

DEMOCRATIC AND POPULAR ALGERIAN REPUBLIC
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
SAAD DAHLEB UNIVERSITY OF BLIDA

Faculty of Technology
Department of Water and Environmental Sciences



Master's Dissertation

Analysis of the Vulnerability of Dams to Flood Risks Using Machine Learning – Case of Beni Haroun Dam

Prepared by:

Ould Chakmakdji Oualid

Supervisor:

Dr. Bouzeria Housseyn

Academic Year: 2024-2025

Acknowledgments

First and foremost, I thank God Almighty for granting me the strength, patience, and resilience to complete this dissertation.

With a heart full of both gratitude and sorrow, I dedicate this work to the memory of my beloved mother — may Allah have mercy on her. Though she is no longer with me, her unwavering support, wise advice, and unconditional love continue to guide and inspire me every day. She always believed in me, even when I doubted myself, and I carry her words and strength in everything I do. This achievement is as much hers as it is mine.

I would also like to express my deepest gratitude to my father. His constant presence, support, and belief in me have been a source of strength throughout this journey. His sacrifices, wisdom, and quiet resilience have shaped the person I am today. I am forever grateful for his unwavering encouragement and love.

I am deeply grateful to my supervisor, Dr. Bouzeria Housseyn, for their constant guidance, valuable insights, and encouragement throughout this journey. Their support has been instrumental in shaping this dissertation.

I would also like to thank the faculty and staff of University of Blida 1, Department of Water and Environmental Science for their dedication and for providing me with the knowledge and tools necessary to grow academically and personally.

To my family and loved ones, thank you for your support, patience, and understanding during this challenging period of my life. Your presence and encouragement have been a source of comfort and motivation.

To my friends and colleagues, thank you for the discussions, collaboration, and moments of kindness that helped me stay focused and positive.

Finally, to everyone who contributed in any way to the realization of this work — whether through academic help, emotional support, or simply believing in me — I sincerely thank you.

ملخص

تُعَدّ الفيضانات المفاجئة من أكثر الظواهر الطبيعية التي تثير القلق، لما قد تُسببه من تأثيرات خطيرة على استقرار السدود وسلامة المناطق المجاورة. تهدف هذه الدراسة إلى فهم مدى هشاشة السدود أمام هذه الظاهرة، من خلال استخدام تقنيات حديثة في التعلم الآلي. وقد تم اختيار سد بني هارون كحالة دراسة. الهدف الرئيسي هو تحليل المخاطر الهيدرولوجية والتنبؤ بمدى تعرض السد للخطر اعتمادًا على مؤشرات مناخية وهيكلية.

تم جمع بيانات متنوعة مثل كميات الأمطار والمعطيات المتعلقة بتصميم السد. وقد تم تطوير نماذج تعلم آلي مثل "Random Forest"، "Gradient Boosting"، "XGBoost" و "CatBoost" وتدريبها باستخدام بيانات كميات الأمطار، بهدف تحديد العوامل الأكثر تأثيرًا على شدة الفيضانات والتنبؤ بدقة بمستوى الضعف.

أظهرت النتائج أن بعض العوامل المناخية والهيكلية تلعب دورًا رئيسيًا في تحديد خطر الفيضانات على سد بني هارون. ونجحت النماذج المطورة في التنبؤ بالحدود الحرجة للفيضانات، ما يساهم في دعم اتخاذ القرار.

تؤكد هذه الدراسة على فعالية تقنيات التعلم الآلي في تحليل مخاطر الفيضانات، ودورها في دعم سلامة السدود وتعزيز الجاهزية لمواجهة الظواهر المناخية القصوى.

الكلمات المفتاحية: خطر الفيضانات، ضعف السدود، التعلم الآلي، النمذجة الهيدرولوجية، سد بني هارون.

Résumé

Les crues figurent parmi les phénomènes naturels les plus redoutés, capables d'affecter sérieusement la stabilité des barrages et la sécurité des zones voisines. Cette étude cherche à comprendre dans quelle mesure les barrages peuvent être vulnérables face aux crues, en utilisant des techniques modernes d'apprentissage automatique. Le barrage de Beni Haroun a été choisi comme cas d'étude. L'objectif principal est d'analyser les aléas hydrologiques et de prédire la vulnérabilité du barrage à partir d'indicateurs climatiques et structurels.

Pour cela, différentes données ont été collectées, notamment les précipitations et les caractéristiques structurelles du barrage. Des modèles d'apprentissage automatique, tels que Random Forest, Gradient Boosting, XGBoost et CatBoost ont été développés et entraînés à partir de données pluviométrique. Ces modèles ont permis d'identifier les facteurs les plus influents sur les risques de crue et de prédire avec précision les niveaux de vulnérabilité.

Les résultats révèlent que certains facteurs hydrologiques et structurels ont un impact significatif sur le risque d'inondation au niveau du barrage de Beni Haroun. Les modèles développés se sont montrés efficaces pour anticiper les seuils critiques d'inondation et orienter les décisions.

Cette recherche met en avant l'efficacité de l'apprentissage automatique dans l'analyse des risques d'inondation et son utilité pour renforcer la sécurité des barrages face aux événements climatiques extrêmes.

Mots-clés: Risque d'inondation, vulnérabilité des barrages, apprentissage automatique, modélisation hydrologique, barrage de Beni Haroun.

Abstract

Flood events are among the most concerning natural phenomena due to their potential to seriously impact dam stability and nearby communities. This study aims to explore how vulnerable dams can be to such events by applying modern machine learning techniques. The Beni Haroun Dam was selected as a case study. The primary goal is to analyze flood hazards and predict the dam's susceptibility based on hydrological and structural indicators.

To accomplish this, diverse datasets were collected, including rainfall records and dam design parameters. Machine learning models such as Random Forest, Gradient Boosting, XGBoost and CatBoost were developed and trained using rainfall data. These models aimed to identify key variables influencing flood impact and accurately predict vulnerability levels.

Results showed that certain climatic and structural factors significantly affect flood risk at Beni Haroun Dam. The trained models effectively identified critical flood thresholds and provided meaningful predictions to support decision-making.

This work demonstrates the power of machine learning in analyzing flood risks and highlights its potential in supporting dam safety and preparedness for extreme weather events.

Keywords: Flood risk, dam vulnerability, machine learning, hydrological modeling, Beni Haroun Dam.

Table of contents

Acknowledgments	I
ملخص	II
Résumé	III
Abstract	IV
Table of contents	V
List of figures	X
List of tables	XII
List of abbreviations and symbols	XIII
General introduction	1
Research questions	2
Methodology overview	3
Chapter I: Theoretical Background	5
I.1 Flood risk, type and factor	6
I.1.1 Definition of flood risk	6
I.1.2 Factors	6
I.1.2.1 physical factors	6
I.1.2.2 human factors	6
I.1.3 Types of flood risks	7
I.1.3.1 Fluvial	7
I.1.3.2 Pluvial	7
I.1.3.3 Coastal	8
I.2 Dam flood risk, vulnerability, safety and importance	8
I.2.1 Dam-Specific flood risks	8
I.2.2 Challenges of dam vulnerability to floods	9
I.2.3 Dam Safety	10
I.2.4 Importance of assessing dam vulnerability	11

I.3 Machin Learning and natural disaster	12
I.3.1 ML and DL applications in disaster management	12
I.3.1.1 Disaster and hazard prediction	12
I.3.1.2 Risk and vulnerability assessment	13
I.3.1.3 Disaster detection	13
I.3.1.4 Early warning systems	13
I.4 Conclusion	14
Chapter II: Overview of Beni Haroun Dam	15
II.1 Beni Haroun dam	16
II.1.1 Situation	16
II.1.2 Importance	16
II.1.3 Objective	17
II.1.4 Geology of the dam site	17
II.1.4.1 General description	17
II.1.4.2 Lithology	18
II.2 Dam Characteristics	20
II.2.1 Structural specifications of the dam	20
II.2.1.1 Main characteristics of the RCC dam	20
II.2.1.2 Main characteristics of the management	25
II.3 Conclusion	27
Chapter III: Machine Learning, Modeling Approaches, Data Preparation and modeling .	28
III.1 Machine Learning	29
III.1.2 Types of Machine Learning	29
III.1.3 Commonly used Machine Learning algorithms	30
III.1.3.1 Linear regression	31
III.1.3.2 Logistic regression	31
III.1.3.3 Neural networks	32

III.1.3.4 Decision trees	32
III.2 Modeling approaches	32
III.2.1 RF (Random Forest)	32
III.2.1.1 Random Forest works	33
III.2.2 XGBoost	34
III.2.2.1 XGBoost Works	35
III.2.3 Gradient Boosting	36
III.2.3.1 Ensemble learning and boosting	36
III.2.4.2 Gradient Boosting works	37
III.2.4 CatBoost	38
III.2.4.1 CatBoost algorithm works	38
III.3 Data Preparation for modeling	40
III.3.1 Historical floodwater release data	40
III.3.2 End-of-month volume	40
III.3.3 Monthly inflow	41
III.3.4 Variable of volume	42
III.3.5 Max daily rainfall	43
III.4 Modeling	44
III.4.1 Feature selection and engineering	44
III.4.2 Target variable construction	44
III.4.3 Handling missing or anomalous data	44
III.4.4 Normalization and scaling	44
III.4.5 Data Splitting: training and testing	45
III.4.6 Encoding and labeling	45
III.4.7 Performance evaluation criteria	45
III.5 Conclusion	46
Chapter IV: Results, Discussion and Comparison	48

IV.1 Predictive model outcomes	49
IV.1.1 Random Forest (RF)	49
IV.1.1.1 VIF and tolerance statistics	49
IV.1.1.2 Model metrics	50
IV.1.1.3 Classification report	50
IV.1.1.4 Confusion matrix	51
IV.1.1.5 Normalized confusion matrix	52
IV.1.1.6 ROC curve – multi-class	53
IV.1.1.7 Feature Importance	55
IV.1.1.8 Temporal flood risk map	56
IV.1.1.9 Annual trend of monthly inflow	57
IV.1.2 XGBoost	58
IV.1.2.1 VIF and tolerance statistics	58
IV.1.2.2 Model metrics	59
IV.1.2.3 Classification report	59
IV.1.2.4 Confusion matrix	60
IV.1.2.5 Normalized confusion matrix	61
IV.1.2.6 ROC curve – multi-class	62
IV.1.2.7 Feature importance	64
IV.1.2.8 Temporal flood risk map	65
IV.1.2.9 Annual trend of monthly inflow	66
IV.1.4 Gradient Boosting	67
IV.1.4.1 VIF and tolerance statistics	67
IV.1.4.2 Model metrics	68
IV.1.4.3 Classification report	68
IV.1.4.4 Confusion matrix	69
IV.1.4.5 Normalized confusion matrix	70

IV.1.4.6 ROC curve – multi-class	71
IV.1.4.7 Feature importance	73
IV.1.4.8 Temporal flood risk map	74
IV.1.4.9 Annual trend of monthly inflow	75
IV.1.5 CatBoost	76
IV.1.5.1 VIF and tolerance statistics	76
IV.1.5.2 Model metrics	77
IV.1.5.3 Classification report	77
IV.1.5.4 Confusion matrix	78
IV.1.5.5 Normalized confusion matrix	79
IV.1.5.6 ROC curve – multi-class	80
IV.1.5.7 Feature importance	82
IV.1.5.8 Temporal flood risk map	83
IV.1.5.9 Annual trend of monthly inflow	84
IV.2 Comparative analysis of Machine Learning models for flood risk classification	85
IV.2.1 Overall model performance	85
IV.2.2 Detection of high inflow events	85
IV.2.3 Feature importance and model transparency	86
IV.2.4 Temporal risk mapping and operational utility	86
IV.3 Conclusion	86
IV.4 Recommendations	87
IV.4.1 Flood risk mitigation strategies	87
IV.4.2 Proposed improvements for dam flood management	88
General conclusion	89
Perspectives	90
References	91

List of figures

Figure 1. Location of Beni Haroun dam

Figure 2. Spillway cross-section

Figure 3. Cross-section of non-overflow block

Figure 4. Plan view

Figure 5. Random Forest diagram

Figure 6. XGBoost diagram

Figure 7. Gradient Boosting diagram

Figure 8. Structure of CatBoost algorithm

Figure 9. Clustered column of end-of-month volume

Figure 10. Histogram of monthly inflow

Figure 11. Variable of volume

Figure 12. Max daily rainfall

Figure 13. Confusion matrix of RF

Figure 14. Normalized confusion matrix of RF

Figure 15. Multi-class ROC curve of RF

Figure 16. Temporal flood risk map of RF

Figure 17. Annual trend of monthly inflow of RF

Figure 18. Confusion matrix of XGBoost

Figure 19. Normalized confusion matrix of XGBoost

Figure 20. Multi-class ROC curve of XGBoost

Figure 21. Temporal flood risk map of XGBoost

Figure 22. Annual trend of monthly inflow of XGBoost

Figure 23. Confusion matrix of Gradient Boosting

Figure 24. Normalized confusion matrix of Gradient boosting

Figure 25. Multi-class ROC curve of Gradient Boosting

Figure 26. Temporal flood risk map of Gradient Boosting

Figure 27. Annual trend of monthly inflow

Figure 28. Confusion matrix of CatBoost

Figure 29. Normalized confusion matrix of CatBoost

Figure 30. Multi-class ROC curve of CatBoost

Figure 31. Temporal flood risk map of CatBoost

Figure 32. Annual trend of monthly inflow of CatBoost

List of tables

Table 1. Main characteristics of the reservoir

Table 2. Main characteristics of the dam

Table 3. Main characteristics of the flood spillway

Table 4. Main characteristics of the half-depth drainage

Table 5. Main characteristics of the injection and drainage of the foundation

Table 6. Main characteristics of the diversion structure

Table 7. Historical floodwater release data, spill (hm³)

Table 8. VIF and tolerance statistics of RF

Table 9. Classification report of RF

Table 10. Variable and importance of RF

Table 11. Vif and tolerance statistics of XGBoost

Table 12. Classification report of XGboost

Table 13. Variable and importance of XGBoost

Table 14. VIF and tolerance statistics of Gradient Boosting

Table 15. Classification report of Gradient Boosting

Table 16. Variable and importance of Gradient Boosting

Table 17. VIF and tolerance statistics of CatBoost

Table 18. Classification report of CatBoost

Table 19. Variable and importance of CatBoost

List of abbreviations and symbols

ML: Machine Learning

DL: Deep Learning

Eml: Eocene marly limestones

Pm: Paleocene marls

El: Eocene limestones

Em: Eocene marls

C: Colluvium

S: Scree

T: Alluvial terrace

A: Alluvium

E: Embankments

RCC: Roller-compacted concrete

NR: Normal Retention

NL: Normal Level

M.W.L: Maximum Water Level

MOL: Minimum Operating Level

RF: Random Forest

XGBoost: eXtreme Gradient Boosting

CatBoost: Category Boosting

VIF: Variance Inflation Factor

AUC: Area Under the Curve

ROC: Receiver Operating Characteristic

OvR: One-vs-Rest

General introduction

In recent decades, the frequency and intensity of flood events have increased significantly due to the combined effects of climate change, urbanization, and environmental degradation. These extreme events pose serious threats to water infrastructure, particularly dams, which are essential for water storage, energy generation, flood control, and irrigation. Ensuring the proper functioning and safety of dams is therefore critical for protecting downstream communities and supporting socio-economic development. ^[1]

One of the key challenges in dam management today is the rising risk of extreme hydrological events that may exceed design capacities and expose structural vulnerabilities. This highlights the urgent need for advanced methods to assess dam vulnerability under various flood scenarios—not only to prevent potential failures but also to enhance decision-making in risk management and emergency planning. ^[2]

In this context, machine learning techniques are proving highly valuable in the fields of hydrology and civil engineering. These methods can model complex, non-linear relationships and deliver reliable predictions based on large and diverse datasets. Their predictive capabilities offer new opportunities to assess flood risks and understand the behavior of dams under extreme conditions. ^[3]

This study focuses on the Beni Haroun Dam, the largest dam in Algeria and a critical element of the country's hydraulic infrastructure. Given its role in supplying water to several wilayas, evaluating its vulnerability to flood events is of utmost importance. Yet, few studies have thoroughly explored this issue using modern predictive approaches based on machine learning, leaving a gap that this research aims to address.

