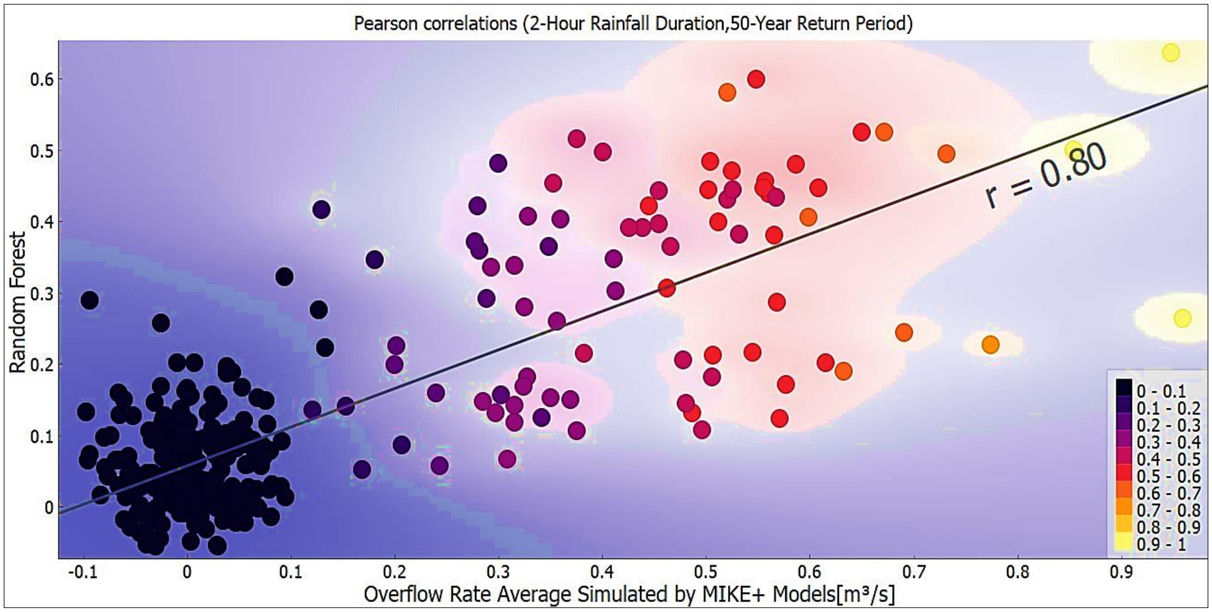


(b)

Figure 5.10: Scatter Plot of Pearson Correlations Between Ensemble Model Predictions and MIKE+ Simulated Overflow Rates for 2-Hour Rainfall Duration at a 20-Year Return Period: a) Gradient Boosting, (b) Random Forest

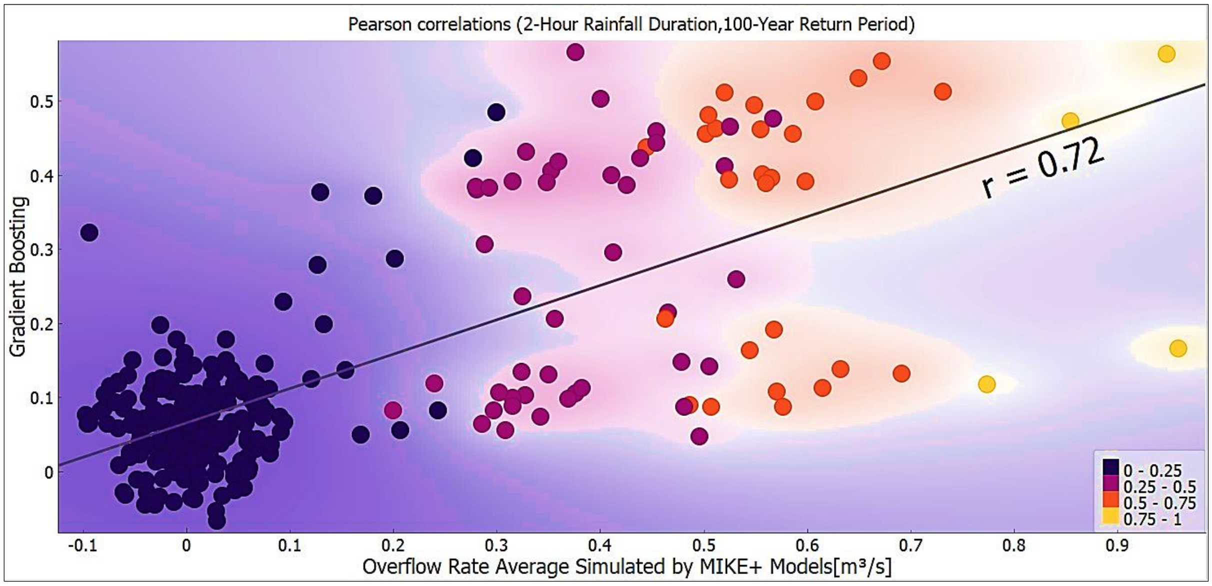


(a)

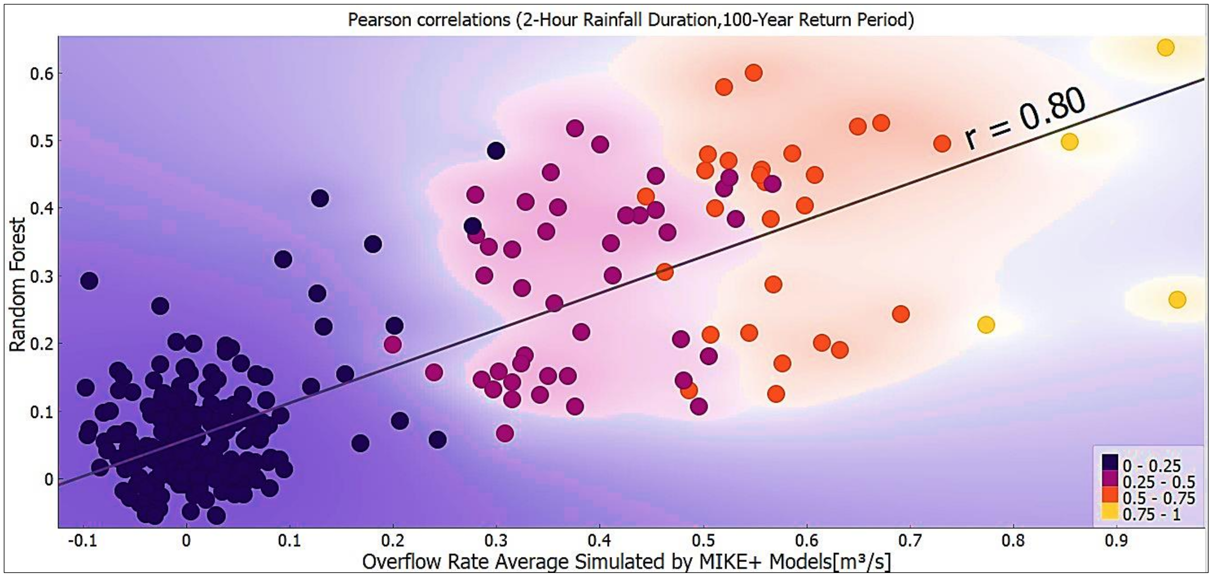


(b)

Figure 5.11: Scatter Plot Analysis of Pearson Correlations for Ensemble Models and MIKE+ Simulations: 2-Hour Rainfall Duration at a 50-Year Return Period (a) Gradient Boosting, (b) Random Forest



(a)



(b)

Figure 5.12: Scatter Plot Analysis of Pearson Correlations for Ensemble Models and MIKE+ Simulations: 2-Hour Rainfall Duration at a 100-Year Return Period: (a) Gradient Boosting, (b) Random Forest

Based on Pearson correlation analysis, the overflow rates predicted by MIKE+ across various return periods show a significant positive correlation with the ensemble methods, particularly Random Forest and Gradient Boosting. This strong correlation highlights the effectiveness of ensemble models in accurately predicting overflow rates, reflecting their ability to capture the intricate dynamics within the stormwater network.

The performance of ensemble models is illustrated in Figures 5.9 to 5.12, highlighting the Average Overflow Rate Comparison for different return periods with a 2-hour rainfall duration. The top-performing models, specifically Random Forest and Gradient Boosting, show considerable potential in predicting overflow rates in environmental and hydrological contexts. Their application could greatly improve flood prediction and management efforts. When selecting models for predicting rainfall durations (2-hour, 3-hour, 4-hour), it is recommended to use the same models that exhibited strong correlations in forecasting the average overflow rate, indicating their suitability for various hydrological prediction and management tasks.

In conclusion, ensemble models are highly effective in enhancing the accuracy and robustness of machine learning models. They are extensively used in real-world applications and have consistently delivered state-of-the-art results across various tasks (Zhou, 2012).

5.5.2 Comparison for 3-Hour and 4-Hour Rainfall Durations

Figures 5.13 and 5.14 display heatmaps showing Pearson correlations between ensemble model predictions and MIKE+ simulated overflow rates across different rainfall durations and return

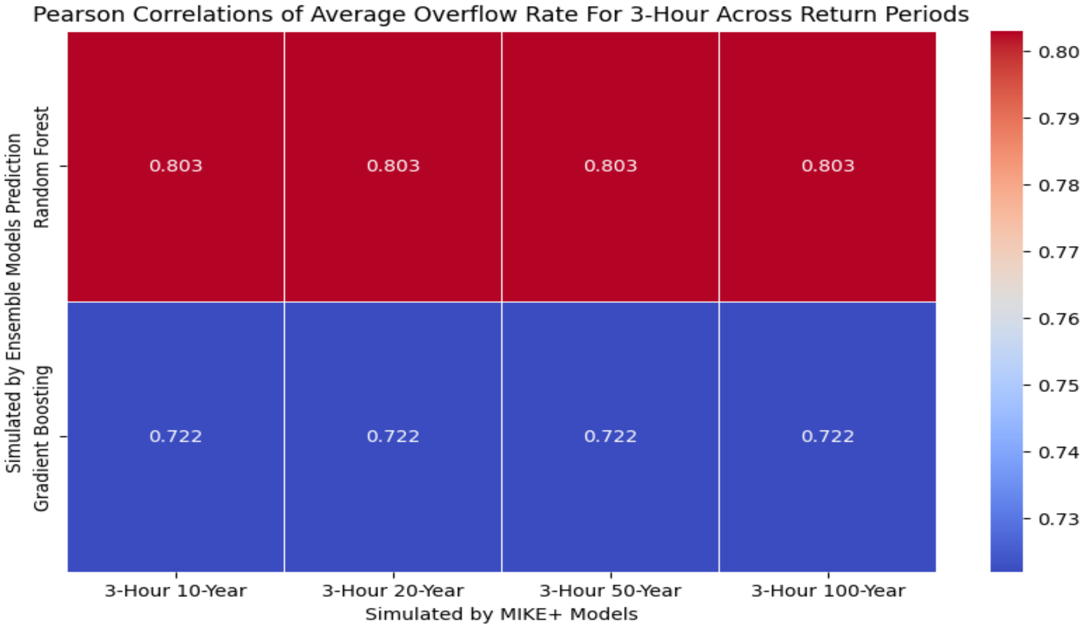


Figure 5.13: Heatmap Analysis of Pearson Correlations: Ensemble Method Predictions and MIKE + Simulated Average Overflow Rates for 3-Hour Rainfall Duration

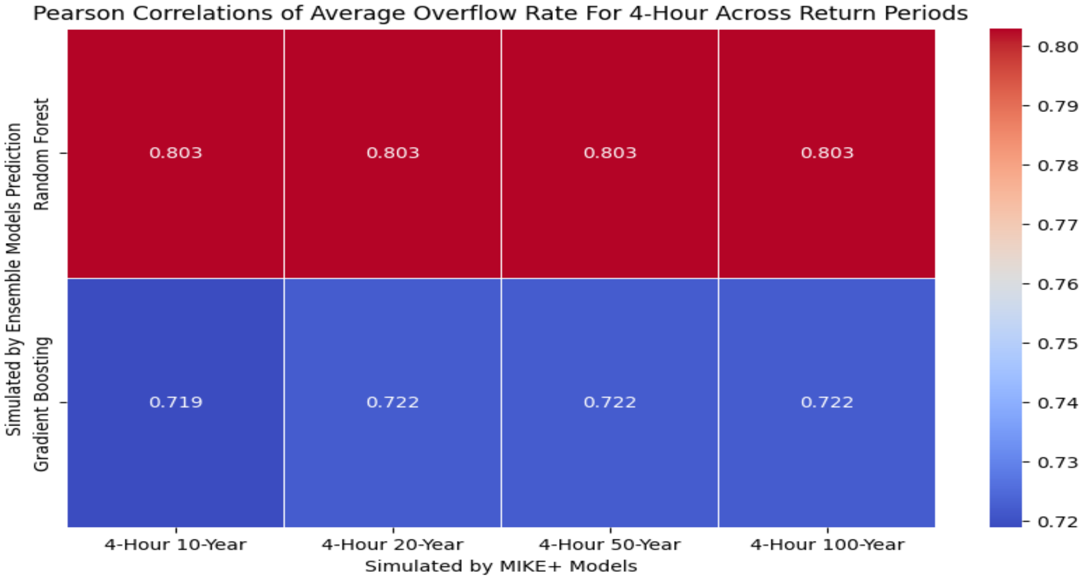


Figure 5.14: Heatmap Analysis of Pearson Correlations: Ensemble Method Predictions and MIKE+ Simulated Average Overflow Rates for 3-Hour Rainfall Duration

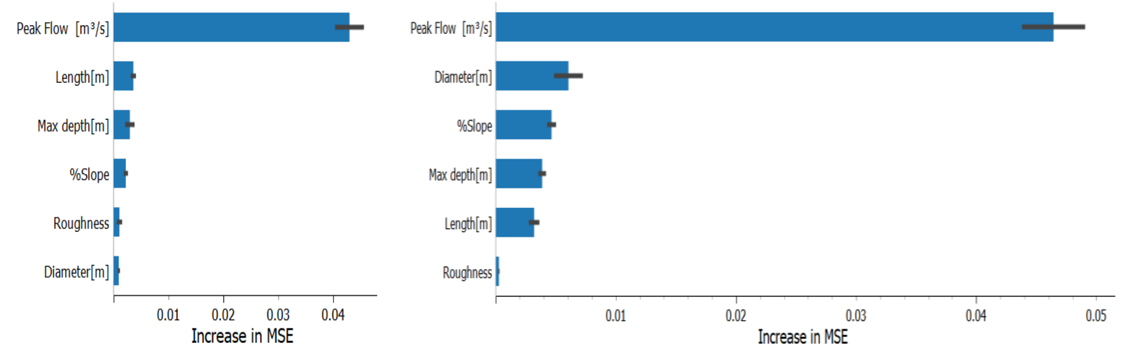
The heatmap reveals consistent Pearson correlation coefficients between the predictions made by ensemble models and MIKE+ simulations for the average overflow rate across different rainfall durations and return periods. Both models show correlations around +0.803 for the 3-hour rainfall duration across all return periods, highlighting their reliability and robustness in predicting the average overflow rate under varying conditions.

The idea of diversity in ensemble learning models is widely acknowledged in the literature and is known for enhancing model performance and accuracy. Breiman (2001) proposes that introducing randomness in data sampling and feature selection within the Random Forest algorithm enhances diversity among individual trees. This reduced correlation between their predictions results in a more robust and accurate ensemble model. The diversity among these uncorrelated models significantly lowers overall variance, leading to improved performance. Similarly, Gradient Boosting, as described by Friedman (2001) , explains that Gradient Boosting maintains minimal correlation among its predictions, leading to improved accuracy. Furthermore, research supports the idea that greater model diversity, reflected in a low Pearson correlation, boosts the effectiveness of ensembles and reduces errors. Dietterich (2000) highlights the critical role of diversity in techniques like bagging and boosting, crediting their effectiveness to reducing correlation among models by using varied training samples or methodologies. Kuncheva and Whitaker (2003) demonstrate that lower Pearson correlation, indicative of greater diversity, significantly boosts the accuracy of ensemble models. Brown et al. (2005) investigate methods for enhancing diversity, such as resampling and feature selection, to decrease correlation and enhance the performance of ensemble models.

Ensemble models offer precise predictions of overflow rates, serving as a cost-effective and efficient alternative to traditional simulation methods. Their consistent and reliable correlations have important implications for urban planning and stormwater management, allowing stakeholders to make well-informed decisions on infrastructure development, flood mitigation, and urban resilience. The integration of machine learning models with MIKE+ simulations is also emphasized, leveraging the strengths of both approaches to create accurate and comprehensive predictive tools for managing stormwater. As a result, ensemble models present a viable alternative to complex simulation systems like MIKE+, providing a practical and computationally efficient method for engineers and planners to evaluate and manage flooding risks in urban environments.

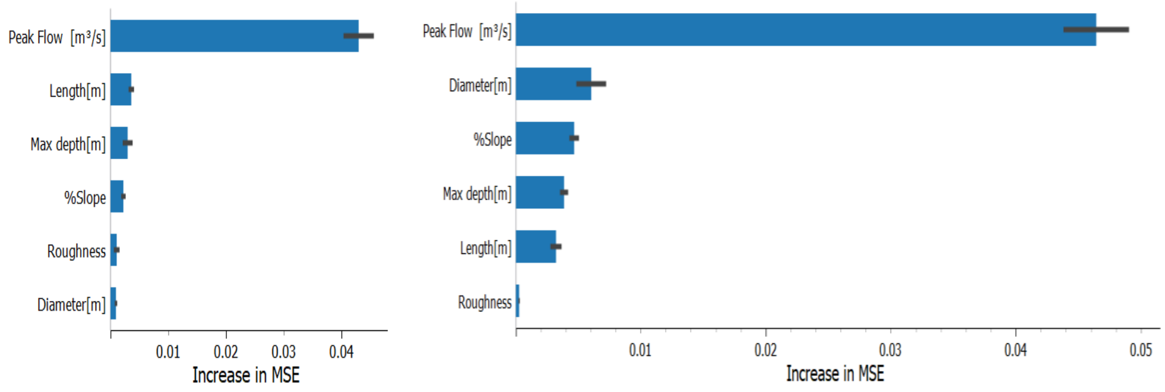
5.6 Feature Importance Analysis

This study investigates the impact of various features on overflow rates and emphasizes the importance of identifying the most influential factors within a stormwater network. Understanding these key features is crucial for effective risk management in such networks. By employing ensemble models, specifically Random Forest and Gradient Boosting, the study evaluates feature importance across different rainfall durations and return periods. The analysis uses the Mean Squared Error (MSE) metric to assess each feature's predictive power. The dataset includes one meta-attribute and six key features: peak flow, maximum depth, length, slope, roughness, and diameter, with six data instances representing different scenarios. This thorough approach sheds light on the complex behavior of stormwater networks and highlights the significance of understanding feature impacts to improve risk mitigation strategies.



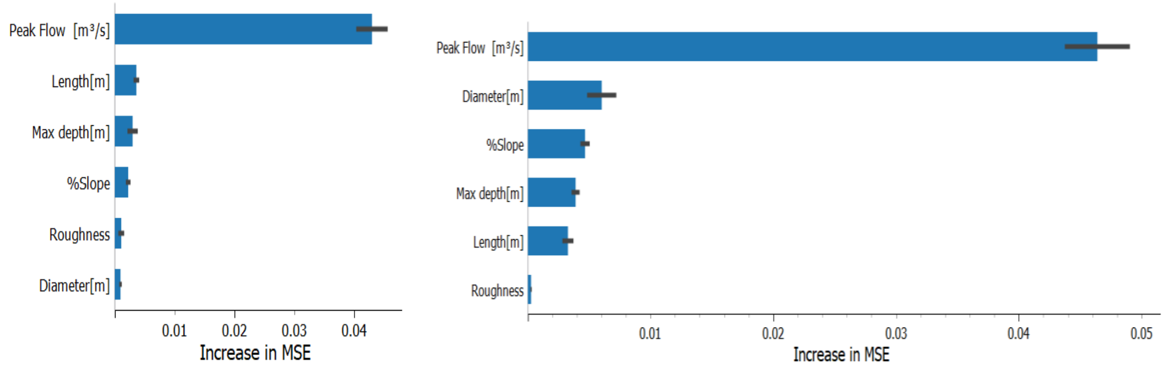
(a) **(b)**

Figure 5.15: MSE-Based Feature Importance for 2-Hour Rainfall Duration and 10-Year Return Period: a) (GB), b) (RF)



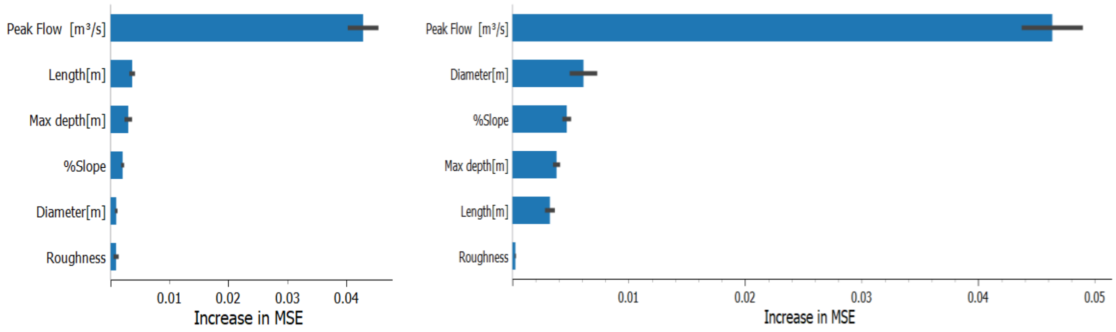
(a) **(b)**

Figure 5.16: MSE-Based Feature Importance for 2-Hour Rainfall Duration and 20-Year Return Period: a) (GB), b) (RF)



(a) **(b)**

Figure 5.17: MSE-Based Feature Importance for 2-Hour Rainfall Duration and 50-Year Return Period: a) (GB), b) (RF).