Dams are critical components of water resource management systems, serving essential functions such as water supply, irrigation, flood control, and hydroelectric power generation. However, these structures are increasingly exposed to hydrological extremes due to climate variability and land use changes. In particular, intense and unpredictable flood events pose serious threats to the structural safety and operational efficiency of dams. ^[4]

The main objective of this study is to assess flood risks and analyze dam vulnerability using hydrological approaches and machine learning (ML) techniques, focusing on the Beni Haroun Dam. This research aims to improve flood prediction accuracy and optimize dam risk management.

Research questions

1. How can ML improve flood risk predictions?
2. Which of the selected models performs best in classifying flood risk levels based on predictive accuracy and other performance metrics?

Methodology overview

The methodology adopted in this study is structured around four principal stages, aimed at assessing the vulnerability of dams to flood risks using Machine Learning techniques. The approach focuses on data integration, hydrological modeling, and predictive analytics to understand and evaluate flood behavior in the context of dam safety.

1. Data collection and preparation

The first stage involves assembling relevant datasets from hydrological, climatological, structural, and environmental sources:

- **Hydrological and climatological data:** Includes average and peak discharges and rainfall data from past episodes. Time series over multiple years are used to capture hydrological trends.
- **Dam-related data:** Covers the structural characteristics of the dam, historical records of water releases during flood periods.

2. Machine Learning, modeling approaches and data preparation

The third phase centers on the use of machine learning techniques to quantify relationships between variables and forecast flood risks:

- **Data preparation:** A tabular dataset is constructed from the climate, hydrological. Standardization is applied to prevent bias in model training.
- **Modeling techniques:**
 1. Random Forest
 2. XGBoost
 3. Gradient Boosting
 4. CatBoost
- **Validation:** The dataset is split into 70% training and 30% testing sets. Cross-validation techniques are used to ensure model robustness and reliability.

3. Expected results

The methodology aims to yield the following outcomes:

- A predictive model capable of estimating the probability and severity of flood events under varying climatic conditions.
- Identification of critical thresholds for the dam, including its maximum capacity before overtopping.
- Strategic recommendations to improve flood management practices and reduce dam vulnerability.

Chapter I

Theoretical Background

I.1 Flood risk, type and factor

I.1.1 Definition of flood risk

“Flood risk” is a combination of the probability and the potential consequences of flooding. Areas at risk of flooding are those at risk of flooding from any source, now or in the future. Sources include rivers and the sea, direct rainfall on the ground surface, rising groundwater, overwhelmed sewers and drainage systems, reservoirs, canals and lakes and other artificial sources. ^[5]

I.1.2 Factors

I.1.2.1 physical factors

Prolonged rainfall

After prolonged rainfall, soil becomes saturated, causing an increase in surface runoff. As rainfall can no longer infiltrate the soil, more water enters the river channel, increasing the likelihood of flooding. ^[6]

Heavy rainfall

Heavy rainfall can result in water arriving too quickly to infiltrate the soil. This increases surface run-off, leading water to reach the river channel quicker, resulting in a greater risk of flooding. ^[6]

Geology

Impermeable surfaces such as clay and granite do not allow infiltration, leading to greater surface run-off. The risk of flooding increases as water reaches the river channel quickly, increasing discharge and the risk of flooding. ^[6]

Relief

Water can move swiftly downhill in regions with steep terrain, such as mountains. The steeper the slope, the more rapid the water flow into a river channel, increasing the risk of flooding. ^[6]

I.1.2.2 human factors

Landuse

Urban development affects water movement as well. Hard, non-absorbent surfaces like concrete, asphalt, and drainage systems can speed up water flow to rivers. Plus, the scarcity of greenery means less rainwater is absorbed or evaporated. Drains also increase water flow into surrounding rivers, increasing the flood risk.

Agricultural practices can influence how water moves. For instance, fields without crops, especially during winter when soils are wet, can quickly channel water. Also, ploughing down slopes can form small channels that accelerate water flow directly to rivers. ^[6]

Deforestation

Vegetation, including trees, intercepts rainfall, slowing its movement. Some of this water is stored before evaporation returns moisture to the atmosphere. Trees also absorb water from the soil, allowing greater infiltration into the soil and reducing surface run-off. When vegetation is removed, infiltration and interception are reduced and surface run-off increases. Without trees, more water flows directly to rivers, raising the potential for flooding. This leads to a greater risk of flooding as more water reaches the river channel. ^[6]

I.1.3 Types of flood risks

I.1.3.1 Fluvial

A fluvial or river flood is a flood that occurs in and around lakes, streams, or rivers. Fluvial flooding can either be overbank flooding—when the water level rises over the edges of a river, stream, or lake—or flash flooding—where there is a high volume of water moving at a high velocity in an existing riverbed with little advance notice.

While the amount, accumulation, and duration of precipitation are the main contributing factors to fluvial flooding, soil water saturation and the area's terrain also impact the likelihood of a fluvial flood. In flatter areas, floodwaters rise more slowly but stick around for days. While in more sloping and mountainous areas, the floods tend to happen more quickly and move faster. ^[7]

I.1.3.2 Pluvial

Pluvial flooding occurs when increased rainfall creates a flood independent of an existing body of water. These floods can occur in any location, even without nearby water bodies. Like fluvial flooding, pluvial floods come in two forms: surface water floods and flash floods. Surface water floods occur when a drainage system is overwhelmed. Flash floods, in this case, are just like fluvial flash floods; however, they are not contained to a river system. These flash floods are triggered by immense rainfall in a short period near low-lying terrain. ^[7]

I.1.3.3 Coastal

Coastal flooding is the inundation of land along the coast by seawater. It can be caused by high tides, storm surge, and tsunamis.

High-tide flooding is generally defined as a coastal water level reaching or exceeding a median height of 0.5 to 0.65 meters above the long-term average daily highest tide. The number of days with minor high tide flooding

Storm surge is the abnormal rise in seawater levels during a tropical cyclone, or hurricane, and is measured as the height of the water above the normal predicted tide. The magnitude of storm surge depends on the orientation of the coast with the track of the hurricane; the hurricane's strength, size, and speed; and coastal and underwater topography. Whether the hurricane hits during high or low tide can matter for the resulting magnitude of flooding. ^[7]

I.2 Dam flood risk, vulnerability, safety and importance

I.2.1 Dam-Specific flood risks

A dam failure is a catastrophic type of failure characterized by the sudden, rapid and uncontrolled release of impounded water accompanied by the trapped silt and debris that erode and accumulate additional debris along the way. ^[8]

The most common cause of dam failure is flood or dam overtopping. The next common cause is piping or seepage. Different causes attributable to the structural failure comprise the third most common category. Sometimes a dam may even fail due to the failure of its spillway gate, earthquake or even poor design/construction. The many types of dam failures may be summarised using five failures scenarios/events: hydrologic, geologic, structural, seismic, and human-influenced. ^[9]

I.2.2 Challenges of dam vulnerability to floods

Dams are susceptible to floods due to several factors, including extreme rainfall, climate change, and structural weaknesses.

- **Extreme Rainfall:** Dams are built to withstand specific flood levels, but sometimes, extreme rainfall can exceed their limits, causing them to overflow or even fail. A real example of this happened in Libya in 2023 when two dams collapsed due to the intense rainfall brought by Storm Daniel, a powerful Mediterranean cyclone. ^[10]
- **Climate Change:** Research in the US has also indicated an increased risk of dam failure due to a changing climate. Analysis of rainfall sequences and events associated with recent hydrologic failures of 552 dams across the country, suggests intensifying precipitation may contribute to increasing failures of dams by overtopping. ^[10]

The decadal rate of dam failures has been increasing since the 1970s, and with over 90,000 ageing dams still in service, the increasing likelihood of intense rainfall is leading to increased concern about future dam failures. ^[11]

- **Structural Weaknesses:** Many manmade structures including dams, bridges and buildings were not built to withstand the extreme rainfall events happening today. Advancing age makes dams more susceptible to failure. ^[11]
- **Inadequate Spillway Capacity:** The extreme rainfall event will cause increased stream flows resulting in the water level in the reservoir to rise to heights that the dam may have never previously experienced. And, if the dam and spillway system are not equipped to safely pass an extreme rainfall event, the reservoir level will rise and water will go over the dam itself. This is called “overtopping.” ^[12]
- **Poor Reservoir Management:** The management of reservoir systems is a very complex issue and is often considered a wicked problem. In order to understand the wickedness of reservoir management, one needs first to understand what a wicked problem is: as a class of social system problems that are ill-formulated, where the information is confusing, where there are many clients and decision makers with conflicting values, and where the ramifications in the whole system are thoroughly confusing. ^[13]

I.2.3 Dam Safety

Water stored behind a dam represents potential energy which can create a hazard to life and property located downstream of a dam. In order for a dam to safely fulfill its intended function, it must be constructed, operated and maintained properly.

The risks associated with the storage of water must be minimized at all times. The height of a dam, its maximum impoundment capacity, the physical characteristics of the dam site and the location of downstream facilities should be assessed to determine the appropriate hazard classification.

The functions of the Dam Safety Section include: safety inspection of dams; technical review of proposed dam construction or modification; monitoring of remedial work for compliance with dam safety criteria; and emergency preparedness. ^[14]

I.2.4 Importance of assessing dam vulnerability

Comprehensive evaluation of dams is a vital process that involves assessing the safety, stability, and overall performance of dams. It evaluates various aspects of the dam, including structural integrity, hydraulic efficiency, environmental impact, and risk management.^[15]

Assessing dam vulnerability is crucial for several reasons:

1. **Ensuring structural integrity:** The primary objectives of comprehensive evaluation of dams are to assess their structural integrity and efficiency. They also aim to identify any existing or potential issues, such as dam deterioration, seepage, or vulnerability to natural hazards. The potential environmental impact of the dam and its reservoir, including potential effects on water quality, wildlife habitats, and downstream ecosystems are also evaluated. As the dam is assessed, engineers consider the dam's location and vulnerabilities to develop effective risk management strategies.^[15]
2. **Adapting to environmental changes:** With global climate change resulting in unpredictable and extreme patterns of rainfall, heavily aged dams are at an even greater risk of succumbing to the effects of extreme weather. As such, dam failure has become a prevalent issue, endangering communities located underneath the infrastructure. This helps in understanding how these factors impact dam stability and guides necessary adaptations.^[16]
3. **Protecting downstream communities:** The purpose of this guide is to outline a procedure for identifying and assessing the potential consequences of dam failure at the community level using readily available information. A good assessment will establish the potential short- and long-term economic, social, and environmental effects of dam failure that will inform planning efforts.^[17]
4. **Informing risk management strategies:** are short-term actions that the Corps can take to reduce risk, based on information gathered during risk assessments or inspections. We implement these to reduce risk while we study potential dam safety issues further, perform maintenance actions, or pursue long-term modifications. Examples of interim risk reduction measures include structural measures, increased inspection and monitoring, dam safety training, stockpiling of emergency materials, operational changes or restrictions, updating and exercising Emergency Action Plans, and increased risk communication and coordination.^[18]

I.3 Machine Learning and natural disaster

Amidst the continually changing climate and the rise in natural disasters, it is crucial to strengthen resilience against these calamities. This chapter explores the dynamic intersection of machine learning and natural disasters, revealing how advanced technologies reshape disaster management. In the face of escalating challenges posed by earthquakes, floods, and wildfires, machine learning emerges as an innovative solution, offering proactive approaches beyond conventional reactive methods. The narrative unfolds by tracing the evolution of disaster management, highlighting the transformative impact of machine learning on early warning systems. It explores predictive analytics and risk assessment, elucidating how machine learning algorithms leverage historical data and real-time information to deepen our understanding of disaster vulnerabilities. Beyond prediction, the discourse extends to the pivotal role of machine learning in optimizing response and recovery efforts—efficiently allocating resources and fostering recovery planning. A critical dimension of this integration emerges in the analysis of remote sensing and satellite imagery, where machine learning algorithms enable more accurate and timely disaster monitoring. The exploration extends further, unraveling the interconnectedness of various hazards and emphasizing how machine learning facilitates a holistic understanding. The synergy between machine learning and traditional knowledge systems comes to the forefront, recognizing the significance of integrating local wisdom into predictive models.

[19]

I.3.1 ML and DL applications in disaster management

I.3.1.1 Disaster and hazard prediction

Disaster and hazard prediction utilizing ML (Machine Learning) and DL (Deep Learning) involve the development of models that analyze historical data and real-time information to forecast potential disasters. This approach proves instrumental in situations where timely predictions can mitigate risks and protect vulnerable communities. [19]

- **Increased Predictive Accuracy:** ML and DL models enhance the accuracy of disaster predictions, allowing for more reliable forecasting.
- **Proactive Risk Mitigation:** Early identification of potential hazards enables proactive measures, reducing the impact on communities.
- **Adaptability to Various Disasters:** ML and DL models can be adapted to predict a wide range of disasters, including floods, earthquakes, wildfires, hurricanes, and more.

I.3.1.2 Risk and vulnerability assessment

ML and DL technologies contribute significantly to risk and vulnerability assessment, offering nuanced insights into susceptibilities in specific regions. This application is particularly useful in situations where targeted mitigation strategies are required. ^[19]

- **Precision in identifying vulnerabilities:** DL models process vast datasets to pinpoint vulnerabilities accurately, aiding in the formulation of targeted mitigation strategies.
- **Adaptability to varied regions:** ML applications in risk assessment can be tailored to different geographical regions, ensuring relevance in diverse contexts.
- **Comprehensive analysis:** The capacity to analyze various factors contributes to a comprehensive understanding of disaster risks.

I.3.1.3 Disaster detection

ML applications in disaster detection focus on swiftly identifying events like floods and wildfires, enabling rapid response in critical situations. ^[19]

- **Swift event identification:** ML algorithms can quickly identify disasters, facilitating timely responses to minimize damage.
- **Adaptive algorithms:** ML models can adapt to different disaster scenarios, enhancing their versatility.
- **Real-time monitoring:** The integration of real-time data allows for continuous monitoring and detection of evolving disaster situations.

I.3.1.4 Early warning systems

ML and DL play a pivotal role in optimizing early warning systems, offering timely alerts and fostering community resilience. ^[19]

- **Timely alert generation:** ML algorithms process data rapidly, enabling the generation of timely alerts for at-risk communities.
- **Community-centric approach:** Early warning systems become more effective by incorporating community voices and experiences through ML applications.
- **Integration with various data sources:** ML models can integrate data from diverse sources, enhancing the comprehensiveness of early warnings.

I.4 Conclusion

Chapter I establishes a comprehensive foundation on flood risks, their causes, and the critical role of dam safety, particularly in the context of increasing environmental challenges. It begins by defining flood risk as a combination of both the probability of flooding and the magnitude of its consequences. The chapter systematically explores both natural (physical) and human-induced factors contributing to flooding. Prolonged and intense rainfall, impermeable geological surfaces, and steep slopes are shown to accelerate runoff into river systems, while urbanization, agriculture, and deforestation further exacerbate the problem by altering natural water absorption and flow patterns.

The chapter distinguishes among three main types of flood risks—*fluvial*, *pluvial*, and *coastal*—highlighting their distinct causes and behaviors. This classification emphasizes the need for multi-faceted flood management strategies, particularly as extreme weather events become more frequent.

In addressing dam-specific flood risks, the chapter brings focus to vulnerabilities associated with structural failures, overtopping, and poor management practices. Real-world examples, such as the catastrophic dam failures in Libya due to Storm Daniel, illustrate the devastating potential of extreme weather compounded by aging infrastructure and design limitations.

A significant portion of the chapter is dedicated to assessing dam safety. It stresses the importance of inspections, emergency preparedness, and structural monitoring, while also underlining how climate change and increased rainfall intensify the risks faced by aging dams. The chapter highlights that dam safety is not a one-time concern but an ongoing responsibility involving hydrological, structural, and management considerations.

Finally, the integration of Machine Learning (ML) into disaster risk management is introduced. ML is presented as a transformative tool in flood prediction, detection, and risk assessment. From early warning systems to real-time monitoring and vulnerability analysis, ML offers promising methods for enhancing both preparedness and response strategies. The synergy between data-driven models and traditional engineering practices is proposed as a powerful approach to managing increasingly complex flood risks in the modern era.

Chapter II

Overview of Beni Haroun Dam

II.1 Beni Haroun dam

II.1.1 Situation

The Beni Haroun Dam is a large strategic hydraulic complex in Algeria.

The dam site is located in the Wilaya of Mila, in northeastern Algeria, on the Oued El Kebir, about forty kilometers from its mouth in the Mediterranean Sea.

The reservoir created by the dam is located to the south of it. It is situated approximately 40 km North / North-West of the city of Constantine and 350 km east of Algiers.

The dam is positioned at the upstream end of the limestone-marl gorge of Beni Haroun and about 4 km from the confluence of the Oued Rhumel and the Oued Enndja.

The total basin of the Oued Kebir at the dam site covers an area of about 7,725 km².

Considering that part of the Rhumel basin is regulated by the Hammam Grouz Dam, the area relevant to the Beni Haroun Dam is 6,595 km². [20]

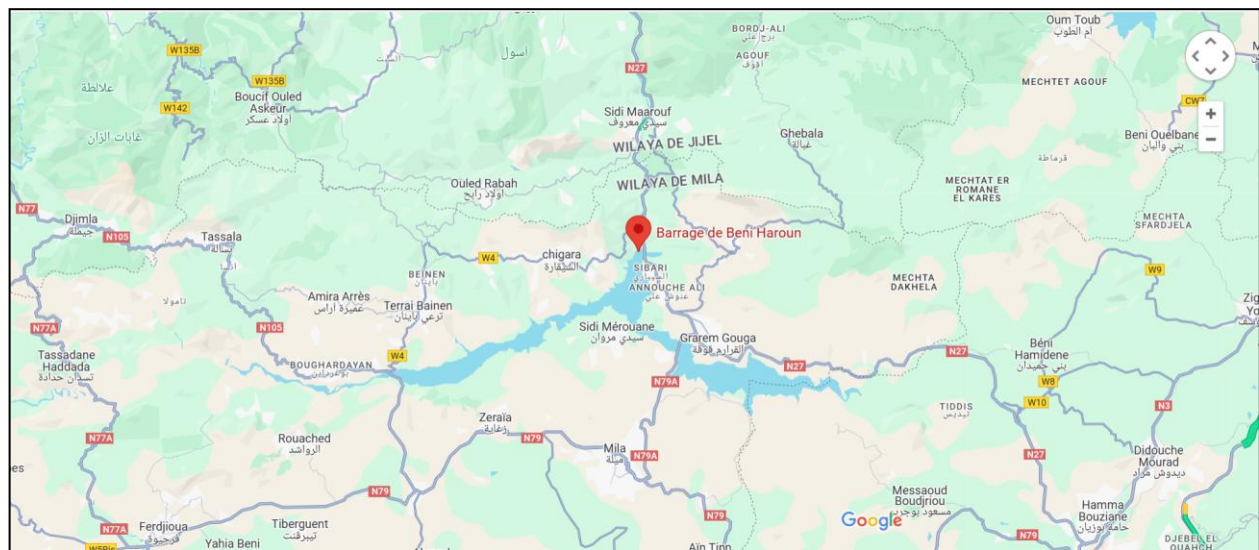


Figure 1. Location of Beni Haroun dam

II.1.2 Importance

The eastern region of Algeria is characterized by a low water demand in the north, despite abundant water resources. In contrast, in the southern high plateaus, water resources are limited, while the demand for drinking water is rapidly increasing, and there are large areas of irrigable agricultural land. [21]

II.1.3 Objective

The Beni Haroun transfer project aims to transfer the water resources mobilized by the Beni Haroun and Bou Siaba dams (partially) to six Wilayas: Batna, Khenchela, Mila, Oum El Bouaghi, Constantine, and Jijel (El Milia region). ^[21]

II.1.4 Geology of the dam site

II.1.4.1 General description

The dam site is characterized by an Eocene limestone ridge approximately 100 meters thick, with a general East-West orientation, resting on Paleocene black marls downstream and overlain by Eocene marls upstream.

This uplift of the limestone ridge results in a reduction of its outcrop surface from the right bank to the left bank.

The direction of the stratification shifts from N80° to 90°E on the right bank to N35°E on the left bank.

Upstream of this limestone bed, Eocene marls (Em) are detected overlying the limestone, while downstream, beneath the limestones, dark Paleocene marls (Pm), more or less fractured, can be distinguished. In the marls, gypsum striations and some pyrite appear, particularly in the upper layers.

In the valley bottom and on the right bank, the contact between the downstream Paleocene marls and the Eocene limestones is conformable. The transition occurs gradually through an increase in the number and thickness of the limestone beds. This transition zone is between 5 and 15 meters thick.

Upstream of the limestone ridge, the contact with the Eocene marls is more tectonized. It is marked by a transition from marl-limestone to black schistose marls, with a transitional zone about 15 meters thick alternating yellowish limestone beds (20 to 40 cm thick) and dark marl beds. The dip of the marl layers shows relatively tight folding with dips that are nearly vertical or even overturned as one moves away from the limestones. In this area, several blocks with slickensided surfaces—indicating shear movements in these layers—are visible, and it is possible that significant limestone lenses may also be observed there.

The limestone ridge itself is crossed by numerous joints and fractures—some filled with clay and showing signs of shearing or slipping, others filled with calcite—and by karstic dissolution features. These have developed mainly along certain nearly vertical joints. The thickness of these karstified zones can reach 30 to 50 centimeters and appears as a clayey or brecciated infill, recemented by calcite.

In the limestones, three facies have been identified: an upper section of marly limestones (Eml) about 30 to 40 meters thick, an intermediate section of limestones with flint nodules about 30 meters thick, and a lower section of black limestones with a thickness of 70 to 100 meters. ^[21]

II.1.4.2 Lithology

The various formations observed in the dam area, from oldest to most recent, are as follows:

Paleocene Marls (Pm)

They outcrop on the slopes and in excavations downstream of the site. They were identified through several boreholes during the early project phases.

These marls are more or less calcareous and slightly pyritiferous. Locally, traces of gypsum and calcite can be seen filling joints. Septaria are frequently found in these marls.

In their upper part, over a thickness of about 5 meters, there is an alternation of marls and limestone beds forming the transition to the overlying layers. ^[21]

Eocene Limestones attributed to the Ypresian (El)

Detailed geological mapping of the site has made it possible to distinguish three subdivisions of the limestone ridge forming the foundation of the dam:

- Marly limestones (Eml)
- Limestones with flint (Elf)
- Basal limestones (Ebl)

Overall, these rocks appear in decimetric to metric layers of fine-grained limestone, dark gray to black, hard and resistant. Some limestone beds are extremely hard due to dolomitization or silicification, while others contain black flint. This flint appears as small nodules (1–3 cm in diameter) and thin bands (2 to 15 cm thick). Thin interbedded layers of calcareous shale (up to 30 cm) are also present.

The limestone formation is relatively resistant to weathering and erosion, but locally shows signs of dissolution.

As the limestones weather, they typically become pale gray to beige. ^[21]

Eocene Marls (Em) (Lutetian age)

The limited outcrops of this formation show compacted shales, dark gray to black, with a lithology that appears almost identical to the facies of the Paleocene marls.

The presence of small septaria allows them to be distinguished.

In most outcrops, the marl is highly deformed by tectonic activity and has subsequently undergone surface weathering. ^[21]

Colluvium and Scree (C and S)

On the right abutment, slope deposits or colluvium are observed, consisting of angular limestone stones, sometimes slightly cemented or containing a clayey matrix.

Their thickness can reach up to 20 meters.

On the left bank, colluvial deposits of stones, limestone blocks, and marls are found.

In general, the upstream and downstream marls are, in many places, covered by these weathering products—either in place or displaced by sliding. ^[21]

Alluvial Terrace (T)

In the lower part of the right bank, there is an alluvial terrace with its base located around elevations 138–140. It is composed of gravel and large blocks within a sandy matrix, transitioning to pebbles in a more clayey matrix further into the slope.

The thickness can reach up to 20 meters and decreases downstream and toward the right abutment. This terrace features several intermediate levels. ^[21]

Alluvium (A)

These are mainly gravels, large blocks, and alluvial sands from the riverbed (oued). ^[21]

Embankments (E)

The main embankments are those of the cofferdams and structures within the riverbed, as well as those covering the former thalweg of the right abutment. They are more than 20 meters thick between the dam abutment and that of the viaduct. ^[21]

II.2 Dam Characteristics

II.2.1 Structural specifications of the dam

II.2.1.1 Main characteristics of the RCC dam

The dam is a straight gravity-type structure made of roller-compacted concrete (RCC) resting on a limestone rock foundation.

The normal retention level (NR), which is also the crest of the free-flow spillway, is at elevation 200.

This corresponds to a structure with a maximum height from the foundation of 118 meters and a crest length of 710 meters. ^[21]

The design of the dam and its auxiliary structures was guided by the goal of:

- Minimizing the quantities of conventional concrete
- Ensuring good continuity of the RCC construction site
- Concentrating as much as possible the structures that require the use of traditional concrete.

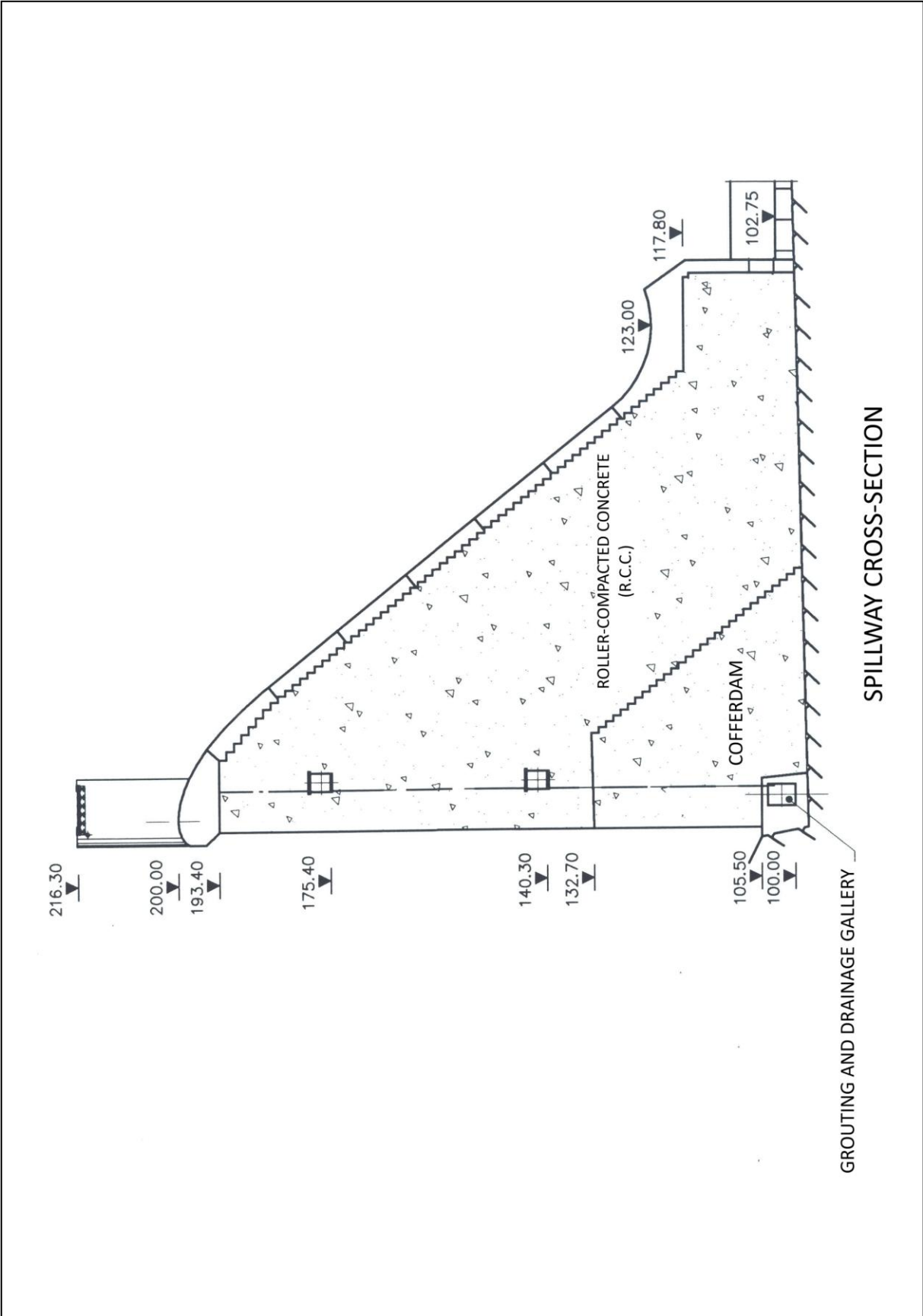


Figure 2. Spillway cross-section [21]

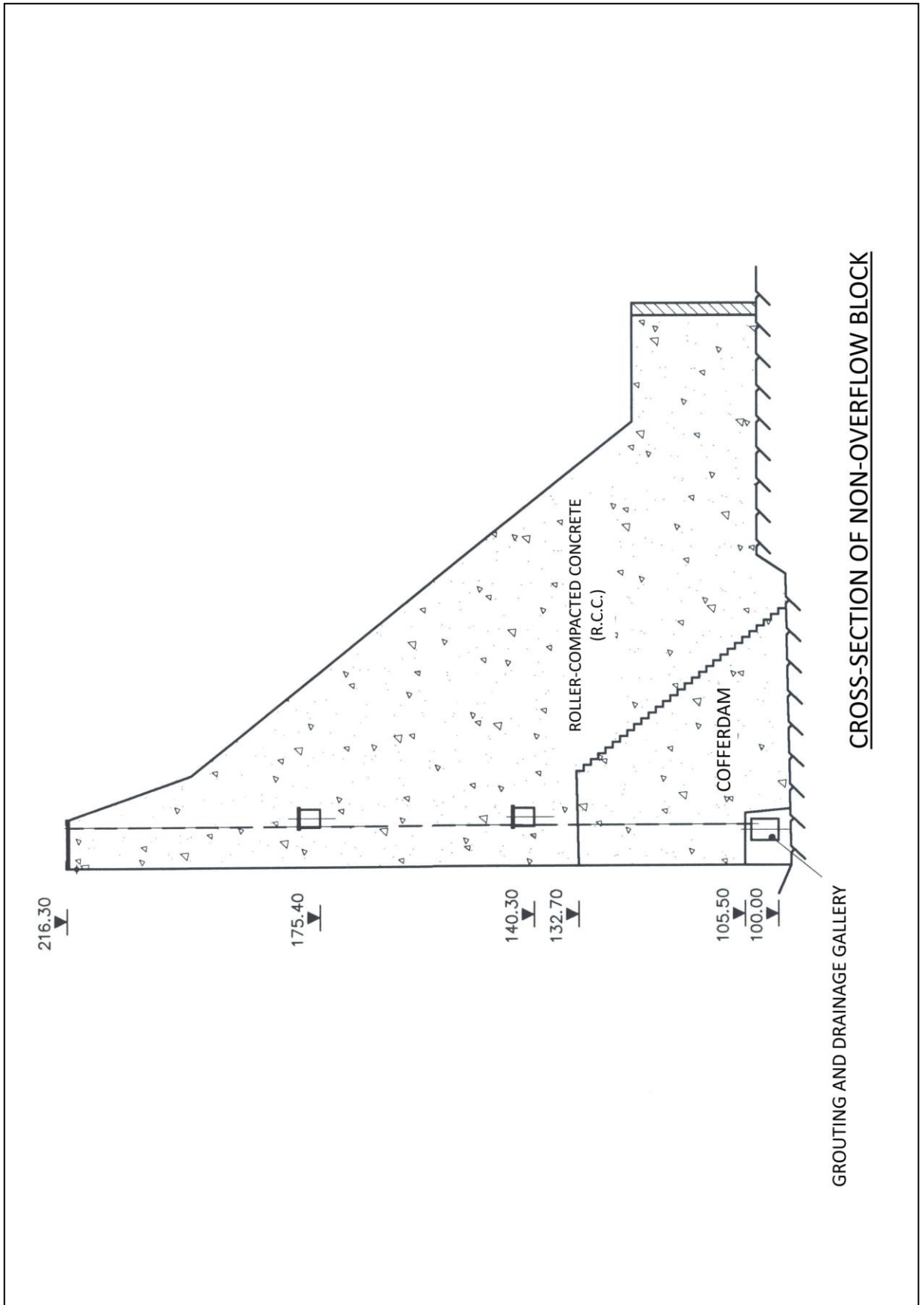


Figure 3. Cross-section of non-overflow block ^[21]

The basic profile of the dam is a triangle with the following characteristics:

- Apex at elevation 214.80 m, maximum water level (M.W.L.)
- Vertical upstream face
- Downstream face inclined at 0.8 horizontal for 1 vertical

The non-overflow section has a crest 8 m wide, leveled at elevation 216.30 m, which provides a freeboard of 1.5 m above the highest water level.