(a) **(b)**

Figure 5.18: MSE-Based Feature Importance for 2-Hour Rainfall Duration and 100-Year Return Period: a) (GB), b) (RF)

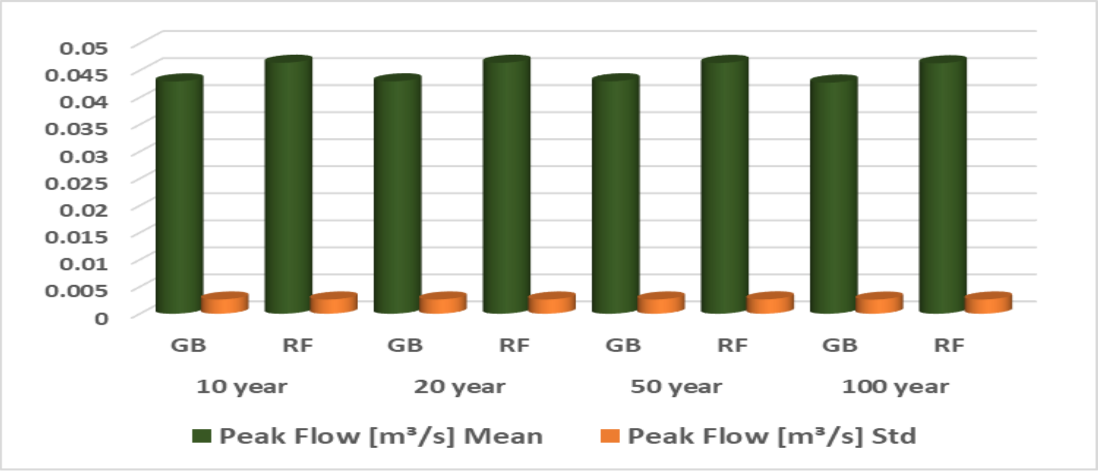


Figure 5.19: Assessment of Top Features' Importance, Including Peak Flow, Using Ensemble Learning Across Return Periods (10, 20, 50, 100 Years): Mean and Standard Deviation Calculation for 2-Hour Rainfall Duration

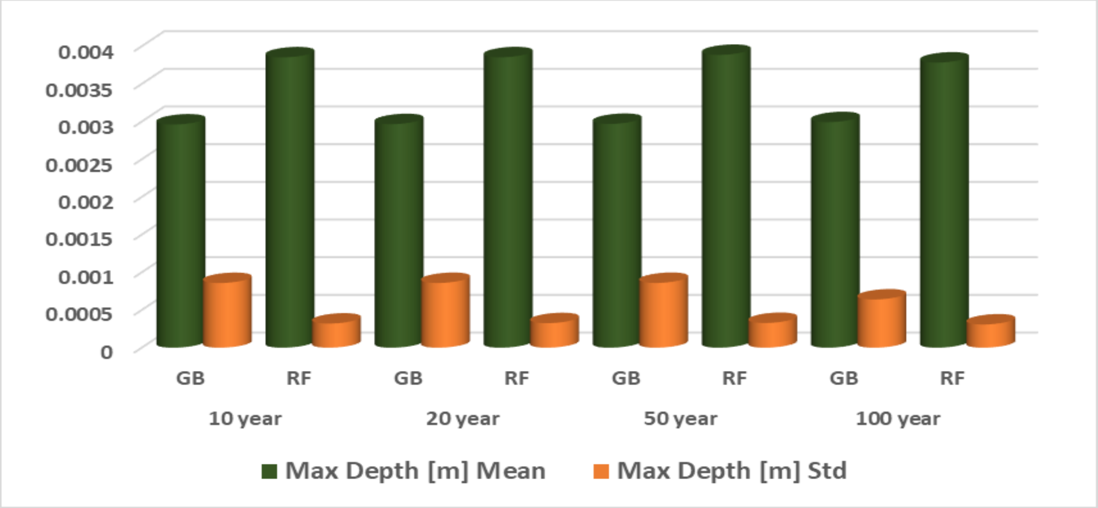


Figure 5.20: Assessment of Top Features' Importance, Including Max Depth, Using Ensemble Learning Across Return Periods (10, 20, 50, 100 Years): Mean and Standard Deviation Calculation for 2-Hour Rainfall Duration

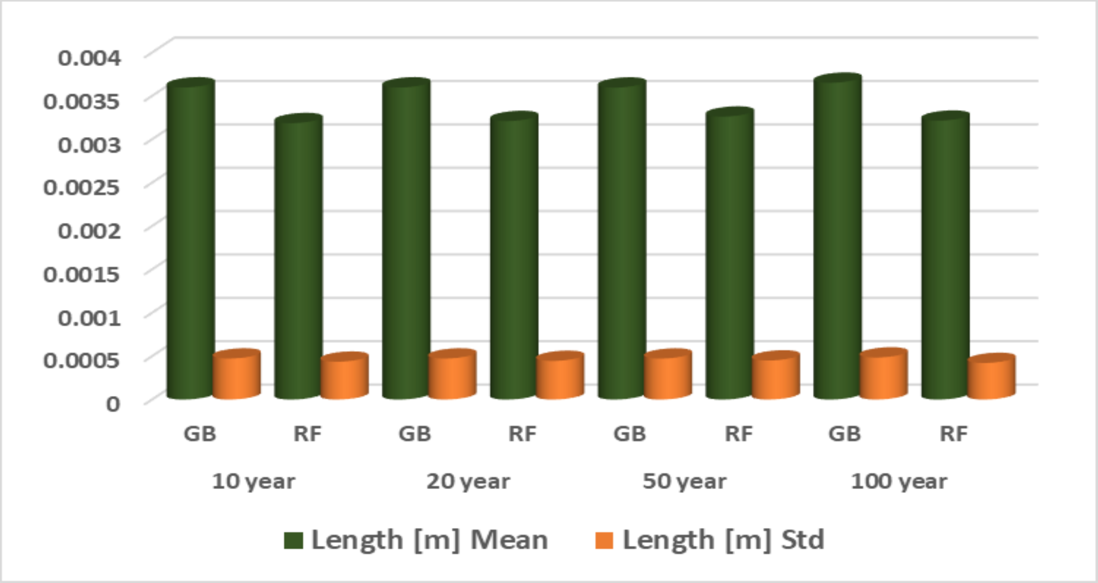


Figure 5.21: Assessment of Top Features' Importance, Including Lenght, Using Ensemble Learning Across Return Periods (10, 20, 50, 100 Years): Mean and Standard Deviation Calculation for 2-Hour Rainfall Duration

Figures 5.15 to 5.18 illustrate the feature importance metric values for the Gradient Boosting (GB) and Random Forest (RF) algorithms using the MSE method across different return periods (10, 20, 50, and 100 years) for a 2-hour rainfall duration, in predicting overflow rates in a stormwater network. Figures 5.19 to 5.21 show the mean and standard deviation (std) of each feature, calculated across the various return periods. These statistics provide insights into the stability and reliability of the feature importance scores. The analysis reveals that:

- Peak Flow is the most influential feature for predicting overflow rates across all rainfall durations and return periods. Since Peak Flow indicates the volume of water flowing through the network, higher values lead to increased overflow rates.

- Max Depth ranks second in importance. As Max Depth represents the water level within the network, higher levels correspond to higher overflow rates.
- Length is the third most significant feature. It represents the distance water travels through the network; longer lengths result in higher overflow rates.

Rainfall duration and return period have minimal impact on feature importance scores due to the capability of Gradient Boosting and Random Forest models to capture the relationship between these factors and overflow rates. The study provides valuable insights into the key variables affecting network performance across different return periods. Understanding the significance of these attributes helps in prioritizing maintenance and improvements to effectively mitigate overflow risks. This research underscores that ensemble learning models, especially Gradient Boosting and Random Forest, offer robust and precise predictions for flood events across varying conditions, highlighting their potential to enhance flood prediction and management. Future research should address the identified limitations to further improve model reliability and applicability.

5.7 Conclusion

This research highlights the potential of combining physically-based models with machine learning techniques to improve stormwater overflow rate predictions in urban drainage systems. The MIKE+ model, despite its reliance on limited observed data, showed strong performance in simulating water depths, as indicated by a high R^2 value. However, the incorporation of ensemble learning techniques, specifically Random Forest and Gradient Boosting, significantly improved the prediction accuracy of overflow rates across various rainfall durations and return periods (Boughandjioua et al., 2024). These results highlight the robustness and reliability of machine learning models in handling complex hydrological processes, particularly in scenarios where observed data is sparse or unavailable. The investigation into ensemble learning as an alternative to classical models underscores the importance of understanding the sensitivity of hydraulic systems to different rainfall durations and the impact factors influencing stormwater networks (Boughandjioua et al., 2024).

An important aspect of this study was the use of the Mean Squared Error (MSE) method to assess feature importance within the machine learning models. By analyzing the contribution of different input variables to the model’s predictions, the study identified key factors that significantly influence stormwater overflow rates. This feature importance analysis provides valuable insights into the underlying dynamics of stormwater systems, helping to prioritize variables that should be closely monitored or managed in practice. The MSE-based feature importance also reinforces the strength of ensemble learning models in not only improving prediction accuracy but also offering a deeper understanding of the factors driving hydrological processes.

General Conclusion

General Conclusion

In conclusion, this thesis presents a comprehensive study on the impact of climate change on urban drainage systems, with a particular focus on the growing challenges caused by increased rainfall intensity and frequency. The research highlights the vulnerabilities of urban areas in Algeria, especially Bir Farina, to flooding due to rapid urbanization and insufficient drainage infrastructure. By integrating traditional hydrological modeling through MIKE+ software with advanced machine learning techniques, particularly ensemble learning models like Random Forest and Gradient Boosting, this study addresses the limitations of classical models in predicting stormwater overflow rates.

The research provides an in-depth examination of the use of ensemble models based on machine learning algorithms to predict average overflow rates within stormwater networks. Through the implementation of the SWN-ML framework, which combines MIKE+ simulations with machine learning models, specifically Random Forest and Gradient Boosting, this study enhances the accuracy of overflow rate predictions. The results validate the efficacy of the proposed approach, demonstrating strong correlations between the predictions made by the ensemble models and the MIKE+ simulations across varying rainfall durations, return periods, and network configurations.

Notably, the study underscores the predictive capabilities of Gradient Boosting and Random Forest models, highlighting their ability to capture complex relationships in the data and provide reliable estimates. These findings suggest that machine learning methodologies offer a computationally efficient and accurate alternative to traditional simulation models for forecasting overflow rates in stormwater networks. The insights from this research make a significant contribution to improving flood risk management strategies and developing early warning systems, which can mitigate flood-related hazards in urban environments.

For future work, particularly with larger datasets and increasing complexity, there is potential to enhance these models by exploring novel deep learning techniques, including reinforcement learning, to further improve predictive accuracy and robustness.

References

References

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O’Connell, P. E., & Rasmussen, J. (1986). An introduction to the European Hydrological System — Systeme Hydrologique Europeen, “SHE”, 1: History and philosophy of a physically-based, distributed modelling system. Journal of Hydrology, 87(1–2), 45–59. https://doi.org/10.1016/0022-1694(86)90114-9

Abdelkebir, B. (2022). Contribution à la gestion des eaux pluviales en milieu urbanisé : Cas de la ville de Guelman (Doctoral dissertation). Université 8 Mai 1945 Guelma.

Acharki, S., Amharref, M., & Kassogue, H. (2017). Modélisation de l ’ impact des changements climatiques sur les ressources en eau : Cas du périmètre du Gharb Bulletin d ’ Information de l ’ Académie Hassan II des Sciences et Techniques n ° Hors-série Modélisation de l ’ impact des changements climatique. 2018(October).

Adnan, R. M., Petroselli, A., Heddam, S., Santos, C. A. G., & Kisi, O. (2021). Short term rainfall-runoff modelling using several machine learning methods and a conceptual event-based model. Stochastic Environmental Research and Risk Assessment, 35(3), 597–616. https://doi.org/10.1007/s00477-020-01910-0

Ahlman, S. (2006). SEWSYS ett verktyg för att bestämma källor till dagvattenföroreningar och pröva olika åtgärder.

Ahmed, F., Moors, E., Khan, M. S. A., Warner, J., & Terwisscha van Scheltinga, C. (2018). Tipping points in adaptation to urban flooding under climate change and urban growth: The case of the Dhaka megacity. Land Use Policy, 79, 496–506. https://doi.org/10.1016/j.landusepol.2018.05.051

Alfaro, E., Gámez, M., & García, N. (2013). adabag: An R Package for Classification with Boosting and Bagging. In JSS Journal of Statistical Software (Vol. 54, Issue 2). http://www.jstatsoft.org/

Amatya, D. M., Walega, A., Callahan, T. J., Morrison, A., Vulava, V., Hitchcock, D. R., Williams, T. M., & Epps, T. (2022). Storm event analysis of four forested catchments on the Atlantic coastal plain using a modified SCS-CN rainfall-runoff model. Journal of Hydrology, 608, 127772. https://doi.org/10.1016/J.JHYDROL.2022.127772

Arabameri, A., Saha, S., Chen, W., Roy, J., Pradhan, B., & Bui, D. T. (2020). Flash flood susceptibility modelling using functional tree and hybrid ensemble techniques. Journal of Hydrology, 587, 125007. https://doi.org/10.1016/J.JHYDROL.2020.125007

Araghinejad, S., Azmi, M., & Kholghi, M. (2011). Application of artificial neural network ensembles in probabilistic hydrological forecasting. Journal of Hydrology, 407(1–4), 94–104. https://doi.org/10.1016/J.JHYDROL.2011.07.011

Aslam, B., Maqsoom, A., Cheema, A. H., Ullah, F., Alharbi, A., & Imran, M. (2022). Water Quality Management Using Hybrid Machine Learning and Data Mining Algorithms: An Indexing Approach. IEEE Access, 10, 119692–119705. https://doi.org/10.1109/ACCESS.2022.3221430

Azaria, N. (2022, June 6). Feature importance: 7 methods and a quick tutorial. Aporia. https://www.aporia.com/learn/feature-importance/feature-importance-7-methods-and-a-quick-tutorial/

Bahrami, E., Salarijazi, M., & Nejatian, S. (2022). Estimation of flood hydrographs in the ungauged mountainous watershed with Gray synthetic unit hydrograph model. Arabian Journal of Geosciences 2022 15:8, 15(8), 1–10. https://doi.org/10.1007/S12517-022-10029-1

Bai, Y., Zhang, Z., & Zhao, W. (2019). Assessing the Impact of Climate Change on Flood Events Using HEC-HMS and CMIP%5. Water, Air, and Soil Pollution, 230(6), 1–13. https://doi.org/10.1007/S11270-019-4159-0/METRICS

Bakhshipour, A. E., Dittmer, U., Haghighi, A., & Nowak, W. (2019). Hybrid green-blue-gray decentralized urban drainage systems design, a simulation-optimization framework. Journal of Environmental Management, 249, 109364. https://doi.org/10.1016/J.JENVMAN.2019.109364

Barua, A., Ahmed, M. U., & Begum, S. (2023). A Systematic Literature Review on Multimodal Machine Learning: Applications, Challenges, Gaps and Future Directions. In IEEE Access (Vol. 11, pp. 14804–14831). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2023.3243854

Batina, L., Bhasin, S., Jap, D., & Picek, S. (2019). CSI NN: Reverse Engineering of Neural Network Architectures Through Electromagnetic Side Channel. USENIX Security Symposium. https://api.semanticscholar.org/CorpusID:199558279

Bbeiman, L. (1996). Bagging Predictors. 24, 123–140.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data 2018 5:1*, *5*(1), 1–12. https://doi.org/10.1038/sdata.2018.214

Benameur, S., Benkhaled, A., Meraghni, D., Chebana, F., & Necir, A. (2017). Complete flood frequency analysis in Abiod watershed, Biskra (Algeria). Natural Hazards, 86(2), 519–534. https://doi.org/10.1007/S11069-016-2703-4/METRICS

Bentley Systems. (2021). StormCAD. https://www.bentley.com/en/products/brands/stormcad

Benzater, B., Elouissi, A., Dabanli, I., Benaricha, B., & Hamimed, A. (2021). Extreme rain trend analysis in Macta watershed North West Algeria. Arabian Journal of Geosciences, 14(4), 1–14. https://doi.org/10.1007/S12517-021-06636-Z/METRICS

Berhail, S., & Katipoğlu, O. M. (2024). Exploring hydrological and meteorological drought trends in Northeast Algeria: implications for water resource management. Theoretical and Applied Climatology, 155(11), 9689–9712. https://doi.org/10.1007/S00704-024-05207-Y/METRICS

Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. Hydrological Processes, 6(3), 279–298. https://doi.org/10.1002/HYP.3360060305

Bhasme, P., Vagadiya, J., & Bhatia, U. (2022). Enhancing predictive skills in physically-consistent way: Physics Informed Machine Learning for hydrological processes. Journal of Hydrology, 615, 128618. https://doi.org/10.1016/J.JHYDROL.2022.128618

Bin, Z. (2015). The Application of SWMM Model in the Urban Planning Study on Sponge City. Earth Sciences, 4(5), 205. https://doi.org/10.11648/J.EARTH.20150405.17

Boughandjioua, S., Laouacheria, F., & Azizi, N. (2023). Machine Learning Algorithms Investigation for Urban Drainage Decision Systems: Overview. 2023 International Conference on Decision Aid Sciences and Applications, DASA 2023, 306–313. https://doi.org/10.1109/DASA59624.2023.10286621

Boughandjioua, S., Laouacheria, F., & Azizi, N. (2024). Enhancing stormwater network overflow prediction: investigation of ensemble learning models. Acta Geophysica. https://doi.org/10.1007/S11600-024-01407-2

Bowes, B. D., Adams, S., Beling, P. A., Saliba, S. M., & Goodall, J. L. (2020). Mitigation of Flooding in Stormwater Systems Utilizing Imperfect Forecasting and Sensor Data with Deep Deterministic Policy Gradient Reinforcement Learning. https://doi.org/10.20944/PREPRINTS202010.0413.V1

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324/METRICS

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). Classification and regression trees. Classification and Regression Trees, 1–358. https://doi.org/10.1201/9781315139470/CLASSIFICATION-REGRESSION-TREES-LEO-BREIMAN-JEROME-FRIEDMAN-OLSHEN-CHARLES-STONE

Brown, G., Wyatt, J., Harris, R., & Yao, X. (2005). Diversity creation methods: a survey and categorisation. Information Fusion, 6(1), 5–20. https://doi.org/10.1016/J.INFFUS.2004.04.004

Cea, L., & Costabile, P. (2022). Flood Risk in Urban Areas: Modelling, Management and Adaptation to Climate Change. A Review. https://doi.org/10.3390/hydrology9030050

Chang, Y. H., Tseng, C. W., & Hsu, H. C. (2023). Predicting the Overflowing of Urban Personholes Based on Machine Learning Techniques. Water 2023, Vol. 15, Page 4100, 15(23), 4100. https://doi.org/10.3390/W15234100

Chatterjee, C., Förster, S., & Bronstert, A. (2008). Comparison of hydrodynamic models of different complexities to model floods with emergency storage areas. Hydrological Processes, 22(24), 4695–4709. https://doi.org/10.1002/HYP.7079

Chafjiri, A. S., Gheibi, M., Chahkandi, B., Eghbalian, H., Waclawek, S., Fathollahi-Fard, A. M., & Behzadian, K. (2024). Enhancing flood risk mitigation by advanced data-driven approach. *Heliyon*, *10*(18), e37758. https://doi.org/10.1016/J.HELIYON.2024.E37758

Chen, J., Li, Y., Zhang, C., Tian, Y., & Guo, Z. (2023). Urban Flooding Prediction Method Based on the Combination of LSTM Neural Network and Numerical Model. International Journal of Environmental Research and Public Health, 20(2). https://doi.org/10.3390/IJERPH20021043

Chen, W., Zhang, S., Li, R., & Shahabi, H. (2018). Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling. Science of the Total Environment, 644, 1006–1018. https://doi.org/10.1016/J.SCITOTENV.2018.06.389

Cheriguene, S., Azizi, N., Dey, N., Ashour, A. S., & Ziani, A. (2019). A new hybrid classifier selection model based on mRMR method and diversity measures. International Journal of Machine Learning and Cybernetics, 10(5), 1189–1204. https://doi.org/10.1007/S13042-018-0797-6/METRICS

CHI. (2022). PCSWMM. https://www.pc-swwm.com

Chiaburu, T., & Biessmann, F. (2024). Interpretable Time Series Models for Wastewater Modeling in Combined Sewer Overflows. http://arxiv.org/abs/2401.02465

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science, 7, 1–24. https://doi.org/10.7717/PEERJ-CS.623

Chow, V. T., Maidment, D. R., & Mays, L. W. (1988). Applied hydrology. McGraw-Hill.Climates to Travel. (2024). Algeria climate: Average weather, temperature, rain. https://www.climatestotravel.com/climate/algeria

Ciello, R. Del. (2020). COUNTRY REPORT ON ENERGY EFFICIENCY AND RENEWABLE ENERGY INVESTMENT CLIMATE ALGERIA. https://www.researchgate.net/publication/341900356

Cox, D. R. (1958). The Regression Analysis of Binary Sequences. Journal of the Royal Statistical Society: Series B (Methodological), 20(2), 215–232. https://doi.org/10.1111/J.2517-6161.1958.TB00292.X

Crowe, C. T., Elger, D. F., Williams, B. C., & Roberson, J. A. (2010). Engineering fluid mechanics (9th ed.). John Wiley & Sons.