The downstream face of this section changes slope above elevation 196, which enhances earthquake resistance and allows the installation of a roadway.

The surface spillway, of the "free overflow" type, is economically built, as in all concrete dams, by shaping the central part of the dam as an overflow section.

It is 124 m long, divided into 6 bays topped by a road bridge.

The spillway crest is extended by a chute connected to a ski jump, positioned above the highest water level. The spillway capacity is 13,230 m³/s at a water level of 214.74 m, which corresponds to the maximum flow released after attenuation of the probable maximum flood (PMF) of 16,640 m³/s.

The valley floor has been widened by cutting into the right bank for about 100 meters to ensure a solid foundation for the part of the dam affected by the spillway and bottom outlet.

The bottom outlet, located next to the spillway, is installed on the right bank at elevation 140.

It consists of two conduits designed to discharge a flow of 670 m³/s (335 m³/s per conduit) below the normal reservoir level, ending in a ski jump.

During construction, river diversion was ensured during the dry season (May–October) by the two diversion tunnels on the left bank, as planned in the original rockfill solution. During the rainy season, these tunnels were also sufficient, and although the overflow of the central part of the dam could have been used to pass major floods, it was not necessary.

Three galleries, distributed over 3 levels and located near the upstream face of the dam, are used for injection and drainage operations within the dam body and foundation.

Positioned at elevations 100 m, 140 m, and 175.10 m, they are located 4.00 m from the upstream face for the lower gallery and 6.50 m for the two upper ones. These galleries extend into the abutments over distances ranging from 70 to 130 m.

An additional rock gallery, at elevation 120.00 m, reinforces drainage on the right bank. ^[21]

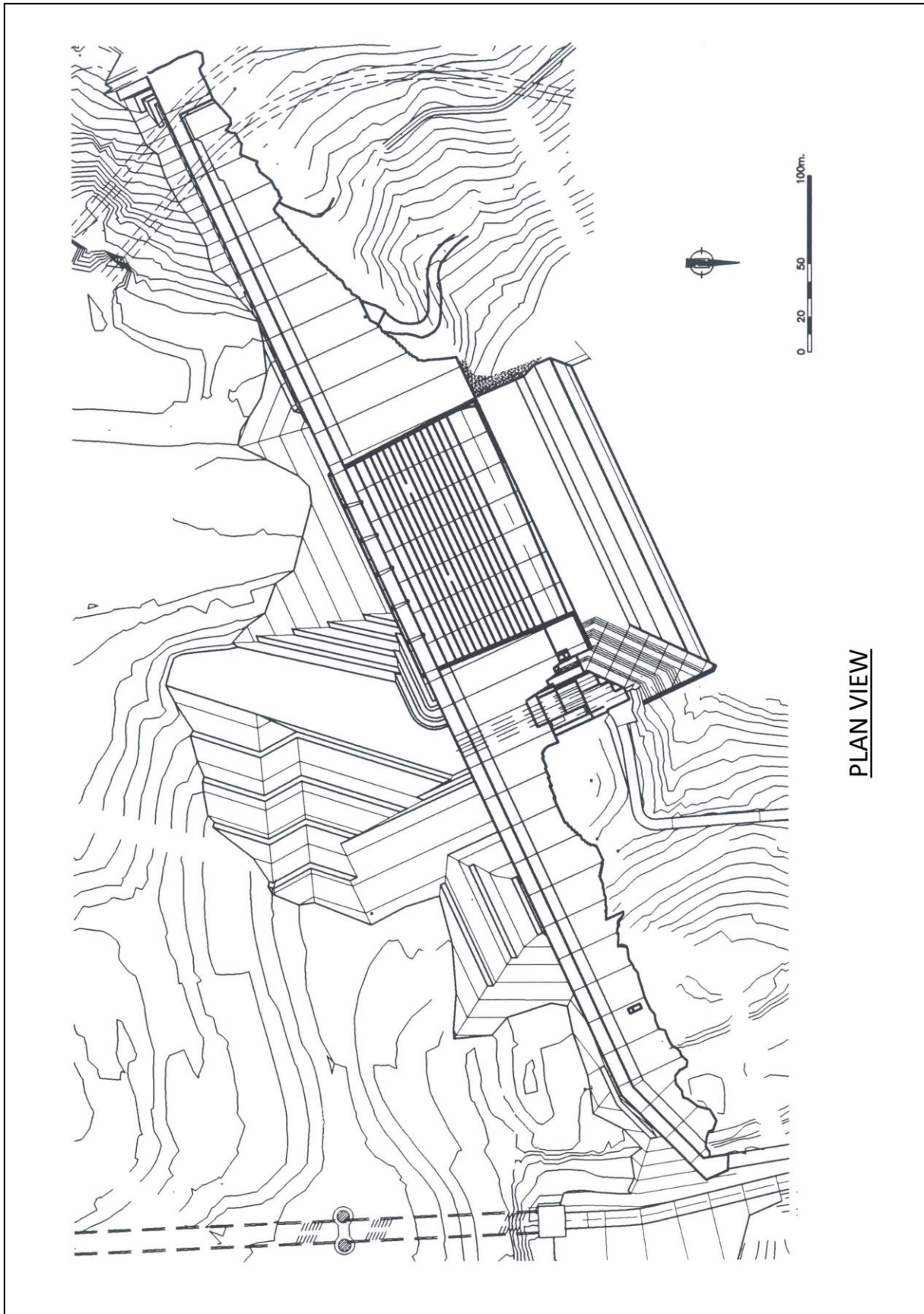


Figure 4. Plan view ^[21]

The gallery system described above allows for the collection of drainage water from the dam body, foundation, and slopes. It also enables monitoring of the dam's behavior and facilitates maintenance and any necessary repair operations.

Beneath the dam, the waterproofing barrier consists of a multi-row grout curtain (generally 3 rows) with a depth ranging from 40 to 120 meters. Deeper control borings were also carried out.

Grouting to consolidate the foundation rock was performed across the entire dam footprint to a depth of 8 meters, using a square grid with 4-meter spacing.

A drainage curtain with a maximum depth of 40 meters was constructed from the inspection gallery at the dam's base.

The aggregates for the RCC (Roller-Compacted Concrete) were sourced from the processing of alluvial deposits in the nearby wadi. ^[21]

II.2.1.2 Main characteristics of the management

Reservoir ^[21]

Table 1. Main characteristics of the reservoir

Normal Level (NL)	200.00 m
Maximum Water Level (HWL)	214.80 m
Minimum Operating Level (MOL)	172.00 m
Dead Storage Volume (110-172)	240 10 ⁶ m ³
Useful Storage Volume (172 - 200)	723 10 ⁶ m ³
Total Storage Volume (110-200)	963 10 ⁶ m ³

Dam ^[21]

Table 2. Main characteristics of the dam

Type	Roller-Compacted Concrete (RCC) Gravity Dam
Maximum Height:	
- Above the natural ground	107 m
- Above the foundations	118 m
Crest Length	710 m
Crest Width	8 m
Maximum Width at the function level	93 m
Upstream face slope	Vertical
Downstream face slope	08H/1V
Crest Elevation	216.30 m

Flood spillway ^[21]

Table 3. Main characteristics of the flood spillway

Type	Free-Flow Surface Spillway - Channel and Ski-Jump
Location	Central Part of the Dam
Crest Elevation	200 m
Elevation of the downstream ski-jump lip	Ranging from 124.00 m to 126.00 m
Total Length of the Spillway at the Crest	124 m
Effective Length of the Spillway at the Crest	114 m distributed over 6 spans of 19 m each
Length of the Bucket	130 m
Maximum Flow	13,230 m ³ /sec for HWL at 214.74 m

Half-Depth Drainage ^[21]

Table 4. Main characteristics of the half-depth drainage

Type	2 armored outlets in the body of the dam with ski-jump
Location	Right bank
Entrance Positioning	140.0 m
Section of the Outlets	(3.0 m x 4.10 m) x 2
Gates	2 sluice gates upstream 2 segment gates downstream
Maximum Flow	670 m ³ /sec below NL 200 > 700 m ³ /sec below NL > 200

Injection and drainage of the foundation ^[21]

Table 5. Main characteristics of the injection and drainage of the foundation

Injection curtain	Vertical multifilar aligned with the upstream footer gallery, depth from 40 to 120 m
Drainage curtain	1 drilled curtain from the upstream footer gallery

Diversion structure ^[21]

Table 6. Main characteristics of the diversion structure

2 underground galleries	Ø 8 m (solution galleries "rock-fill embankment")
-------------------------	---

II.3 Conclusion

Chapter II provides a detailed and technical overview of the Beni Haroun Dam, positioning it as one of Algeria's most significant hydraulic infrastructures. Strategically located in the Wilaya of Mila and serving several eastern Algerian provinces, the dam plays a vital role in water storage, distribution, and inter-regional transfer, particularly addressing the imbalanced demand between the northern and southern regions.

The geological characterization of the dam site reveals a complex stratigraphy dominated by Eocene limestones and Paleocene marls, intersected by numerous fractures and karstic features. This heterogeneity demands careful engineering to ensure foundation stability. The presence of tectonic deformation, weathered materials, and varying lithological units illustrates the geotechnical challenges addressed during design and construction.

The dam itself is a straight gravity structure made of roller-compacted concrete (RCC) with a maximum height of 118 meters and a crest length of 710 meters. The technical specifications reflect advanced engineering tailored to optimize durability, minimize conventional concrete use, and enhance seismic resistance. Key safety components, such as the multi-bay spillway, ski-jump discharge, bottom outlets, and drainage galleries, demonstrate the dam's robust flood management capacity.

Detailed operational parameters are also highlighted, including normal and maximum water levels, reservoir volumes, and outlet capacities, which are essential for understanding the dam's performance under various hydrological conditions. The RCC structure is supported by a grout curtain and drainage systems that ensure effective seepage control and structural integrity.

What stands out is the dam's dual role: it is both a vital water resource infrastructure and a critical point of vulnerability. Its capacity to withstand high inflow events—especially amid changing climate conditions—is central to regional water security and flood protection. As such, the chapter emphasizes not only the engineering prowess behind Beni Haroun Dam but also the importance of continuous monitoring and proactive management to adapt to future environmental challenges.

Moreover, climatic variability—amplified by the global trend of climate change—is emerging as a significant challenge. The recurrence of intense rainfall events and shifting seasonal patterns raises the stakes for dam safety and water resource management. In dry years, low inflows could jeopardize the dam's primary purpose of supplying water to multiple wilayas. In wet years, unanticipated surges in runoff could overwhelm spillway and drainage systems if not properly anticipated and managed.

Therefore, while Beni Haroun Dam stands as a robust engineering feat, its performance and long-term sustainability are intimately tied to hydrological, meteorological, and climatic factors. These factors must be continuously monitored.

Chapter III

Machine Learning, Modeling Approaches, Data Preparation and modeling

III.1 Machine Learning

The basic concept of machine learning in data science involves using statistical learning and optimization methods that let computers analyze datasets and identify patterns. Machine learning techniques leverage data mining to identify historic trends and inform future models. ^[22]

The typical supervised machine learning algorithm consists of roughly three components:

1. **A decision process:** A recipe of calculations or other steps that takes in the data and “guesses” what kind of pattern your algorithm is looking to find.
2. **An error function:** A method of measuring how good the guess was by comparing it to known examples (when they are available). Did the decision process get it right? If not, how do you quantify “how bad” the miss was?
3. **An updating or optimization process:** A method in which the algorithm looks at the miss and then updates how the decision process comes to the final decision, so next time the miss won’t be as great.

For example, if you’re building a movie recommendation system, you can provide information about yourself and your watch history as input. The algorithm will take that input and learn how to return an accurate output: movies you will enjoy. Some inputs could be movies you watched and rated highly, the percentage of movies you’ve seen that are comedies, or how many movies feature a particular actor. The algorithm’s job is to find these parameters and assign weights to them. If the algorithm gets it right, the weights it used stay the same. If it gets a movie wrong, the weights that led to the wrong decision get turned down so it doesn’t make that kind of mistake again.

Since a machine learning algorithm updates autonomously, the analytical accuracy improves with each run as it teaches itself from the data it analyzes. This iterative nature of learning is both unique and valuable because it occurs without human intervention — empowering the algorithm to uncover hidden insights without being specifically programmed to do so. ^[22]

III.1.2 Types of Machine Learning

There are many types of machine learning models defined by the presence or absence of human influence on raw data — whether a reward is offered, specific feedback is given, or labels are used.

- **Supervised learning:** The dataset being used has been pre-labeled and classified by users to allow the algorithm to see how accurate its performance is.
- **Unsupervised learning:** The raw dataset being used is unlabeled and an algorithm identifies patterns and relationships within the data without help from users.

- **Semi-supervised learning:** The dataset contains structured and unstructured data, which guides the algorithm on its way to making independent conclusions. The combination of the two data types in one training dataset allows machine learning algorithms to learn to label unlabeled data.
- **Reinforcement learning:** The dataset uses a “rewards/punishments” system, offering feedback to the algorithm to learn from its own experiences by trial and error.

Finally, there’s the concept of *deep learning*, which is a newer area of machine learning that automatically learns from datasets without introducing human rules or knowledge. This requires massive amounts of raw data for processing — and the more data that is received, the more the predictive model improves. ^[22]

III.1.3 Commonly used Machine Learning algorithms

The purpose of machine learning is to use machine learning algorithms to analyze data. By leveraging machine learning, a developer can improve the efficiency of a task involving large quantities of data without the need for manual human input. Around the world, strong machine learning algorithms can be used to improve the productivity of professionals working in data science, computer science, and many other fields.

There are a number of machine learning algorithms that are commonly used by modern technology companies. Each of these machine learning algorithms can have numerous applications in a variety of educational and business settings. ^[22]

III.1.3.1 Linear regression

Linear regression is an algorithm used to analyze the relationship between independent input variables and at least one target variable. This kind of regression is used to predict continuous outcomes — variables that can take any numerical outcome. For example, given data on the neighborhood and property, can a model predict the sale value of a home? Linear relationships occur when the data relationship being observed tends to follow a straight line overall — and as such, this model can be used to observe whether a data point is increasing, decreasing, or remaining the same relative to some independent variable, such as time elapsed or position.

Machine learning models can be employed to analyze data in order to observe and map linear regressions. Independent variables and target variables can be input into a linear regression machine learning model, and the model will then map the coefficients of the best fit line to the data. In other words, the linear regression models attempt to map a straight line, or a linear relationship, through the dataset. ^[22]

III.1.3.2 Logistic regression

Logistic regression is a supervised learning algorithm that is used for classification problems. Instead of continuous output like in linear regression, a logistic model predicts the probability of a binary event occurring. For example in dam risk management, logistic regression can be used to classify whether a specific inflow event will lead to a flood risk or not.

For instance, based on input variables such as:

- Monthly inflow (m^3/s)
- Maximum daily rainfall (mm)
- End-of-month reservoir volume (hm^3)
- Volume change compared to previous month (hm^3)

The logistic regression model can be trained to predict the probability that the dam will enter a "flood alert" status (1) or remain in "normal operation" (0).

So the output is binary:

- 1 → Flood risk present
- 0 → No flood risk

The dam operator can then use this prediction to trigger early warning systems or initiate preventive release protocols.

Machine learning algorithms can use logistic regression models to determine categorical outcomes. When given a dataset, the logistic regression model can check any weights and biases and then use the given dependent categorical target variables to understand how to correctly categorize that dataset. ^[22]

III.1.3.3 Neural networks

Neural networks are artificial intelligence algorithms that attempt to replicate the way the human brain processes information to understand and intelligently classify data. These neural network learning algorithms are used to recognize patterns in data and speech, translate languages, make financial predictions, and much more through thousands, or sometimes millions, of interconnected processing nodes. Data is “fed-forward” through layers that process and assign weights, before being sent to the next layer of nodes, and so on.

Crucially, neural network algorithms are designed to quickly learn from input training data in order to improve the proficiency and efficiency of the network’s algorithms. As such, neural networks serve as key examples of the power and potential of machine learning models. ^[22]

III.1.3.4 Decision trees

Decision trees are data structures with nodes that are used to test against some input data. The input data is tested against the leaf nodes down the tree to attempt to produce the correct, desired output. They are easy to visually understand due to their tree-like structure and can be designed to categorize data based on some categorization schema.

Decision trees are one method of supervised learning, a field in machine learning that refers to how the predictive machine learning model is devised via the training of a learning algorithm. ^[22]

III.2 Modeling approaches

III.2.1 RF (Random Forest)

Random forest is a commonly-used machine learning algorithm, trademarked by Leo Breiman and Adele Cutler, that combines the output of multiple decision trees to reach a single result. Its ease of use and flexibility have fueled its adoption, as it handles both classification and regression problems. ^[23]

III.2.1.1 Random Forest works

Random forest algorithms have three main hyperparameters, which need to be set before training. These include node size, the number of trees, and the number of features sampled. From there, the random forest classifier can be used to solve for regression or classification problems.

The random forest algorithm is made up of a collection of decision trees, and each tree in the ensemble is comprised of a data sample drawn from a training set with replacement, called the bootstrap sample. Of that training sample, one-third of it is set aside as test data, known as the out-of-bag (oob) sample, which we'll come back to later. Another instance of randomness is then injected through feature bagging, adding more diversity to the dataset and reducing the correlation among decision trees. Depending on the type of problem, the determination of the prediction will vary. For a regression task, the individual decision trees will be averaged, and for a classification task, a majority vote—i.e. the most frequent categorical variable—will yield the predicted class. Finally, the oob sample is then used for cross-validation, finalizing that prediction.

[23]

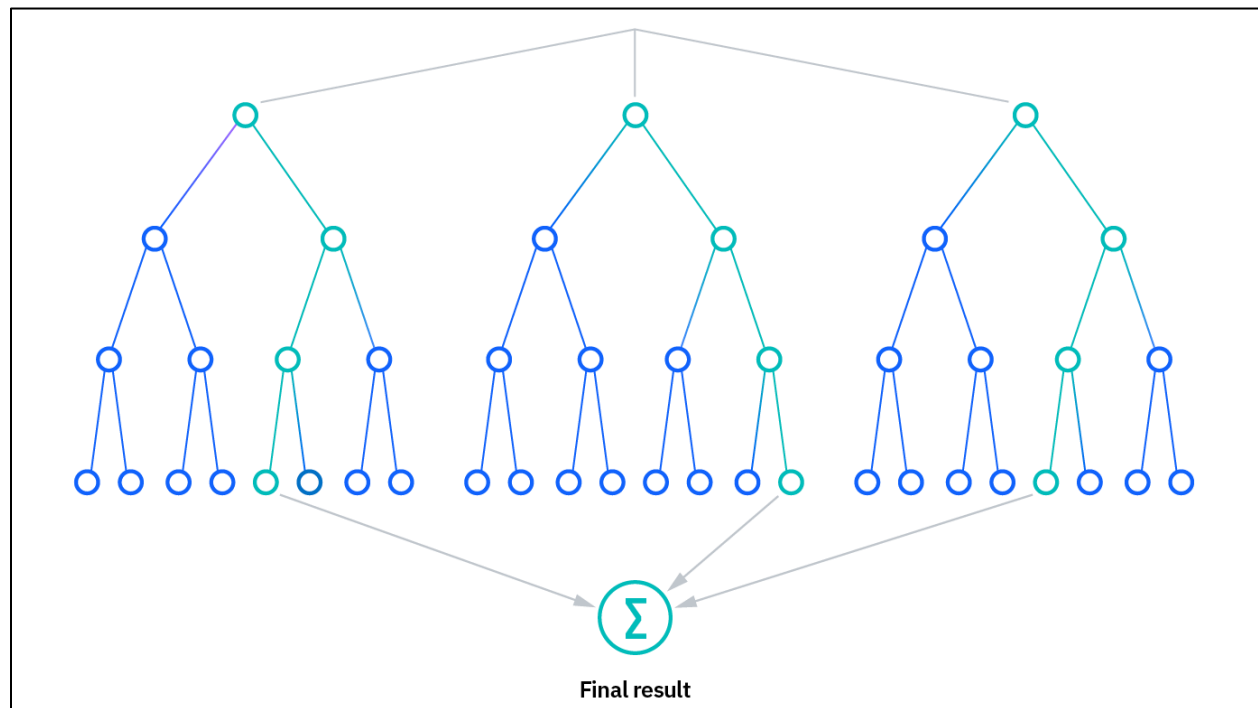


Figure 5. Random Forest diagram [23]

III.2.2 XGBoost

XGBoost short form for eXtreme Gradient Boosting is an advanced machine learning algorithm designed for efficiency, speed and high performance.

It is an optimized implementation of Gradient Boosting and is a type of ensemble learning method that combines multiple weak models to form a stronger model. ^[24]

- XGBoost uses decision trees as its base learners and combines them sequentially to improve the model's performance. Each new tree is trained to correct the errors made by the previous tree and this process is called boosting.
- It has built-in parallel processing to train models on large datasets quickly. XGBoost also supports customizations allowing users to adjust model parameters to optimize performance based on the specific problem.

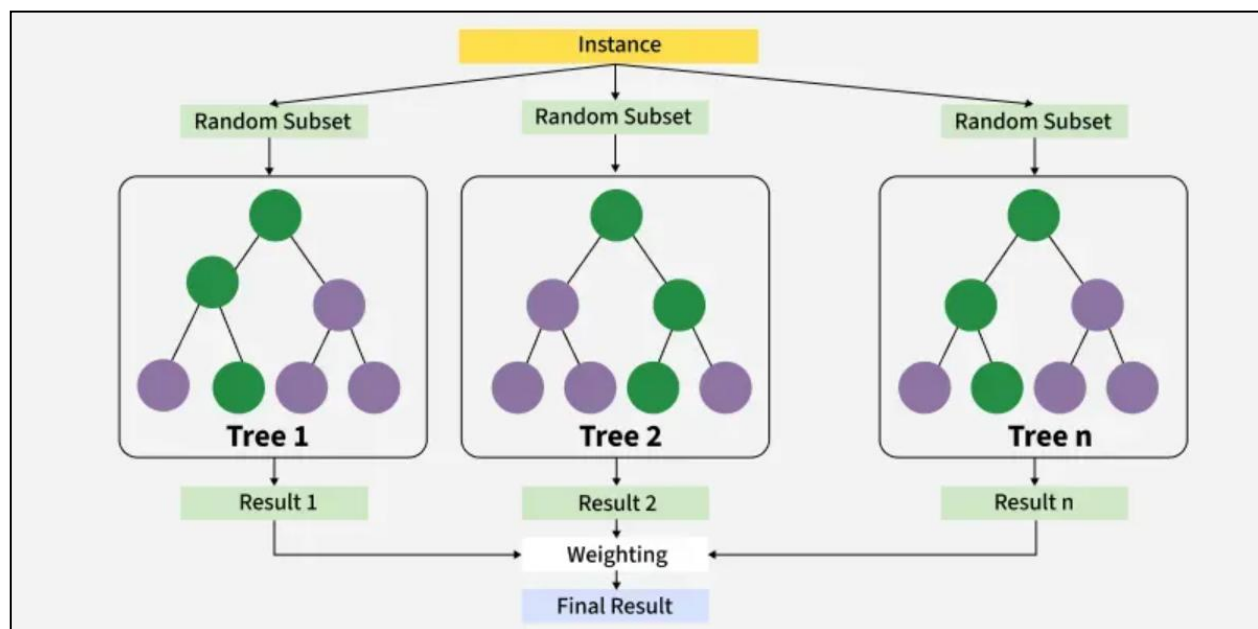


Figure 6. XGBoost diagram ^[24]

III.2.2.1 XGBoost Works

It builds decision trees sequentially with each tree attempting to correct the mistakes made by the previous one. The process can be broken down as follows: ^[24]

1. **Start with a base learner:** The first model decision tree is trained on the data. In regression tasks this base model simply predicts the average of the target variable.
2. **Calculate the errors:** After training the first tree the errors between the predicted and actual values are calculated.
3. **Train the next tree:** The next tree is trained on the errors of the previous tree. This step attempts to correct the errors made by the first tree.
4. **Repeat the process:** This process continues with each new tree trying to correct the errors of the previous trees until a stopping criterion is met.
5. **Combine the predictions:** The final prediction is the sum of the predictions from all the trees.

III.2.3 Gradient Boosting

Gradient boosting is an ensemble learning algorithm that produces accurate predictions by combining multiple decision trees into a single model. This algorithmic approach to predictive modeling, introduced by Jerome Friedman, uses base models to build upon their strengths, correcting errors and improving predictive capabilities. By capturing complex patterns in data, gradient boosting excels at diverse predictive modeling tasks. ^[25]

III.2.3.1 Ensemble learning and boosting

Ensemble learning is a machine learning approach that combines multiple models or methods to boost predictive performance. It often employs techniques such as bagging and boosting. Bagging involves training numerous models on different data subsets with some randomness, which helps reduce variance by averaging out individual errors. A great example of this approach is random forests.

In contrast, boosting is an ensemble technique that iteratively trains models to correct previous mistakes. It gives more weight to misclassified instances in subsequent models, allowing them to focus on challenging data points and ultimately enhancing overall performance. AdaBoost, widely regarded as the first applicable boosting algorithm, is a classic illustration of this method. Both bagging and boosting optimize the bias variance tradeoff in models, leading to more robust performance.

These techniques are extensively used in machine learning to improve model accuracy, especially when dealing with complex or noisy datasets. By combining multiple perspectives, ensemble learning provides a way to overcome the limitations of individual models and achieve improved optimization. ^[25]

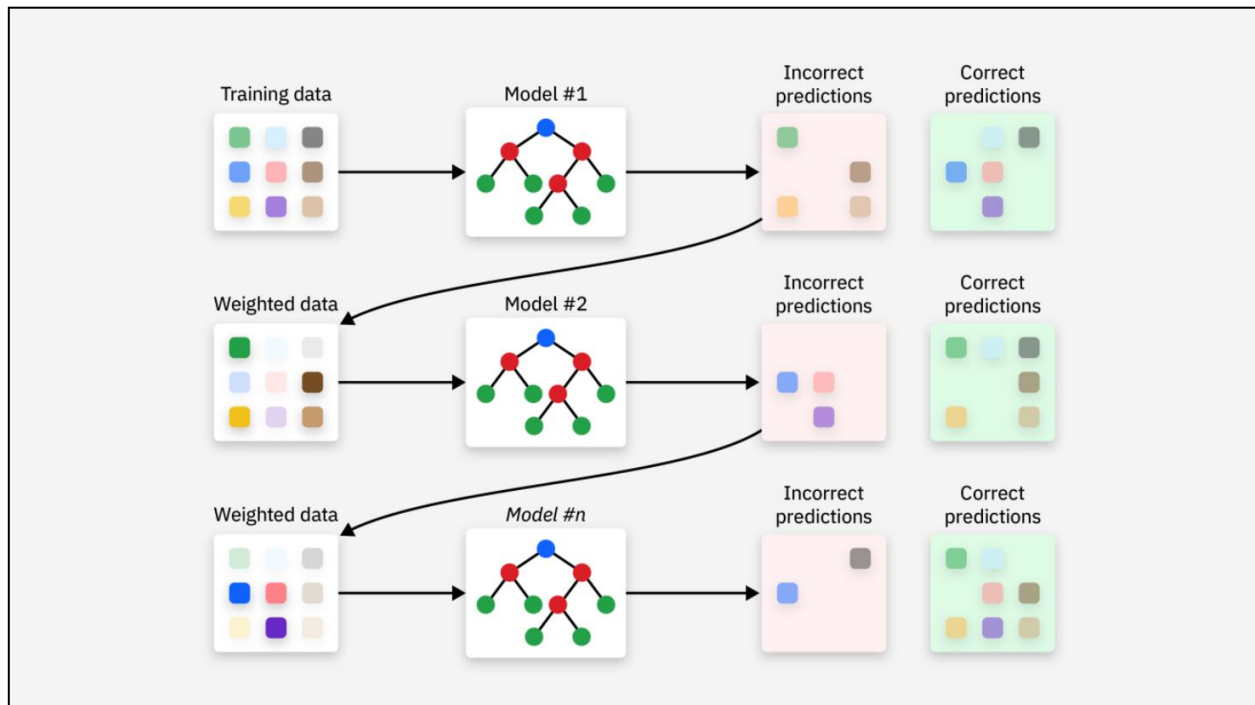


Figure 7. Gradient Boosting diagram ^[25]

III.2.4.2 Gradient Boosting works

Gradient boosting is a machine learning technique that combines multiple weak prediction models into a single ensemble. These weak models are typically decision trees, which are trained sequentially to minimize errors and improve accuracy. By combining multiple decision tree regressors or decision tree classifiers, gradient boosting can effectively capture complex relationships between features. ^[25]

III.2.4 CatBoost

CatBoost, short for “Category Boosting” is an open-source gradient boosting library developed by Yandex. It’s specifically designed to work with categorical variables straight out of the box. Unlike other machine learning algorithms that require categorical variables to be converted into numerical format through one-hot encoding or similar techniques, CatBoost can process these variables natively, which significantly simplifies the data preparation process and enhances model performance. ^[26]

III.2.4.1 CatBoost algorithm works

CatBoost is a supervised machine learning method that is used by the Train Using AutoML tool and uses decision trees for classification and regression. As its name suggests, CatBoost has two main features, it works with categorical data (the Cat) and it uses gradient boosting (the Boost). Gradient boosting is a process in which many decision trees are constructed iteratively. Each subsequent tree improves the result of the previous tree, leading to better results. CatBoost improves on the original gradient boost method for a faster implementation.

CatBoost overcomes a limitation of other decision tree-based methods in which, typically, the data must be pre-processed to convert categorical string variables to numerical values, one-hot-encodings, and so on. This method can directly consume a combination of categorical and non-categorical explanatory variables without preprocessing. It preprocesses as part of the algorithm. CatBoost uses a method called ordered encoding to encode categorical features. Ordered encoding considers the target statistics from all the rows prior to a data point to calculate a value to replace the categorical feature.

Another unique characteristic of CatBoost is that it uses symmetric trees. This means that at every depth level, all the decision nodes use the same split condition.