Cruz, F. R. G., Binag, M. G., Ga, M. R. G., & Uy, F. A. A. (2018). Flood Prediction Using Multi-Layer Artificial Neural Network in Monitoring System with Rain Gauge, Water Level, Soil Moisture Sensors. IEEE Region 10 Annual International Conference, Proceedings/TENCON, 2018-October, 2499–2503. https://doi.org/10.1109/TENCON.2018.8650387

Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random Forests. In Ensemble Machine Learning (pp. 157–175). Springer New York. https://doi.org/10.1007/978-1-4419-9326-7_5

Dharmarathne, G., Waduge, A. O., Bogahawaththa, M., Rathnayake, U., & Meddage, D. P. P. (2024). Adapting cities to the surge: A comprehensive review of climate-induced urban flooding. *Results in Engineering*, *22*, 102123. https://doi.org/10.1016/J.RINENG.2024.102123

D’ambrosio, R., Longobardi, A., Balbo, A., & Rizzo, A. (2021). Hybrid Approach for Excess Stormwater Management: Combining Decentralized and Centralized Strategies for the Enhancement of Urban Flooding Resilience. Water 2021, Vol. 13, Page 3635, 13(24), 3635. https://doi.org/10.3390/W13243635

Dang, T. Q., Tran, B. H., Le, Q. N., Dang, T. D., Tanim, A. H., Pham, Q. B., Bui, V. H., Mai, S. T., Thanh, P. N., & Anh, D. T. (2024). Application of machine learning-based surrogate models for urban flood depth modeling in Ho Chi Minh City, Vietnam. Applied Soft Computing, 150, 111031. https://doi.org/10.1016/J.ASOC.2023.111031

de Oliveira-Júnior, J. F., Mendes, D., Szabo, S., Singh, S. K., Jamjareegulgarn, P., Cardoso, K. R. A., Bertalan, L., da Silva, M. V., da Rosa Ferraz Jardim, A. M., da Silva, J. L. B., Lyra, G. B., Abreu, M. C., Filho, W. L. F. C., de Sousa, A., de Barros Santiago, D., da Silva Santos, I. G., & Maksudovna, V. K. (2024). Impact of the El Niño on Fire Dynamics on the African Continent. Earth Systems and Environment, 8(1), 45–61. https://doi.org/10.1007/S41748-023-00363-Z

Deltares. (2021). SOBEK. https://www.deltares.nl/en/software/sobek/

Dept., I. M. Fund. M. E. and C. A. (2024). Algeria: Selected Issues. IMF Staff Country Reports, 2024(089). https://doi.org/10.5089/9798400271922.002.A001

Devia, G. K., Ganasri, B. P., & Dwarakish, G. S. (2015). A Review on Hydrological Models. Aquatic Procedia, 4, 1001–1007. https://doi.org/10.1016/J.AQPRO.2015.02.126

DHI. (2022). MIKE+ Collection System. https://manuals.mikepoweredbydhi.help/2022/Cities/MIKE_Plus_Collection_System.pdf. Accessed May 27, 2024.

DHI. (2023). MIKE+ 2D Overland. https://manuals.mikepoweredbydhi.help/latest/Cities/MIKE_Plus_2DOverland.pdf. Accessed May 27, 2024.

Di Nardo, A., Boccelli, D. L., Herrera, M., Creaco, E., Cominola, A., Sitzenfrei, R., & Taormina, R. (2021). Smart Urban Water Networks: Solutions, Trends and Challenges. Water, 13(4), 501. https://doi.org/10.3390/W13040501

Dietterich, T. G. (2000). Ensemble Methods in Machine Learning. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 1857 LNCS, 1–15. https://doi.org/10.1007/3-540-45014-9_1

Diogo, A. F., & do Carmo, J. A. (2019). Peak flows and stormwater networks design-current and future management of urban surface watersheds. Water (Switzerland), 11(4). https://doi.org/10.3390/w11040759

Djordjević, S., Butler, D., Gourbesville, P., Mark, O., & Pasche, E. (2011). New policies to deal with climate change and other drivers impacting on resilience to flooding in urban areas: The CORFU approach. Environmental Science and Policy, 14(7), 864–873. https://doi.org/10.1016/J.ENVSCI.2011.05.008

Dottori, F., Salamon, P., Bianchi, A., Alfieri, L., Hirpa, F. A., & Feyen, L. (2016). Development and evaluation of a framework for global flood hazard mapping. Advances in Water Resources, 94, 87–102. https://doi.org/10.1016/J.ADVWATRES.2016.05.002

Dtissibe, F. Y., Ari, A. A. A., Titouna, C., Thiare, O., & Gueroui, A. M. (2020). Flood forecasting based on an artificial neural network scheme. Natural Hazards, 104(2), 1211–1237. https://doi.org/10.1007/S11069-020-04211-5

Ducharne, A. (2007). Changement climatique et ressources en eau. 10–11.

Duncker, L. (2019). Sanitation and climate change adaptation. Green Building Handbook, 12, 86–99.

EM-DAT, Centre for Research on the Epidemiology of Disasters. (2015). EM-DAT: The international disaster database. https://www.emdat.be/

Eslamian, S., Huda, M. B., Rather, N. A., & Eslamian, F. A. (2024). Handbook of Climate Change Impacts on River Basin Management. *Handbook of Climate Change Impacts on River Basin Management*. https://doi.org/10.1201/9781003473336

Fredholm, B. B. (1985). On the mechanism of action of theophylline and caffeine. Acta Medica Scandinavica, 217(1), 149-153. https://doi.org/10.1111/j.0954-6820.1944.tb06850.x

Friedman, J. H. (1999). Greedy Function Approximation: A Gradient Boosting Machine. http://www.salford-systems.com/treenet.html

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5), 1189–1232. https://doi.org/10.1214/AOS/1013203451

Fu, G., Jin, Y., Sun, S., Yuan, Z., & Butler, D. (2022). The role of deep learning in urban water management: A critical review. Water Research, 223, 118973. https://doi.org/10.1016/J.WATRES.2022.118973

Game, P., Wang, M., Audra, P., & Gourbesville, P. (2023). Flood modelling for a real-time decision support system of the covered Lower Paillons River, Nice, France. Journal of Hydroinformatics, 25(5), 1884–1908. https://doi.org/10.2166/HYDRO.2023.181/1303201/JH2023181.PDF

Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J., & Hochreiter, S. (2020). Rainfall-Runoff Prediction at Multiple Timescales with a Single Long Short-Term Memory Network. https://doi.org/10.5194/hess-25-2045-2021

GeeksforGeeks. (2023). Cross-validation in machine learning. GeeksforGeeks. https://www.geeksforgeeks.org/cross-validation-machine-learning/

GeeksforGeeks. (2024). *Machine learning - Linear regression*. GeeksforGeeks https://www.geeksforgeeks.org/ml-linear-regression/

Greenacre, M., Groenen, P. J. F., Hastie, T., D’Enza, A. I., Markos, A., & Tuzhilina, E. (2022). Principal component analysis. Nature Reviews Methods Primers 2022 2:1, 2(1), 1–21. https://doi.org/10.1038/s43586-022-00184-w

Grum, M., Jørgensen, A. T., Johansen, R. M., & Linde, J. J. (2006). The effect of climate change on urban drainage: an evaluation based on regional climate model simulations. Water Science and Technology, 54(6–7), 9–15. https://doi.org/10.2166/WST.2006.592

Gu, X., Zhang, Q., Li, J., Singh, V. P., & Sun, P. (2019). Impact of urbanization on nonstationarity of annual and seasonal precipitation extremes in China. Journal of Hydrology, 575, 638–655. https://doi.org/10.1016/J.JHYDROL.2019.05.070

Güneralp, B., Güneralp, I., & Liu, Y. (2015). Changing global patterns of urban exposure to flood and drought hazards. Global Environmental Change, 31, 217–225. https://doi.org/10.1016/J.GLOENVCHA.2015.01.002

Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1999). Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. Journal of Hydrologic Engineering, 4(2), 135–143. https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135)

Gupta, S., Kumar, D., Negi, R., Singh, R., Ansari, M. J., Kashyap, S., & Mehrotra, S. (2024). Genomic Intelligence. https://doi.org/10.1201/9781003570233

Guresen, E., & Kayakutlu, G. (2011). Definition of artificial neural networks with comparison to other networks. Procedia Computer Science, 3, 426–433. https://doi.org/10.1016/J.PROCS.2010.12.071

Hachemi, A., & Benkhaled, A. (2016). Flood- duration-frequency modeling application to wadi abiodh, biskra (algeria). Larhyss Journal, 13(3), 277–297. https://www.asjp.cerist.dz/en/article/55365

Hafnaoui, M. A., & Dabanlı, İ. (2023). FLOODS IN ALGERIA: ANALYZES AND STATISTICS. https://www.researchgate.net/publication/377002510

Hamitouche, Y., Zeroual, A., Meddi, M., Assani, A. A., Alkama, R., Şen, Z., & Zhang, X. (2024). Projected Changes in Extreme Precipitation Patterns across Algerian Sub-Regions. Water (Switzerland), 16(10), 1353. https://doi.org/10.3390/W16101353/S1

Han, J., Kamber, M., & Pei, J. (2011). Data mining: Concepts and techniques (3rd ed.). Morgan Kaufmann.

Hassan, B. T., Yassine, M., & Amin, D. (2022). Comparison of Urbanization, Climate Change, and Drainage Design Impacts on Urban Flashfloods in an Arid Region: Case Study, New Cairo, Egypt. Water 2022, Vol. 14, Page 2430, 14(15), 2430. https://doi.org/10.3390/W14152430

Heaton, J. (2018). Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning: The MIT Press, 2016, 800 pp, ISBN: 0262035618 Article. Genetic Programming and Evolvable Machines, 19(1–2), 305–307.

Hossain, S., Hewa, G. A., & Wella-Hewage, S. (2019). A Comparison of Continuous and Event-Based Rainfall–Runoff (RR) Modelling Using EPA-SWMM. Water 2019, Vol. 11, Page 611, 11(3), 611. https://doi.org/10.3390/W11030611

Houta, D., & Gadiyatov, V. G. (2023). Climatic Factors of the Algerian Coast. Springer Geology, 93–105. https://doi.org/10.1007/978-3-031-16575-7_10

Huang, D., Liu, Z., & Wu, D. (2024). Research on Ensemble Learning-Based Feature Selection Method for Time-Series Prediction. Applied Sciences (Switzerland), 14(1). https://doi.org/10.3390/app14010040

Hukkeri, G. S., Naganna, S. R., Pruthviraja, D., Bhat, N., & Goudar, R. H. (2023). Drought Forecasting: Application of Ensemble and Advanced Machine Learning Approaches. IEEE Access, 11, 141375–141393. https://doi.org/10.1109/ACCESS.2023.3341587

Innovyze. (2020). InfoWorks ICM. https://www.innovyze.com/products/infoworks-icm

International Federation of Red Cross and Red Crescent Societies. (2024). IFRC Tunisia DREF MDRDZ010. https://www.ifrc.org/document/ifrc-dref-brochure-2024

Islam, M. S., & Rony, M. A. T. (2024). Navigating the global stock market: correlation, prediction, and the influence of external factors. Iran Journal of Computer Science 2024 7:3, 7(3), 397–422. https://doi.org/10.1007/S42044-024-00177-W

IPCC. (2014). Climate change 2014: Synthesis report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

IPCC. (2018). Global warming of 1.5 °C: An IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways. https://www.ipcc.ch/sr15/

Iradukunda, P., Mwanaumo, E. M., & Kabika, J. (2024). Modelling the future climate impacts on hydraulic infrastructure development in tropical (peri-)urban region: Case of Kigali, Rwanda. Heliyon, 10(5). https://doi.org/10.1016/J.HELIYON.2024.E27126

Irvin, V., & Nakashima, S. (2023). The Query Search Method. Proceedings of the ALISE Annual Conference. https://doi.org/10.21900/J.ALISE.2023.1281

Jagaba, A. H., Kutty, S. R. M., Hayder, G., Baloo, L., Noor, A., Yaro, N. S. A., Saeed, A. A. H., Lawal, I. M., Birniwa, A. H., & Usman, A. K. (2021). A Systematic Literature Review on Waste-to-Resource Potential of Palm Oil Clinker for Sustainable Engineering and Environmental Applications. Materials 2021, Vol. 14, Page 4456, 14(16), 4456. https://doi.org/10.3390/MA14164456

Jehanzaib, M., Shah, S. A., Yoo, J., & Kim, T. W. (2020). Investigating the impacts of climate change and human activities on hydrological drought using non-stationary approaches. Journal of Hydrology, 588, 125052. https://doi.org/10.1016/J.JHYDROL.2020.125052

Jiang, S., Sarica, S., Song, B., Hu, J., & Luo, J. (2021). Patent Data for Engineering Design: A Critical Review and Future Directions. Journal of Computing and Information Science in Engineering, 22(6). https://doi.org/10.1115/1.4054802

Jianping, L., Minrong, L., Jinnan, W., Jianjian, L., Hongwen, S., & Maoxing, H. (2014). Report on Global Environment Competitiveness of Algeria. 309–312. https://doi.org/10.1007/978-3-642-54678-5_17

Johnson, S. C. (1967). Hierarchical clustering schemes. Psychometrika, 32(3), 241–254. https://doi.org/10.1007/BF02289588/METRICS

Kan, G., Liang, K., Yu, H., Sun, B., Ding, L., Li, J., He, X., & Shen, C. (2020). Hybrid machine learning hydrological model for flood forecast purpose. Open Geosciences, 12(1), 813–820. https://doi.org/10.1515/GEO-2020-0166/HTML

Kang, N., Kim, S., Kim, Y., Noh, H., Hong, S. J., & Kim, H. S. (2016). Urban Drainage System Improvement for Climate Change Adaptation. Water 2016, Vol. 8, Page 268, 8(7), 268. https://doi.org/10.3390/W8070268

Kaspersen, P. S., & Halsnæs, K. (2017). Integrated climate change risk assessment: A practical application for urban flooding during extreme precipitation. Climate Services, 6, 55–64. https://doi.org/10.1016/J.CLISER.2017.06.012

Ke, Q., Tian, X., Bricker, J., Tian, Z., Guan, G., Cai, H., Huang, X., Yang, H., & Liu, J. (2020). Urban pluvial flooding prediction by machine learning approaches – a case study of Shenzhen city, China. Advances in Water Resources, 145, 103719. https://doi.org/10.1016/J.ADVWATRES.2020.103719

Ketrouci, K., Meddi, M., & Abdesselam, B. (2012). Étude des crues extrêmes en Algérie : cas du bassin-versant de la Tafna. Science et Changements Planetaires - Secheresse, 23(4), 297–305. https://doi.org/10.1684/SEC.2012.0353

Keum, H. J., Han, K. Y., & Kim, H. Il. (2020). Real-Time Flood Disaster Prediction System by Applying Machine Learning Technique. Ksce Journal of Civil Engineering, 24(9), 2835–2848. https://doi.org/10.1007/S12205-020-1677-7

Khan, K., Ahmad, W., Amin, M. N., & Ahmad, A. (2022). A Systematic Review of the Research Development on the Application of Machine Learning for Concrete. Materials, 15(13). https://doi.org/10.3390/MA15134512

Kolena. (2024). Feature importance: Methods, tools, and best practices. https://www.kolena.com/guides/feature-importance-methods-tools-and-best-practices/

Kourtis, I. M., & Tsihrintzis, V. A. (2021). Adaptation of urban drainage networks to climate change: A review. Science of The Total Environment, 771, 145431. https://doi.org/10.1016/J.SCITOTENV.2021.145431

Kumar, K. S. R., & Biradar, R. V. (2022). An Intelligent Flood Forecasting System Using Artificial Neural Network in WSN. 279–289. https://doi.org/10.1007/978-981-16-3346-1_23

Kuncheva, L. I., & Whitaker, C. J. (2003). Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. Machine Learning, 51(2), 181–207. https://doi.org/10.1023/A:1022859003006/METRICS

Kushawaha, J., Borra, S., Kushawaha, A. K., Singh, G., & Singh, P. (2021). Climate change and its impact on natural resources. Water Conservation in the Era of Global Climate Change, January, 333–346. https://doi.org/10.1016/B978-0-12-820200-5.00002-6

Kwon, S. H., & Kim, J. H. (2021). Machine learning and urban drainage systems: State-of-the-art review. Water (Switzerland), 13(24), 1–14. https://doi.org/10.3390/w13243545

Lange, H., & Sippel, S. (2020). Machine Learning Applications in Hydrology. 233–257. https://doi.org/10.1007/978-3-030-26086-6_10

Laouacheria, F., & Mansouri, R. (2015). Comparison of WBNM and HEC-HMS for Runoff Hydrograph Prediction in a Small Urban Catchment. Water Resources Management, 29(8), 2485–2501. https://doi.org/10.1007/s11269-015-0953-7

Laouacheria, F., Kechida, S., & Chabi, M. (2019). Modelling the impact of design rainfall on the urban drainage system by Storm Water Management Model. Journal of Water and Land Development, 40(1), 119–125. https://doi.org/10.2478/jwld-2019-0013

Le, X. H., Ho, H. V., Lee, G., & Jung, S. (2019). Application of Long Short-Term Memory (LSTM) neural network for flood forecasting. Water (Switzerland), 11(7). https://doi.org/10.3390/w11071387

Leal Filho, W., Sierra, J., Kalembo, F., Ayal, D. Y., Matandirotya, N., de Victoria Pereira Amaro da Costa, C. I., Sow, B. L., Aabeyir, R., Mawanda, J., Zhou, L., & Baldeh, D. (2024). The role of African universities in handling climate change. Environmental Sciences Europe, 36(1), 1–16. https://doi.org/10.1186/S12302-024-00933-6/TABLES/2

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature 2015 521:7553, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Lee, S., Kim, J. C., Jung, H. S., Lee, M. J., & Lee, S. (2017). Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. Geomatics, Natural Hazards and Risk, 8(2), 1185–1203. https://doi.org/10.1080/19475705.2017.1308971

Li, H., Wen, G., Yu, Z., & Zhou, T. (2013). Random subspace evidence classifier. Neurocomputing, 110, 62–69. https://doi.org/10.1016/J.NEUCOM.2012.11.019

Li, S., Ma, K., Jin, Z., & Zhu, Y. (2016). A new flood forecasting model based on SVM and boosting learning algorithms. 2016 IEEE Congress on Evolutionary Computation, CEC 2016, 1343–1348. https://doi.org/10.1109/CEC.2016.7743944

Lind, J. (2015). UPTEC W 14040 Examensarbete 30 hp Mars.

Liu, T., Ramirez-Marquez, J. E., Jagupilla, S. C., & Prigiobbe, V. (2021). Combining a statistical model with machine learning to predict groundwater flooding (or infiltration) into sewer networks. Journal of Hydrology, 603, 126916. https://doi.org/10.1016/J.JHYDROL.2021.126916

Liu, W., Fu, J., Liang, Y., Cao, M., & Han, X. (2022). A Well-Overflow Prediction Algorithm Based on Semi-Supervised Learning. Energies, 15(12), 4324. https://doi.org/10.3390/EN15124324/S1

Liu, X., Liu, W., Tang, Q., Liu, B., Wada, Y., & Yang, H. (2022). Global Agricultural Water Scarcity Assessment Incorporating Blue and Green Water Availability Under Future Climate Change. Earth's Future, 10(4). https://doi.org/10.1029/2021EF002567

Lu, M., Hou, Q., Qin, S., Zhou, L., Hua, D., Wang, X., & Cheng, L. (2023). A Stacking Ensemble Model of Various Machine Learning Models for Daily Runoff Forecasting. Water 2023, Vol. 15, Page 1265, 15(7), 1265. https://doi.org/10.3390/W15071265

Lundy, L., Blecken, G., Österlund, H., & Viklander, M. (2012). Skandinaviens nordligaste tekniska universitet Forskning & utbildning i världsklass Systematic review of urban stormwater research in Sweden.