CatBoost can also be faster than other methods such as XGBoost. It retains certain features—such as cross-validation, regularization, and missing value support—from the prior algorithms. This method performs well with both small data and large data. ^[27]

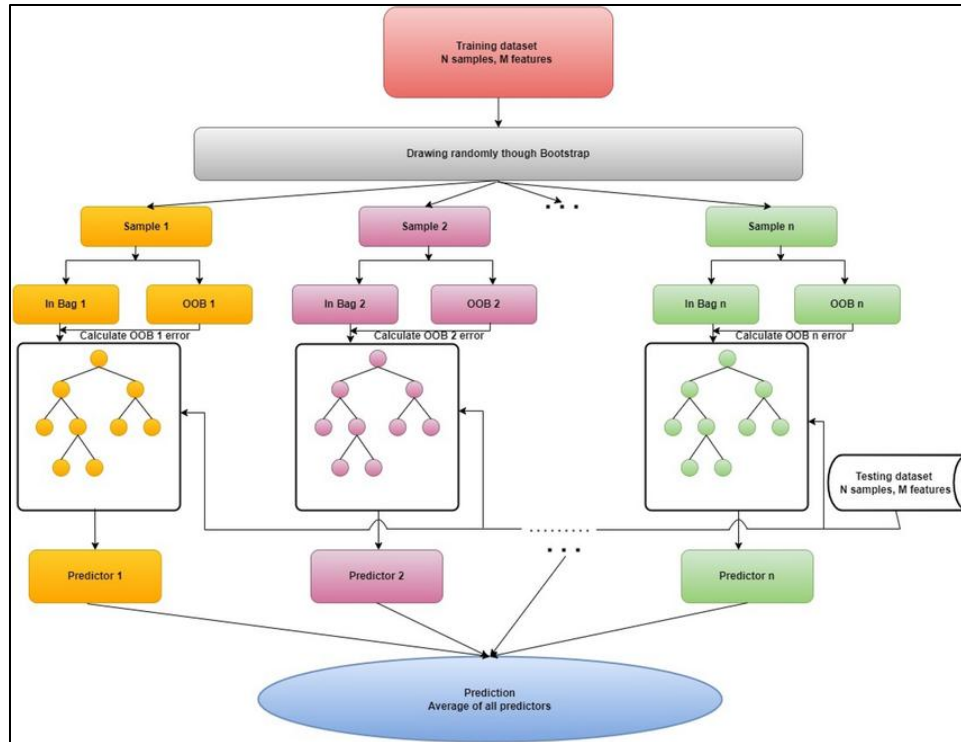


Figure 8. Structure of CatBoost algorithm ^[28]

III.3 Data Preparation for modeling

III.3.1 Historical floodwater release data ^[29]

Table 7. Historical floodwater release data, spill (hm³)

Months/Years	2009	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2024
January	0.00	0.00	0.00	122.33	419.86	106.69	0.00	0.00	299.996	106.958	17.782	0.00
February	0.00	292.53	200.93	0.00	673.25	100.40	142.13	0.00	356.24	40.517	58.779	1.79
March	168.33	357.16	259.44	0.00	690.63	278.15	44.72	177.6	183.707	94.601	36.568	303.88
April	0.00	131.89	101.05	121.24	253.26	156.96	2.28	165.628	139.906	84.25	28.69	6.42
May	0.00	56.77	57.25	29.40	116.86	110.06	0.00	119.045	52.253	33.93	0.00	0.00
June	0.00	40.15	6.90	34.70	69.65	1.80	0.00	31.841	13.535	0.00	0.00	0.00
July	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.443	0.00	0.00	0.00
August	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
September	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
October	0.00	9.44	0.00	0.00	28.40	0.00	0.00	29.902	0.00	0.00	0.00	
November	0.00	0.00	110.88	0.00	55.99	0.00	0.00	44.979	108.309	0.00	0.00	
December	0.00	0.00	121.20	0.00	83.96	0.00	0.00	27.617	216.077	0.00	0.00	

III.3.2 End-of-month volume

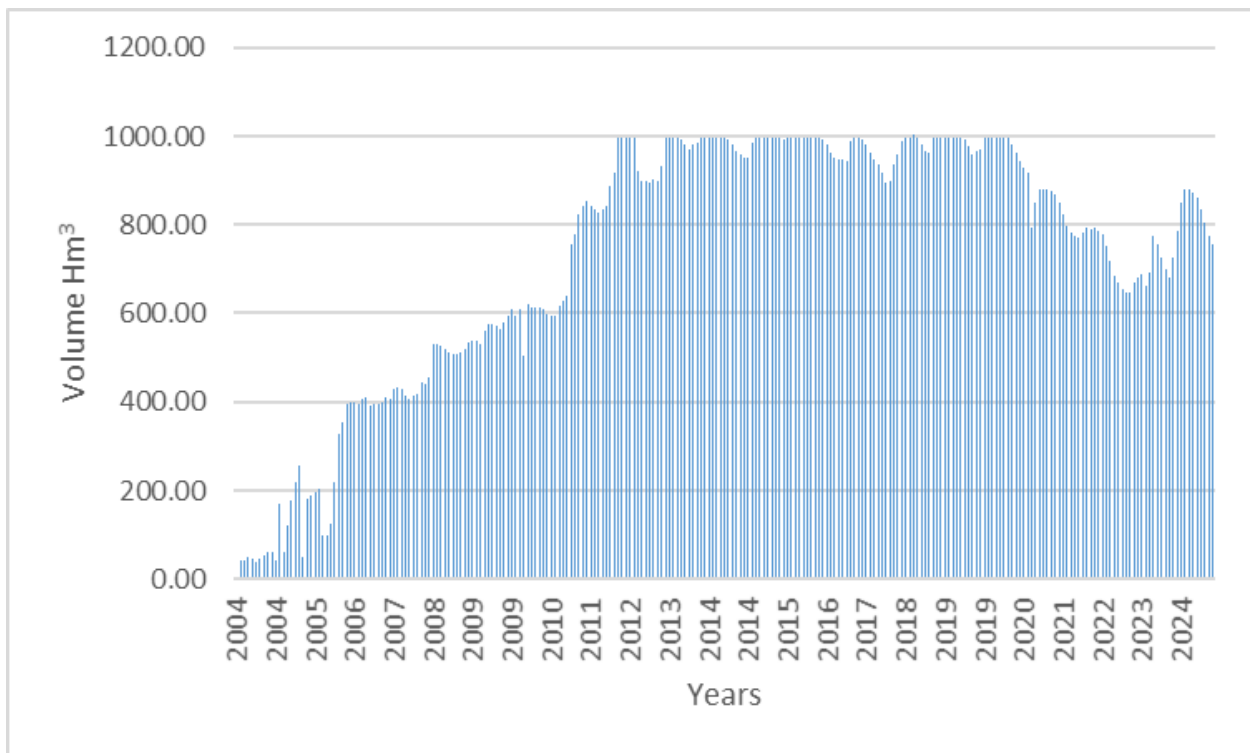


Figure 9. Clustered column of end-of-month volume

III.3.3 Monthly inflow

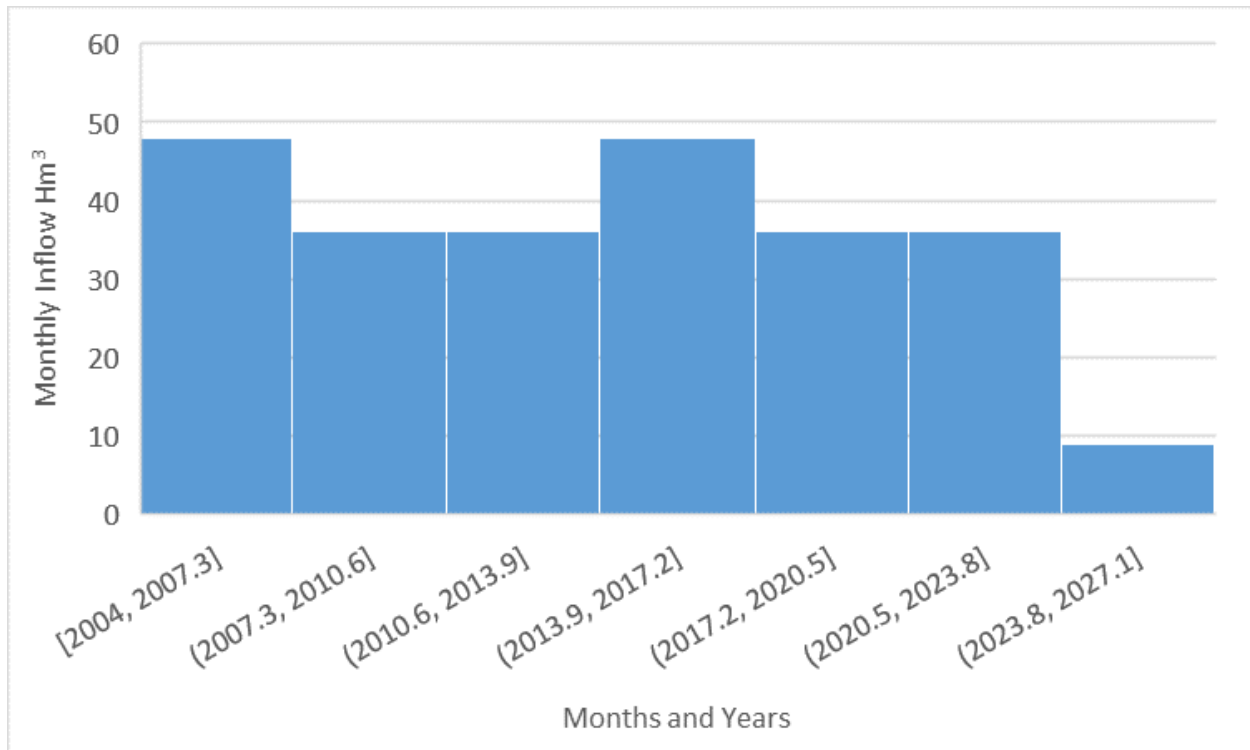


Figure 10. Histogram of monthly inflow

III.3.4 Variable of volume

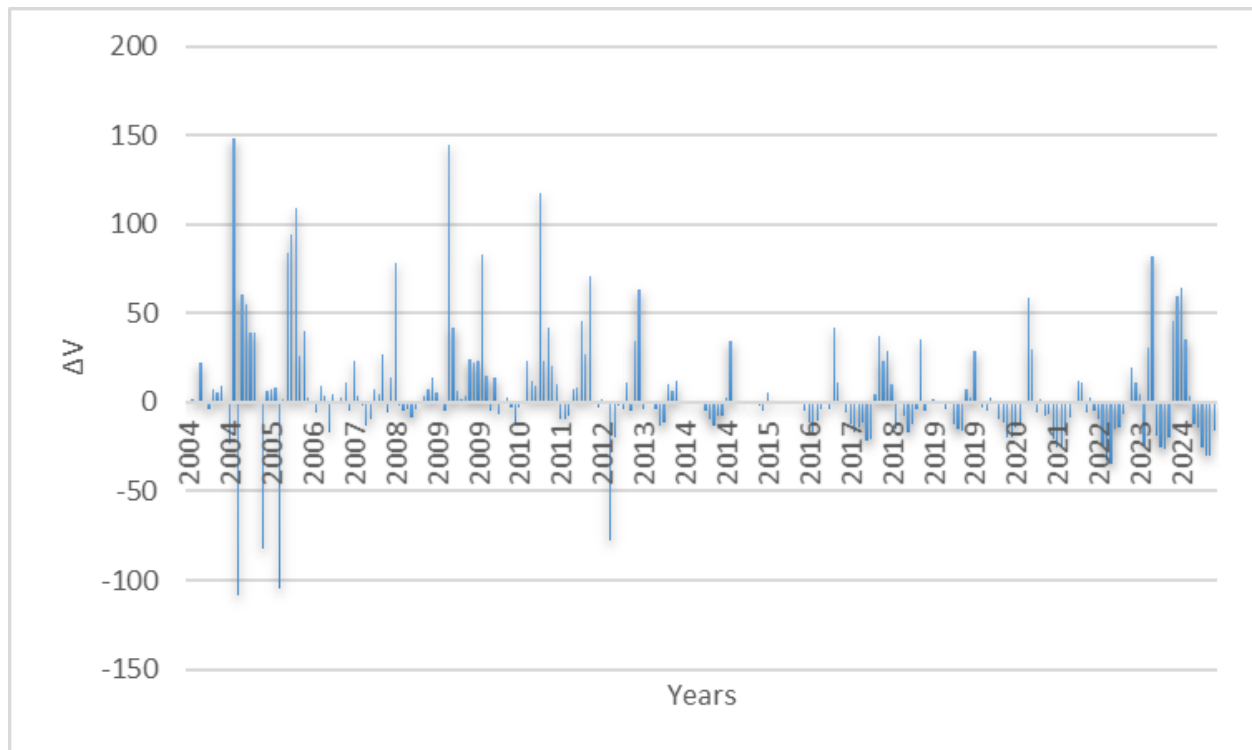


Figure 11. Variable of volume

III.3.5 Max daily rainfall

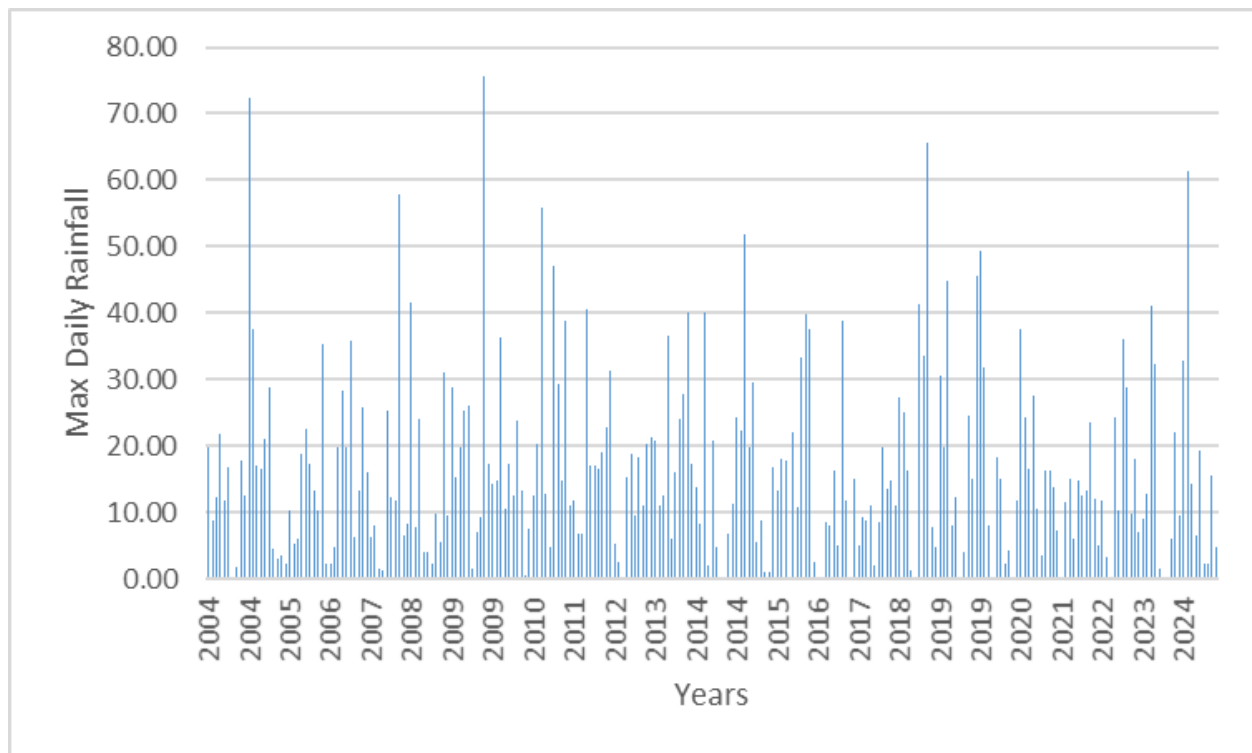


Figure 12. Max daily rainfall

III.4 Modeling

The quality and structure of input data are critical in developing robust machine learning models. In this study, data preparation was conducted using systematic methods to ensure model accuracy, generalizability, and relevance for predicting flood risk at Beni Haroun Dam.

III.4.1 Feature selection and engineering

The initial dataset comprised a wide range of hydrological and operational variables. Based on domain relevance and correlation analysis, the following key features were selected for modeling:

- Monthly Inflow
- Max Daily Rainfall
- Monthly Rainfall
- End of Month Volume
- Volume Change

These variables represent a combination of direct hydrological inputs and reservoir state indicators, both of which influence flood occurrence.

III.4.2 Target variable construction

To frame the task as a classification problem, flood risk labels were created from the variable `Flood_Risk_Combined`, which integrates expert-defined thresholds from two sub-indices: `Flood_Risk_1` and `Flood_Risk_2`. This combined label was converted into three classes (e.g., Low, Medium, High), to serve as the target output in multi-class classification models.

III.4.3 Handling missing or anomalous data

Initial exploratory analysis involved checking for missing values, duplicates, and anomalies. Imputation was not necessary, as the dataset was clean, with no null values. However, extremely low or high values (outliers) in inflow and rainfall were reviewed and retained, considering their physical plausibility and importance in flood scenarios.

III.4.4 Normalization and scaling

Because tree-based ensemble models (Random Forest, XGBoost, CatBoost, Gradient Boosting) are not sensitive to feature scaling, normalization or standardization was not applied. These algorithms naturally handle data in its raw units, preserving interpretability and improving training efficiency.

III.4.5 Data Splitting: training and testing

To evaluate model performance reliably, the dataset was split into training and testing sets, typically using an 80/20 ratio. This ensures that the model learns from one subset and is evaluated on unseen data, simulating real-world deployment.

Additionally, cross-validation (e.g., 5-fold) was applied during model training to reduce overfitting and validate generalizability.