Ma, S., Zayed, T., Xing, J., & Shao, Y. (2024). A state-of-the-art review for the prediction of overflow in urban sewer systems. Journal of Cleaner Production, 434, 139923. https://doi.org/10.1016/J.JCLEPRO.2023.139923

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. Https://Doi.Org/, 5.1, 281–298. https://projecteuclid.org/ebooks/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Fifth-Berkeley-Symposium-on-Mathematical-Statistics-and/chapter/Some-methods-for-classification-and-analysis-of-multivariate-observations/bsmsp/1200512992

Maidment, D. R. (1993). Handbook of Hydrology. McGraw-Hill. https://dl.watereng.ir/HANDBOOK_OF_HYDROLOGY.PDF

Marie, D., Jensen, R., Sandoval, S., Bertrand-Krajewski, J.-L., Steen Mikkelsen, P., Vezzaro, L., Steen, P., Jensen, D. M. R., Sandoval, S., Bertrand-Krajewski, J.-L., Mikkelsen, P. S., & Vezzaro, L. (2022). First flush assumptions in pollution management for separate stormwater systems: debunking a classical urban (drainage) myth. https://hal.science/hal-04132767

Meijer, D., Korving, H., Langeveld, J., & Clemens-Meyer, F. (2023). A method to identify the weakest link in urban drainage systems. Water Science and Technology, 87 5(5), 1273–1293. https://doi.org/10.2166/WST.2023.057

Menad, W., Douvinet, J., Beltrando, G., & Arnaud-Fassetta, G. (2012). Evaluating the influence of urbanised areas face to a remarkable meteorological hazard: The flood event of 9-10 November 2001 at Bab-el-Oued (Alger, Algeria). Géomorphologie: Relief, Processus, Environnement, 18(3), 337-350. (In French)

Miloudi, Y., Labbaci, B., Fezzioui, N., Oumeziane, Y. A., Benidir, A., & Bennaceur, S. (2021). Temperature effects on sorption behaviour of compressed and cement-stabilised earth blocks. Magazine of Concrete Research, 73(9), 476–486. https://doi.org/10.1680/JMACR.19.00169

Mitchell, T. M. (2010). CHAPTER 1 GENERATIVE AND DISCRIMINATIVE CLASSIFIERS : NAIVE BAYES AND LOGISTIC REGRESSION Learning Classifiers based on Bayes Rule. Machine Learning, 1(Pt 1-2), 1–17. https://doi.org/10.1093/bioinformatics/btq112

Mitchell, T. M. . (1997). Machine learning. McGraw-Hill.

Mitchell, V. G., Mein, R. G., & McMahon, T. A. (2001). Modelling the urban water cycle. Environmental Modelling & Software, 16(7), 615–629. https://doi.org/10.1016/S1364-8152(01)00029-9

Moradkhani, H., & Sorooshian, S. (2009). General Review of Rainfall-Runoff Modeling: Model Calibration, Data Assimilation, and Uncertainty Analysis. Hydrological Modelling and the Water Cycle, 1–24. https://doi.org/10.1007/978-3-540-77843-1_1

Motta, M., de Castro Neto, M., & Sarmento, P. (2021). A mixed approach for urban flood prediction using Machine Learning and GIS. International Journal of Disaster Risk Reduction, 56, 102154. https://doi.org/10.1016/j.ijdrr.2021.102154

Mousavi, F. S., Yousefi, S., Abghari, H., & Ghasemzadeh, A. (2021). Design of an IoT-based Flood Early Detection System using Machine Learning. 26th International Computer Conference, Computer Society of Iran, CSICC 2021, 1–7. https://doi.org/10.1109/CSICC52343.2021.9420594

Mujumdar, P. P. (2001). Flood wave propagation. Resonance 2001 6:5, 6(5), 66–73. https://doi.org/10.1007/BF02839085

Mukaka, M. M. (2012). A guide to appropriate use of Correlation coefficient in medical research. Malawi Medical Journal : The Journal of Medical Association of Malawi, 24(3), 69. /pmc/articles/PMC3576830/

Mullapudi, A., Lewis, M. J., Gruden, C. L., & Kerkez, B. (2020a). Deep reinforcement learning for the real time control of stormwater systems. Advances in Water Resources, 140, 103600. https://doi.org/10.1016/J.ADVWATRES.2020.103600

Mullapudi, A., Lewis, M. J., Gruden, C. L., & Kerkez, B. (2020b). Deep reinforcement learning for the real time control of stormwater systems. Advances in Water Resources, 140, 103600. https://doi.org/10.1016/J.ADVWATRES.2020.103600

Mullapudi, A., Lewis, M. J., Gruden, C. L., & Kerkez, B. (2020c). Deep reinforcement learning for the real time control of stormwater systems. Advances in Water Resources, 140, 103600. https://doi.org/10.1016/J.ADVWATRES.2020.103600

Nanda, M. K., Chandra, B., Viswavidyalaya, K., Kohli, R. K., Garg, V. K., & Dhawan, A. (2018). Climatic Classification Environmental Sciences Atmospheric Processes Climatic Classification Paper No: 8 Atmospheric Processes Module: 24 Climatic Classification Development Team Principal Investigator & Co-Principal Investigator. https://www.researchgate.net/publication/325130734

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. Frontiers in Neurorobotics, 7(DEC), 63623. https://doi.org/10.3389/FNBOT.2013.00021/BIBTEX

Noymanee, J., & Theeramunkong, T. (2019). Flood Forecasting with Machine Learning Technique on Hydrological Modeling. Procedia Computer Science, 156, 377–386. https://doi.org/10.1016/J.PROCS.2019.08.214

Ogallo, L. O. (2009). Climate variability and change in Africa: a review of potential impacts on terrestrial water resources. http://erepository.uonbi.ac.ke/handle/11295/44317

Okoro, C. S. (2023). Sustainable Facilities Management in the Built Environment: A Mixed-Method Review. Sustainability 2023, Vol. 15, Page 3174, 15(4), 3174. https://doi.org/10.3390/SU15043174

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. Hydrology and Earth System Sciences, 11(5), 1633–1644. https://doi.org/10.5194/HESS-11-1633-2007

Pham, B. T., Jaafari, A., Phong, T. Van, Yen, H. P. H., Tuyen, T. T., Luong, V. Van, Nguyen, H. D., Le, H. Van, & Foong, L. K. (2021). Improved flood susceptibility mapping using a best first decision tree integrated with ensemble learning techniques. Geoscience Frontiers, 12(3). https://doi.org/10.1016/J.GSF.2020.11.003

Polikar, R. (2009). Ensemble learning. Scholarpedia, 4(1), 2776. https://doi.org/10.4249/SCHOLARPEDIA.2776

Power & Motion. (2024). Climate change and what it means for your hydraulic equipment. Power & Motion. https://www.powermotiontech.com/hydraulics-at-work/article/21884596/climate-change-and-what-it-means-for-your-hydraulic-equipment

Prakash, G., Rao, G. V., Pratap, D., & Kumar Gupta, P. (2021). Flood Inundation Mapping and Depth Modelling using Machine Learning algorithms and Microwave data. In Journal of Geomatics (Vol. 15, Issue 2). https://www.researchgate.net/publication/357536008

Rajaee, T., Khani, S., & Ravansalar, M. (2020). Artificial intelligence-based single and hybrid models for prediction of water quality in rivers: A review. Chemometrics and Intelligent Laboratory Systems, 200, 103978. https://doi.org/10.1016/J.CHEMOLAB.2020.103978

Ramovha, N., Chadyiwa, M., Ntuli, F., & Sithole, T. (2024). The Potential of Stormwater Management Strategies and Artificial Intelligence Modeling Tools to Improve Water Quality: A Review. Water Resources Management, 38(10), 3527–3560. https://doi.org/10.1007/S11269-024-03841-9/FIGURES/3

Rjeily, Y. A., Abbas, O., Sadek, M., Shahrour, I., & Chehade, F. H. (2017). Flood forecasting within urban drainage systems using NARX neural network. Water Science and Technology, 76(9), 2401–2412. https://doi.org/10.2166/WST.2017.409

Rosin, T. R., Romano, M., Keedwell, E., & Kapelan, Z. (2021). A Committee Evolutionary Neural Network for the Prediction of Combined Sewer Overflows. Water Resources Management, 35(4), 1273–1289. https://doi.org/10.1007/S11269-021-02780-Z/TABLES/1

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature 1986 323:6088, 323(6088), 533–536. https://doi.org/10.1038/323533a0

Saheed, Y. K., & Misra, S. (2024). A voting gray wolf optimizer-based ensemble learning models for intrusion detection in the Internet of Things. International Journal of Information Security, 23(3), 1557–1581. https://doi.org/10.1007/S10207-023-00803-X/FIGURES/3

Sahnoune, F., Belhamel, M., Zelmat, M., & Kerbachi, R. (2013). Climate change in Algeria: Vulnerability and strategy of mitigation and adaptation. Energy Procedia, 36, 1286–1294. https://doi.org/10.1016/J.EGYPRO.2013.07.145

Saini, R., & Rawat, S. (2023). Land Use Land Cover Classification in Remote Sensing Using Machine Learning Techniques. 1st IEEE International Conference on Innovations in High Speed Communication and Signal Processing, IHCSP 2023, 99–104. https://doi.org/10.1109/IHCSP56702.2023.10127126

Salim, F., & Bhattacharyya, S. (2023). Ensemble learning bagging and boosting.

Sami, G., Abdelwahhab, F., Yahyaoui, H., & Abdelghani, F. (2021). Flood Hazard In The City Of Chemora (Algeria). Analele Universităţii Din Oradea, Seria Geografie, 31(1), 22–27. https://doi.org/10.30892/AUOG.311103-835

Sardou, M., Maouche, S., & Missoum, H. (2016). Compilation of historical floods catalog of northwestern Algeria: first step towards an atlas of extreme floods. Arabian Journal of Geosciences, 9(6), 1–15. https://doi.org/10.1007/S12517-016-2490-Y/METRICS

Schapire, R. E. (2002). The Boosting Approach to Machine Learning An Overview. In MSRI Workshop on Nonlinear Estimation and Classification. www.research.att.com/

Schilling, J., Hertig, E., Tramblay, Y., & Scheffran, J. (2020). Climate change vulnerability, water resources and social implications in North Africa. Regional Environmental Change, 20(1). https://doi.org/10.1007/S10113-020-01597-7

Shafizadeh-Moghadam, H., Valavi, R., Shahabi, H., Chapi, K., & Shirzadi, A. (2018). Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. Journal of Environmental Management, 217, 1–11. https://doi.org/10.1016/J.JENVMAN.2018.03.089

Shahed Behrouz, M., Zhu, Z., Matott, L. S., & Rabideau, A. J. (2020). A new tool for automatic calibration of the Storm Water Management Model (SWMM). Journal of Hydrology, 581, 124436. https://doi.org/10.1016/J.JHYDROL.2019.124436

Shamsi, U. M. . (2005). GIS applications for water, wastewater, and stormwater systems. Taylor & Francis. https://www.routledge.com/GIS-Applications-for-Water-Wastewater-and-Stormwater-Systems/Shamsi/p/book/9780849320972

Shin, T. (2023, November 7). Understanding feature importance in machine learning. Built In. https://builtin.com/data-science/feature-importance

Shu, C., & Burn, D. H. (2004). Artificial neural network ensembles and their application in pooled flood frequency analysis. Water Resources Research, 40(9), 1821. https://doi.org/10.1029/2003WR002816

Sifa, S. F., Nila, A. A., Hossain, A., & Hayat, T. (2023). Rainfall-Runoff Estimation in the North-Eastern Region of Bangladesh Using SCS-CN Method. The Dhaka University Journal of Earth and Environmental Sciences, 12(2), 83–96. https://doi.org/10.3329/DUJEES.V12I2.73165

Solomatine, D. P., & Wagener, T. (2011). Hydrological Modeling. Treatise on Water Science, 2, 435–457. https://doi.org/10.1016/B978-0-444-53199-5.00044-0

Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction Second edition, in progress.

Szeląg, B., Bąk, Ł., Suligowski, R., & Górski, J. (2018). Statistical models to predict discharge overflow. Water Science and Technology, 78(5), 1208–1218. https://doi.org/10.2166/WST.2018.392

Thi, N. T. V. (2008). One dimensional Saint-Venant system. 45. https://dumas.ccsd.cnrs.fr/dumas-00597434

Tian, Y., Shu, M., & Jia, Q. (2020). Artificial Neural Network. In B. S. Daya Sagar, Q. Cheng, J. McKinley, & F. Agterberg (Eds.), Encyclopedia of Mathematical Geosciences (pp. 1–4). Springer International Publishing. https://doi.org/10.1007/978-3-030-26050-7_44-1

Tina Moni Boruah, M. (2021). Development of 1d and 2d Hydrodynamic Flood Model Using Mike Software by Dhi". International Journal of Advances in Engineering and Management (IJAEM), 3, 61. https://doi.org/10.35629/5252-03086178

Tiwari, M. K., & Chatterjee, C. (2010). Development of an accurate and reliable hourly flood forecasting model using wavelet-bootstrap-ANN (WBANN) hybrid approach. Journal of Hydrology, 394(3–4), 458–470. https://doi.org/10.1016/J.JHYDROL.2010.10.001

Todini, E. (2007). Hydrological catchment modelling: past, present and future. Hydrology and Earth System Sciences, 11(1), 468–482. https://doi.org/10.5194/HESS-11-468-2007

Torres, F. L. R., Lima, L. M. M., Reboita, M. S., de Queiroz, A. R., & Lima, J. W. M. (2024). Integrating Hydrological and Machine Learning Models for Enhanced Streamflow Forecasting via Bayesian Model Averaging in a Hydro-Dominant Power System. Water 2024, Vol. 16, Page 586, 16(4), 586. https://doi.org/10.3390/W16040586

Tramblay, Y., Ribes, A., Somot, S., Neppel, L., Lucas-Picher, P., Vinet, F., & Sauquet, E. (2021). Impacts du changement climatique sur les pluies intenses et les crues en Méditerranée. LHB: Hydroscience Journal, 107(1), 1–5. https://doi.org/10.1080/00186368.2021.1912971

Tran, H. D., Molavi, S., & Muttil, N. (2011). Assessment Framework for the Impacts of Climate Change and Urbanization on Urban Drainage Systems. Pipelines 2011: A Sound Conduit for Sharing Solutions - Proceedings of the Pipelines 2011 Conference, 1403–1412. https://doi.org/10.1061/41187(420)129

Uludağ, A. S. (2022). Deep image compression with a unified spatial and channel context auto-regressive model. https://open.metu.edu.tr/handle/11511/99476

United States Environmental Protection Agency. (2019). SWMM 5. https://www.epa.gov/water-research/storm-water-management-model-swmm

Van Der Maaten, L., & Hinton, G. (2008). Visualizing Data using t-SNE. In Journal of Machine Learning Research (Vol. 9).

Vanpoucke, D. E. P., Van Knippenberg, O. S. J., Hermans, K., Bernaerts, K. V., & Mehrkanoon, S. (2020). Small data materials design with machine learning: When the average model knows best. Journal of Applied Physics, 128(5). https://doi.org/10.1063/5.0012285/1063225

Venkatesan, E., & Mahindrakar, A. B. (2019). Cite this Article: Ellakkia Venkatesan and Amit B. Mahindrakar, Forecasting Floods using Extreme Gradient Boosting A New Approach.

International Journal of Civil Engineering and Technology, 10(2), 1336–1346. http://iaeme.com/Home/issue/IJCIET?Volume=10&Issue=2http://iaeme.com

Vidyapriya, V., & Ramalingam, M. (2012). Flood Modelling using MIKE URBAN Software: An Application to Flood Modelling using MIKE URBAN Software: An Application to Jafferkhanpet watershed. International Journal of Pollution Abatement Technology, 1(1), 20–24. www.ijpat.info

Vu, M. T., Raghavan, S. V., Liu, J., & Liong, S. Y. (2018). Constructing short-duration IDF curves using coupled dynamical–statistical approach to assess climate change impacts. International Journal of Climatology, 38(6), 2662–2671. https://doi.org/10.1002/JOC.5451

Walczak, S., & Cerpa, N. (2003). Artificial Neural Networks. Encyclopedia of Physical Science and Technology, 631–645. https://doi.org/10.1016/B0-12-227410-5/00837-1

Wang, H., & Song, L. (2020). Water Level Prediction of Rainwater Pipe Network Using an SVM-Based Machine Learning Method. International Journal of Pattern Recognition and Artificial Intelligence, 34(02), 2051002. https://doi.org/10.1142/S0218001420510027

Wang, H., Song, L., Wang, H., & Song, L. (2019). Accepted manuscript to appear in IJPRAI Accepted Manuscript International Journal of Pattern Recognition and Artificial Intelli- gence. https://doi.org/10.1142/S0218001420510027

Wang, L., Cui, S., Li, Y., Huang, H., Manandhar, B., Nitivattananon, V., Fang, X., & Huang, W. (2022). A review of the flood management: from flood control to flood resilience. Heliyon, 8(11), e11763. https://doi.org/10.1016/J.HELIYON.2022.E11763

Wang, Y., Wen, J., Sun, G., & Zhang, W. (2019). Development of Intelligent Water Resources System Combined with Artificial Intelligence in Flood Forecasting. Advances in Intelligent Systems and Computing, 1084 AISC, 277–283. https://doi.org/10.1007/978-3-030-34387-3_34

Wang, Yeqiao. (2023). Handbook of natural resources. Volume 4, Fresh water and watersheds. CRC Press. https://www.routledge.com/Fresh-Water-and-Watersheds/Wang/p/book/9781032474397

Watson-Hernández, F., Guzmán-Arias, I., Chavarría-Pizarro, L., & Quesada-Alvarado, F. (2022). The Effect of Climate Change on the Water Supply and Hydraulic Conditions in the Upper Pejibaye River Basin, Cartago, Costa Rica. Hydrology 2022, Vol. 9, Page 76, 9(5), 76. https://doi.org/10.3390/HYDROLOGY9050076

World Bank. (2023, August 29). Learning from past shocks to scale up disaster and climate resilience in Algeria. World Bank Blogs. https://blogs.worldbank.org/en/arabvoices/learning-from-past-shocks-to-scale-up-disaster-and-climate-resilience-in-algeria

World Bank. (2024). Climate Knowledge Portal: Algeria - Climate data historical. https://climateknowledgeportal.worldbank.org/country/algeria/climate-data-historical

World Health Organization. (2015). Operational framework for building climate resilient health systems. https://iris.who.int/bitstream/handle/10665/246137/WHO-FWC-PHE-EPE-15.32-eng.pdf?sequence=1

Wu, Y., Ke, Y., Chen, Z., Liang, S., Zhao, H., & Hong, H. (2020). Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. CATENA, 187, 104396. https://doi.org/10.1016/J.CATENA.2019.104396

Xu, T., & Liang, F. (2021). Machine learning for hydrologic sciences: An introductory overview. In Wiley Interdisciplinary Reviews: Water (Vol. 8, Issue 5). John Wiley and Sons Inc. https://doi.org/10.1002/wat2.1533

Yahiaoui, A. (2012). Inondations torrentielles : cartographie des zones vulnérables en Algérie du nord (Cas de l'oued Mekerra, Wilaya de sidi Bel Abbès). http://repository.enp.edu.dz/jspui/handle/123456789/700

Yan, J., & Tao, T. (2022). Unsupervised anomaly detection in hourly water demand data using an asymmetric encoder–decoder model. Journal of Hydrology, 613, 128389. https://doi.org/10.1016/J.JHYDROL.2022.128389

Yang, Z. R., & Yang, Z. (2014). Artificial Neural Networks. Comprehensive Biomedical Physics, 6, 1–17. https://doi.org/10.1016/B978-0-444-53632-7.01101-1

Yin, Z., Zahedi, L., Leon, A. S., Hadi Amini, M., & Bian, L. (2022). A Machine Learning Framework for Overflow Prediction in Combined Sewer Systems. World Environmental and Water Resources Congress 2022: Adaptive Planning and Design in an Age of Risk and Uncertainty - Selected Papers from the World Environmental and Water Resources Congress 2022, 194–205. https://doi.org/10.1061/9780784484258.019

Yuan, W., Liu, M., & Wan, F. (2019). Calculation of Critical Rainfall for Small-Watershed Flash Floods Based on the HEC-HMS Hydrological Model. Water Resources Management, 33(7), 2555–2575. https://doi.org/10.1007/S11269-019-02257-0/METRICS

Zaidi, N., Douafer, L., & Hamdani, A. (2021). Diversity and abundance of terrestrial gastropods in Skikda region (North-East Algeria): correlation with soil physicochemical factors. The Journal of Basic and Applied Zoology, 82(1). https://doi.org/10.1186/s41936-021-00239-6

ZEGAIT, R., & PIZZO, H. S. (2023). Flood control reservoir using vba simulation case of idles basin in southern algeria. Larhyss Journal P-ISSN 1112-3680 / E-ISSN 2521-9782, 0(53), 41–60. http://larhyss.net/ojs/index.php/larhyss/article/view/4553

Zegait, R., Şen, Z., Pulido-Bosch, A., Madi, H., & Hamadeha, B. (2022). Flash Flood Risk and Climate Analysis in the Extreme South of Algeria (the Case of In-Guezzam City). Geomatics and Environmental Engineering, 16(4), 157–185. https://doi.org/10.7494/GEOM.2022.16.4.157

Zhang, M., Xu, Z., Wang, Y., Zeng, S., & Dong, X. (2022). Evaluation of uncertain signals' impact on deep reinforcement learning-based real-time control strategy of urban drainage systems. Journal of Environmental Management, 324, 116448. https://doi.org/10.1016/J.JENVMAN.2022.116448

Zhou, Z. H. (2012). Ensemble methods: Foundations and algorithms. Ensemble Methods: Foundations and Algorithms, 1–218. https://doi.org/10.1201/B12207

Zoppou, C. (2001). Review of urban storm water models. Environmental Modelling & Software, 16(3), 195–231. https://doi.org/10.1016/S1364-8152(00)00084-0

Zounemat-Kermani, M. (2017). Assessment of several nonlinear methods in forecasting suspended sediment concentration in streams. Hydrology Research, 48(5), 1240–1252. https://doi.org/10.2166/NH.2016.219

Zounemat-Kermani, M., & Scholz, M. (2013). Computing Air Demand Using the Takagi–Sugeno Model for Dam Outlets. Water 2013, Vol. 5, Pages 1441-1456, 5(3), 1441–1456. https://doi.org/10.3390/W5031441

Zounemat-Kermani, M., Batelaan, O., Fadaee, M., & Hinkelmann, R. (2021). Ensemble machine learning paradigms in hydrology: A review. Journal of Hydrology, 598, 126266. https://doi.org/10.1016/J.JHYDROL.2021.126266