III.4.6 Encoding and labeling

Since the selected features were all numerical, categorical encoding was not required. The target classes were encoded into integers (e.g., Low = 0, Medium = 1, High = 2) for compatibility with classification algorithms.

III.4.7 Performance evaluation criteria

To assess model effectiveness, a combination of classification metrics was used:

- **Accuracy:** Overall correctness of predictions.
- **Precision, Recall, F1-score (macro-averaged):** To evaluate performance across all classes equally.
- **AUC (Area Under the Curve):** Particularly useful for distinguishing high-risk flood events from others.
- **Confusion Matrix:** To visualize class-specific prediction strengths and weaknesses.
- **Feature Importance Scores:** Extracted from models to interpret which variables had the most predictive power.

These metrics provided both a global and class-specific understanding of model performance.

III.5 Conclusion

Chapter III delves into the application of machine learning (ML) techniques in the context of natural disaster management, with a specific focus on flood prediction and dam vulnerability assessment. It begins by laying the groundwork for understanding how ML systems function—using algorithms to learn patterns from data, adapt autonomously, and improve their predictive accuracy over time.

The chapter systematically introduces the four main types of machine learning—supervised, unsupervised, semi-supervised, and reinforcement learning—while also distinguishing deep learning as a powerful subset that requires large datasets and minimal human intervention. These foundational concepts are crucial for appreciating how ML moves beyond traditional, reactive disaster management toward proactive, data-driven strategies.

A single family of widely used machine learning algorithms is explained in accessible terms, including decision tree-based ensemble methods such as Random Forest, XGBoost, Gradient Boosting, and CatBoost, which are known for their interpretability, accuracy, and effectiveness in classification tasks.

The chapter then transitions into a more specialized discussion of modeling techniques used for flood risk prediction, specifically focusing on four high-performing ensemble models:

1. Random Forest (RF)
2. XGBoost (eXtreme Gradient Boosting)
3. Gradient Boosting
4. CatBoost (Category Boosting)

Each method is described in detail—outlining how it builds on decision trees, handles data, improves accuracy, and adapts to classification tasks. Special attention is given to boosting methods, which iteratively focus on correcting previous prediction errors, thereby enhancing the model's performance on complex datasets.

By integrating these models into flood risk analysis, particularly for infrastructures like Beni Haroun Dam, the chapter illustrates how ML empowers engineers and decision-makers to:

- Anticipate high inflow conditions,
- Optimize emergency responses,
- Enhance early warning systems,
- And ultimately reduce the risk of dam failure during extreme weather events.

In parallel, the chapter emphasizes that successful ML implementation hinges not only on algorithm selection but also on rigorous data preparation. This included selecting hydrologically relevant variables—such as monthly inflow, rainfall patterns, and reservoir volume metrics—based on domain knowledge and statistical analysis. The target variable was constructed by combining flood risk indicators into a multi-class label, enabling more granular risk classification.

Given that tree-based models are robust to varying data scales, normalization was not necessary. Instead, efforts were focused on ensuring data integrity, managing outliers, and applying intelligent splitting strategies for model training and evaluation. An 80/20 training-testing split was used alongside k-fold cross-validation to validate model stability. Performance metrics such as accuracy, precision, recall, F1-score, AUC, and confusion matrices provided a comprehensive evaluation framework, while feature importance scores offered insight into the hydrological drivers of flood risk.

In essence, Chapter III shows that effective flood prediction is achieved through the synergy of advanced machine learning models and well-prepared, contextually meaningful data. This dual emphasis ensures that predictive outcomes are both technically sound and practically useful for dam safety and disaster risk reduction.

Chapter IV

Results, discussion and Comparison

IV.1 Predictive model outcomes

IV.1.1 Random Forest (RF)

Among the models used, Random Forest served as a strong starting point for flood risk classification due to its reliability, ability to handle complex feature interactions, and resistance to overfitting. The model works by constructing a large number of decision trees and combining their outputs to improve predictive accuracy. Before training the Random Forest model, it was essential to first verify that the input variables did not exhibit problematic multicollinearity, which could distort the results. This was done through an analysis of Variance Inflation Factor (VIF) and tolerance values.

IV.1.1.1 VIF and tolerance statistics

Table 8. VIF and tolerance statistics of RF

Variable	VIF	Tolerance
Monthly Rainfall	3.830721	0.261047
Max Daily Rainfall	2.976299	0.335988
Monthly Inflow	1.358036	0.736358
Volume Change	1.292619	0.773623
End of Month Volume	1.042016	0.959678

Before model training, input variables were assessed for multicollinearity using Variance Inflation Factor (VIF) and tolerance values. All variables showed acceptable levels of collinearity, with no VIF exceeding the commonly accepted threshold of 5. Monthly Rainfall exhibited the highest VIF (3.83), while End of Month Volume had the lowest (1.04), indicating no significant redundancy or distortion among predictors. These results confirm the statistical reliability of the input data used to predict flood-related inflow behavior.

IV.1.1.2 Model metrics

Accuracy : 0.92
Precision (macro): 0.87
Recall (macro) : 0.88
F1-score (macro) : 0.88
AUC (macro) : 0.98

The Random Forest model achieved an overall accuracy of 92%, indicating a high level of predictive reliability. This level of accuracy suggests that the model is well-suited for forecasting inflow levels that contribute to flood risk at Beni Haroun Dam. Furthermore, a macro-averaged Area Under the Curve (AUC) value of 0.98 demonstrates the model's excellent ability to distinguish between different inflow conditions, including potentially hazardous ones.

IV.1.1.3 Classification report

Table 9. Classification report of RF

	Precision	Recall	F1-score	Support
Low	0.87	0.81	0.84	16
Medium	0.75	0.86	0.80	14
High	1.00	0.98	0.99	45

Class-specific performance results revealed that the model was particularly strong in detecting high inflow conditions—the most critical category for dam vulnerability assessment. The high inflow class was predicted with near-perfect scores, achieving a precision of 1.00, recall of 0.98, and F1-score of 0.99. For the medium inflow class, the model maintained a strong recall of 0.86 despite a slightly lower precision of 0.75, yielding an F1-score of 0.80. The low inflow class also showed solid results, with a precision of 0.87, recall of 0.81, and F1-score of 0.84. These outcomes highlight the model's ability to reliably identify extreme inflow events, which are directly associated with periods of elevated flood risk and operational stress on the dam.

IV.1.1.4 Confusion matrix

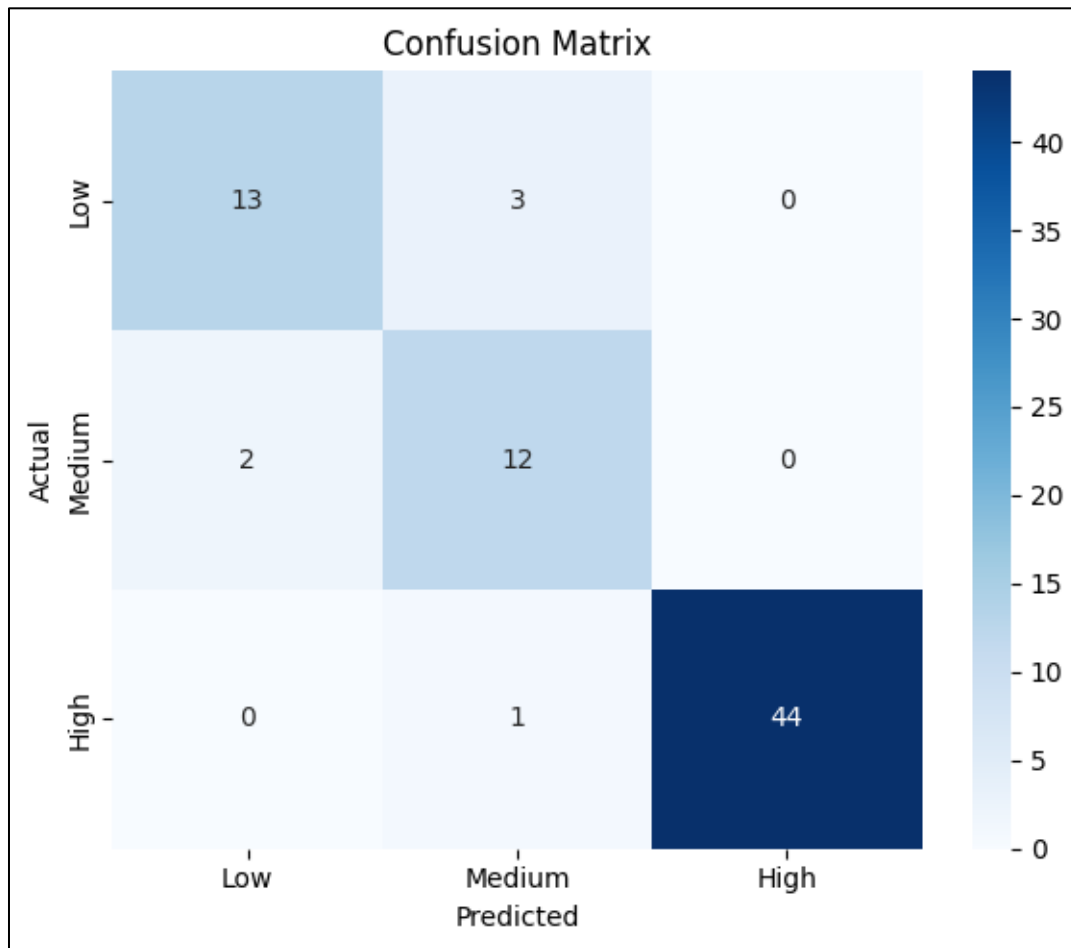


Figure 13. Confusion matrix of RF

The confusion matrix provides a summary of the classification performance of the Random Forest model by comparing actual inflow classes with those predicted by the model. In this matrix, each row represents the actual class, while each column represents the predicted class.

The RF model showed particularly strong performance in predicting high inflow levels, with most actual high inflow instances correctly classified. A small number of misclassifications occurred between the medium and low inflow classes, indicating some overlap in the feature patterns of these two categories. However, the overall structure of the matrix demonstrates a strong alignment between actual and predicted values, particularly for the high-risk (flood-prone) category, which is most critical for dam vulnerability analysis.

IV.1.1.5 Normalized confusion matrix

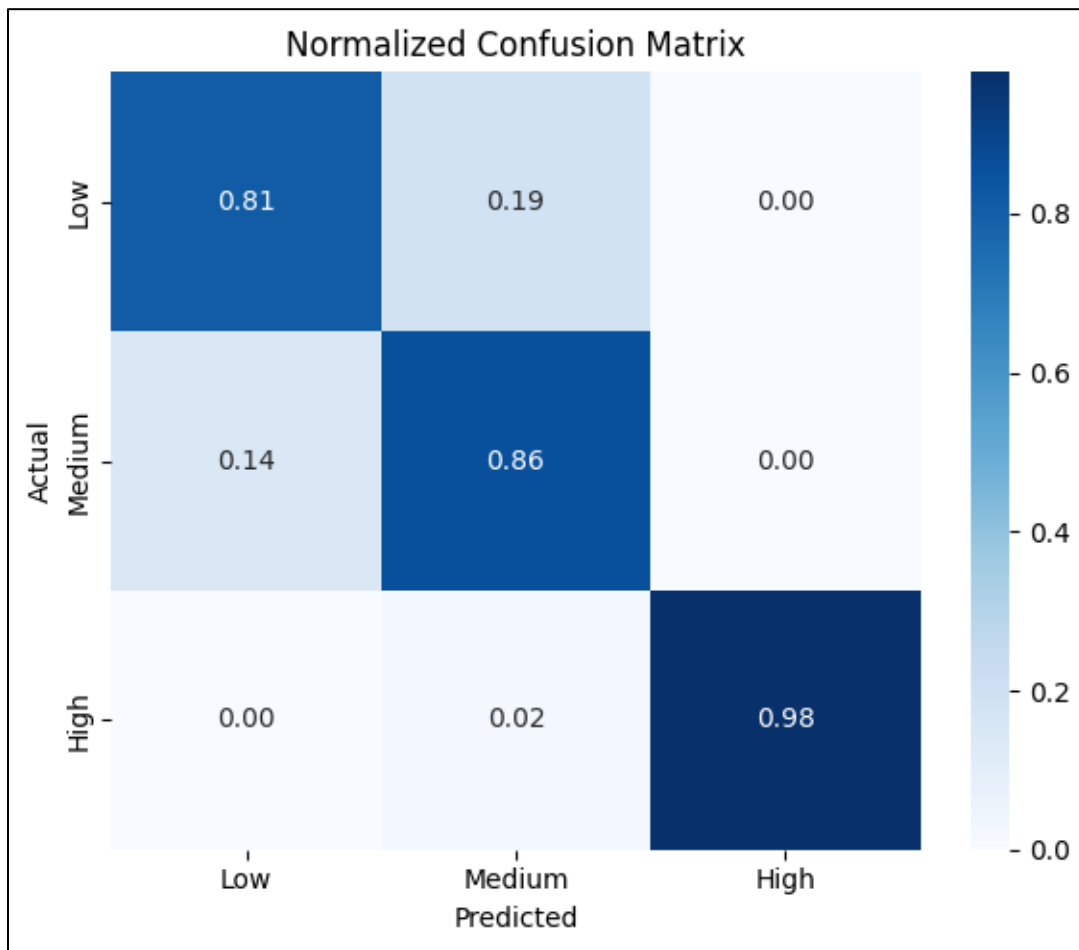


Figure 14. Normalized confusion matrix of RF

The normalized confusion matrix refines the raw confusion matrix by expressing values as proportions or percentages, allowing for a clearer comparison of classification effectiveness across classes regardless of sample size.

From the normalized matrix, it is evident that the high inflow class achieved a near-perfect classification rate, with over 98% of high inflow cases correctly identified by the model. The low inflow class also showed a high true positive rate, while the medium inflow class had a slightly lower precision due to some confusion with the low inflow class. This suggests that while the model is highly effective in distinguishing flood-risk conditions (high inflows), it is slightly more challenged by transitional inflow states that lack clear distinguishing features.

IV.1.1.6 ROC curve – multi-class

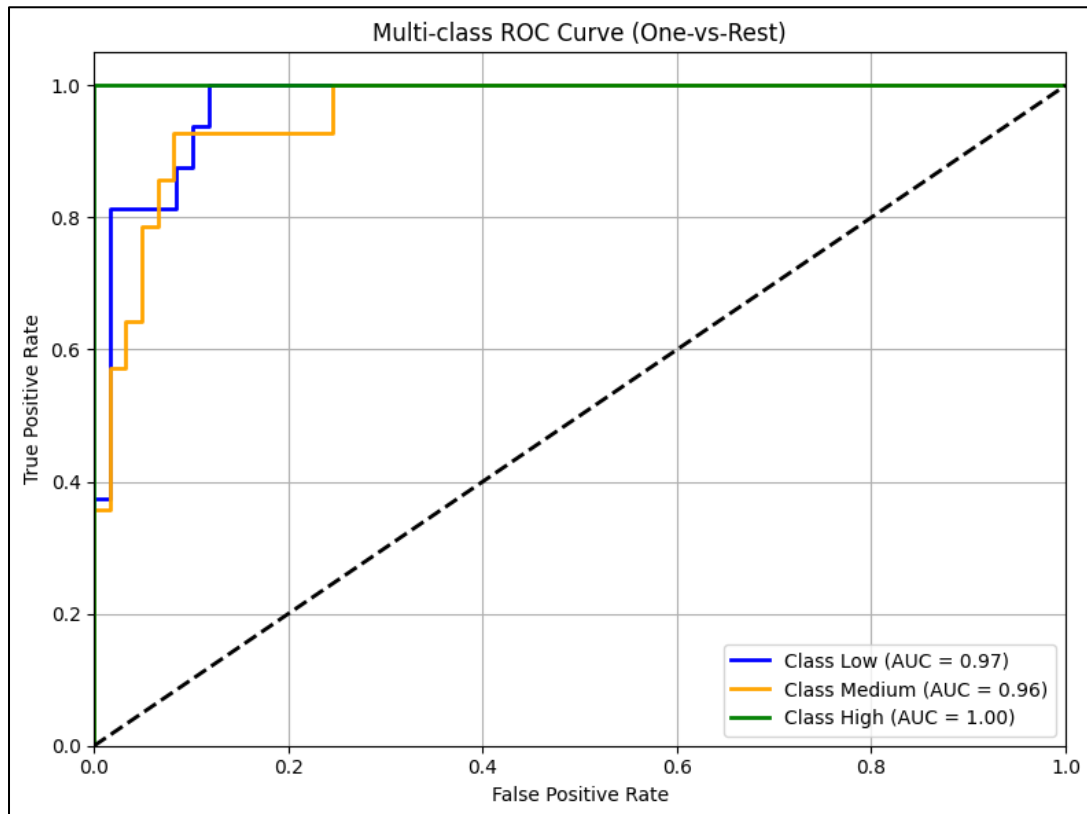


Figure 15. Multi-class ROC curve of RF

The multi-class ROC (Receiver Operating Characteristic) curve for the Random Forest model offers a clear visualization of the classifier's ability to distinguish between the three inflow categories: low, medium, and high. This is achieved using the One-vs-Rest (OvR) strategy, in which a separate ROC curve is plotted for each class by comparing it against the combined set of the other two classes.

The ROC analysis confirms that the Random Forest model performs exceptionally well in multi-class classification, particularly in the context of flood risk prediction at Beni Haroun Dam:

- The high inflow class—the most crucial category for dam safety and flood preparedness—achieved an AUC (Area Under the Curve) value of 1. This indicates that the model has a perfect ability to distinguish flood-prone inflow conditions from normal or low-risk scenarios. Such performance is vital for triggering timely warnings and preventive actions during periods of elevated hydrological stress.
- The medium inflow class also exhibited a high AUC, demonstrating the model's competence in handling transitional inflow cases. Although this class is often more difficult

to classify due to overlaps with neighboring classes, the Random Forest model maintained a strong true positive rate while limiting false positives.

- The low inflow class likewise showed a strong AUC, confirming the model's reliability in identifying dry or non-flood conditions. Accurate identification of these periods is also important for optimizing water storage and release operations.

The ROC curves for all three classes were consistently skewed toward the top-left corner of the plot, which reflects a combination of high sensitivity and specificity. This shape indicates that the model is not only effective at correctly classifying each inflow category, but also maintains a low rate of false alarms—particularly important in operational dam settings where overreaction can lead to unnecessary releases, while underreaction may result in flood damage.

IV.1.1.7 Feature importance

Table 10. Variable and importance of RF

Variable	Importance
Monthly Inflow	0.495830
Volume Change	0.179827
End of Month Volume	0.130346
Monthly Rainfall	0.102944
Max Daily Rainfall	0.091053

An analysis of feature importance showed that Monthly Inflow was the most influential predictor, contributing nearly 50% of the model’s total decision-making weight. Volume Change and End of Month Volume were the next most important variables, accounting for 17.98% and 13.03% respectively. Meanwhile, rainfall variables—including Monthly Rainfall and Max Daily Rainfall—played a secondary yet relevant role. This ranking reflects the physical reality of flood generation and reservoir behavior, underscoring the model’s ability to extract hydrologically meaningful patterns that inform flood vulnerability analysis.

IV.1.1.8 Temporal flood risk map

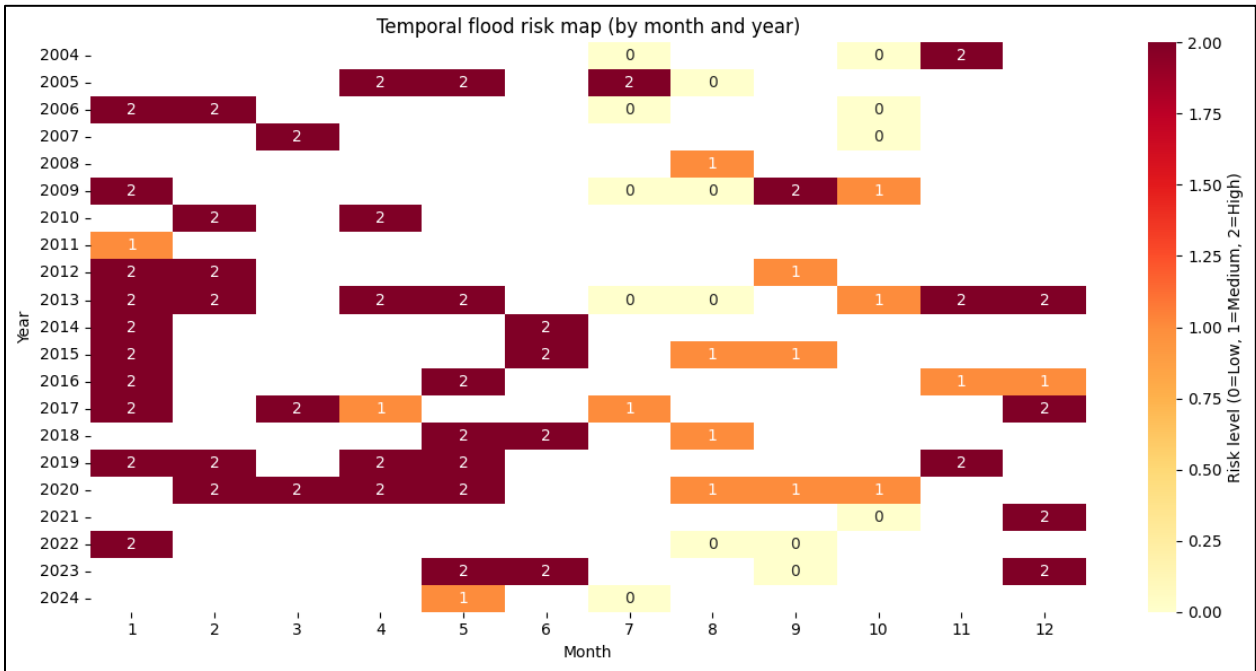


Figure 16. Temporal flood risk map of RF

The temporal risk map generated using Random Forest classification results provides a month-by-month visualization of predicted inflow categories over a multi-year period. This visualization highlights temporal patterns in flood risk and allows stakeholders to track the historical evolution of inflow conditions.

Analysis of the map shows clear seasonal trends, with high inflow levels clustering around the rainy seasons—typically in the winter and spring months. These repeated peaks confirm the model’s ability to capture seasonal flood dynamics. The map serves as a valuable operational tool, allowing dam managers at Beni Haroun to identify recurring high-risk periods and plan preventive actions in advance, such as adjusting storage levels or reinforcing monitoring during vulnerable months.

IV.1.1.9 Annual trend of monthly inflow

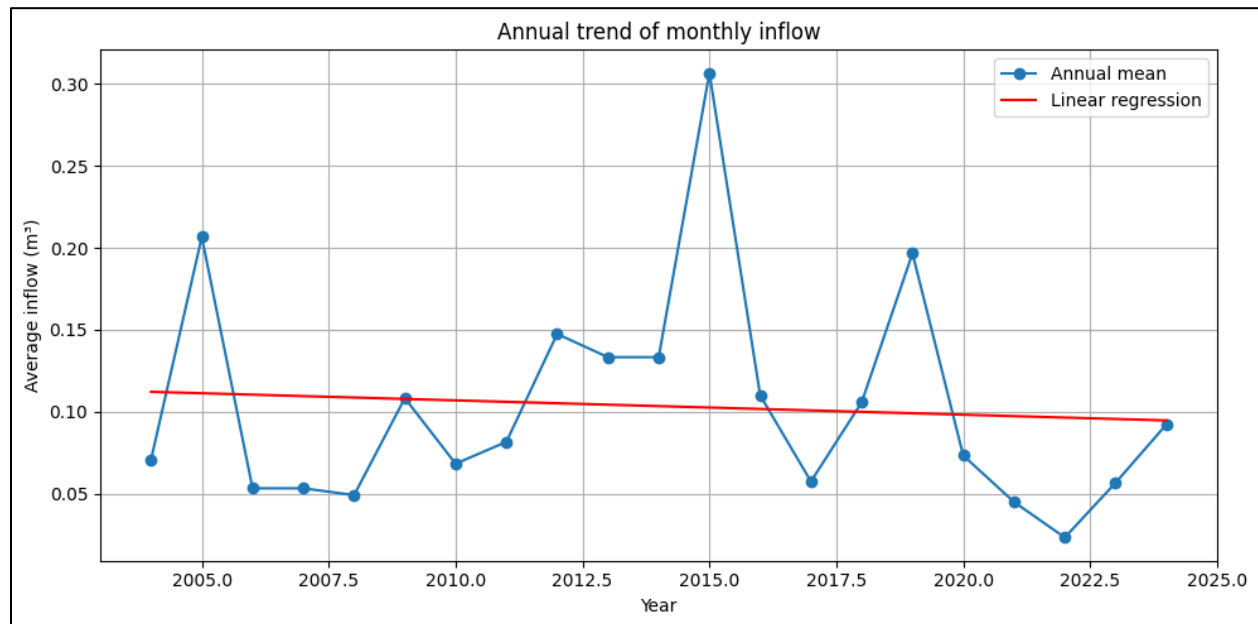


Figure 17. Annual trend of monthly inflow of RF

The annual trend plot illustrates how the classification of inflow levels has varied year by year. This long-term perspective is critical for understanding whether the frequency or intensity of flood-prone conditions is increasing, decreasing, or remaining stable.

The chart indicates that certain years experience significantly more high inflow events, which may correlate with observed regional climatic anomalies or extreme rainfall years. For example, peaks in high inflow frequency during specific years highlight periods of elevated flood risk. This insight supports strategic reservoir planning, risk communication, and investment in adaptive infrastructure. Moreover, the trend chart could serve as a reference point for climate impact studies and long-term water resource management at Beni Haroun Dam.

IV.1.2 XGBoost

Following the Random Forest model, XGBoost was implemented to further explore flood risk classification. Known for its high speed, accuracy, and ability to handle imbalanced or noisy datasets, XGBoost improves model performance by building trees sequentially—each one correcting the errors of the previous. Its efficiency and flexibility make it particularly effective for structured data like hydrological records. As with the previous model, a critical first step involved evaluating the predictor variables for multicollinearity. This was assessed using Variance Inflation Factor (VIF) and tolerance values to ensure that the input data would not introduce redundancy or instability during model training.

IV.1.2.1 VIF and tolerance statistics

Table 11. Vif and tolerance statistics of XGBoost

Variable	VIF	Tolerance
Monthly Rainfall	3.830721	0.261047
Max Daily Rainfall	2.976299	0.335988
Monthly Inflow	1.358036	0.736358
Volume Change	1.292619	0.773623
End of Month Volume	1.042016	0.959678

Prior to training, the predictor variables were evaluated for multicollinearity using Variance Inflation Factor (VIF) and tolerance values. All variables were found to be within acceptable statistical thresholds, with no VIF exceeding the limit of 5. The highest VIF was associated with Monthly Rainfall (3.83), while End of Month Volume had the lowest (1.04), paired with a high tolerance of 0.96. These results confirm the absence of problematic multicollinearity and validate the use of all input variables in the model for flood risk classification.

IV.1.2.2 Model metrics

Accuracy : 0.93
Precision (macro): 0.90
Recall (macro) : 0.91
F1-score (macro) : 0.90
AUC (macro) : 0.98

The XGBoost classifier demonstrated excellent predictive performance. It achieved an overall accuracy of 93%, correctly classifying the majority of inflow levels associated with potential flood scenarios. Macro-averaged values for precision, recall, and F1-score were all close to or above 0.90, indicating a well-balanced model capable of minimizing both false positives and false negatives. The macro-averaged Area Under the Curve (AUC) was 0.98, reinforcing the model's ability to effectively separate inflow levels associated with various degrees of flood risk.

IV.1.2.3 Classification report

Table 12. Classification report of XGboost

	Precision	Recall	F1-score	Support
Low	0.93	0.81	0.87	16
Medium	0.76	0.93	0.84	14
High	1.00	0.98	0,99	45

The model performed strongly across all inflow categories, with particularly high reliability in identifying high inflow events — the most critical category for flood risk management and dam vulnerability assessment. The high inflow class was predicted with perfect precision (1.00) and near-perfect recall (0.98), meaning that the model successfully identified nearly all extreme inflow events with no false alarms. The medium inflow class was also well captured, with a recall of 0.93 and a slightly lower precision of 0.76, indicating occasional overlap with other classes. Low inflow events were classified with a precision of 0.93 and a recall of 0.81. These results highlight the model's strong discriminative ability, particularly its effectiveness in detecting flood-prone conditions at Beni Haroun Dam.

IV.1.2.4 Confusion matrix

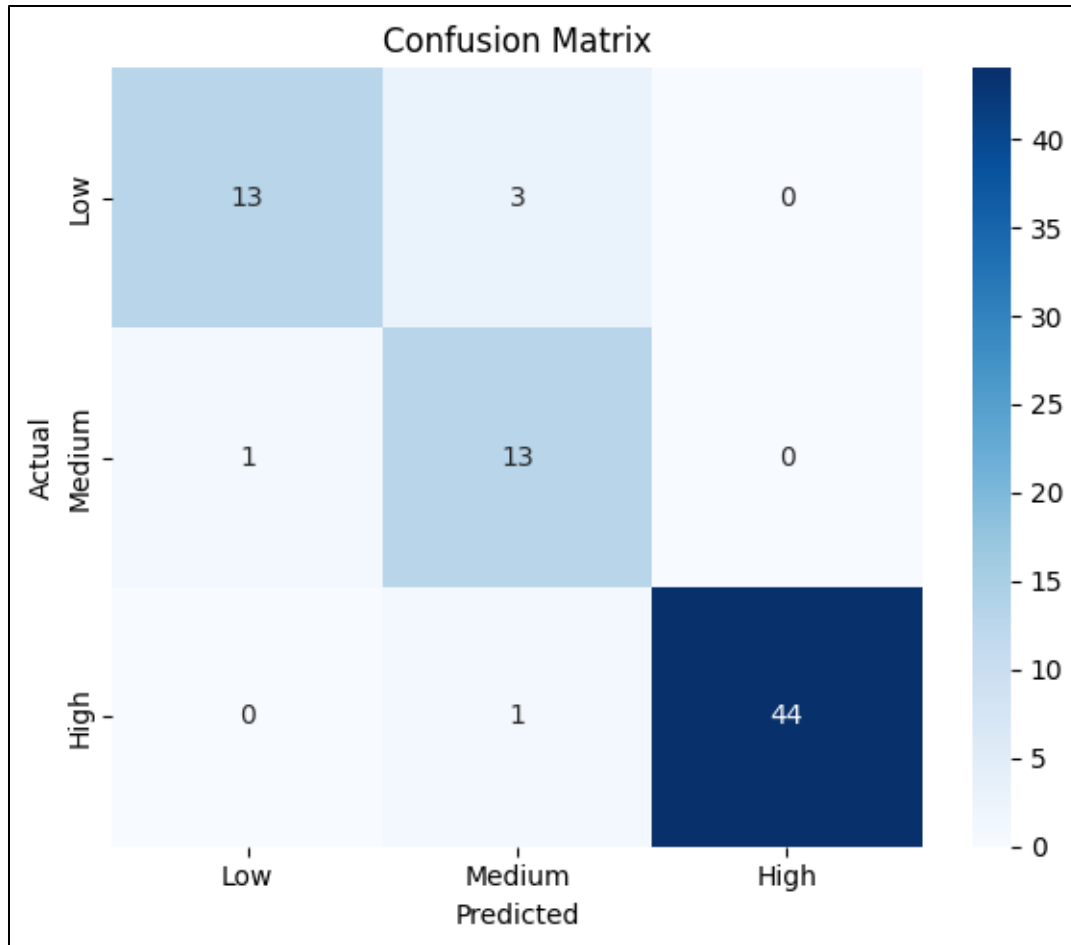


Figure 18. Confusion matrix of XGBoost

The confusion matrix of the XGBoost classifier presents a comprehensive view of the model's ability to distinguish between low, medium, and high inflow levels. It compares the predicted classifications to the actual observed values.

The matrix shows very high accuracy in predicting high inflow events, with minimal misclassification. This confirms that the model reliably identifies flood-prone scenarios, which are most critical for dam safety and flood preparedness. A few misclassifications were observed between medium and low inflow levels, suggesting some overlap in feature patterns for those two classes. However, this does not significantly detract from the overall performance, as the majority of predictions align closely with true inflow categories.

IV.1.2.5 Normalized confusion matrix

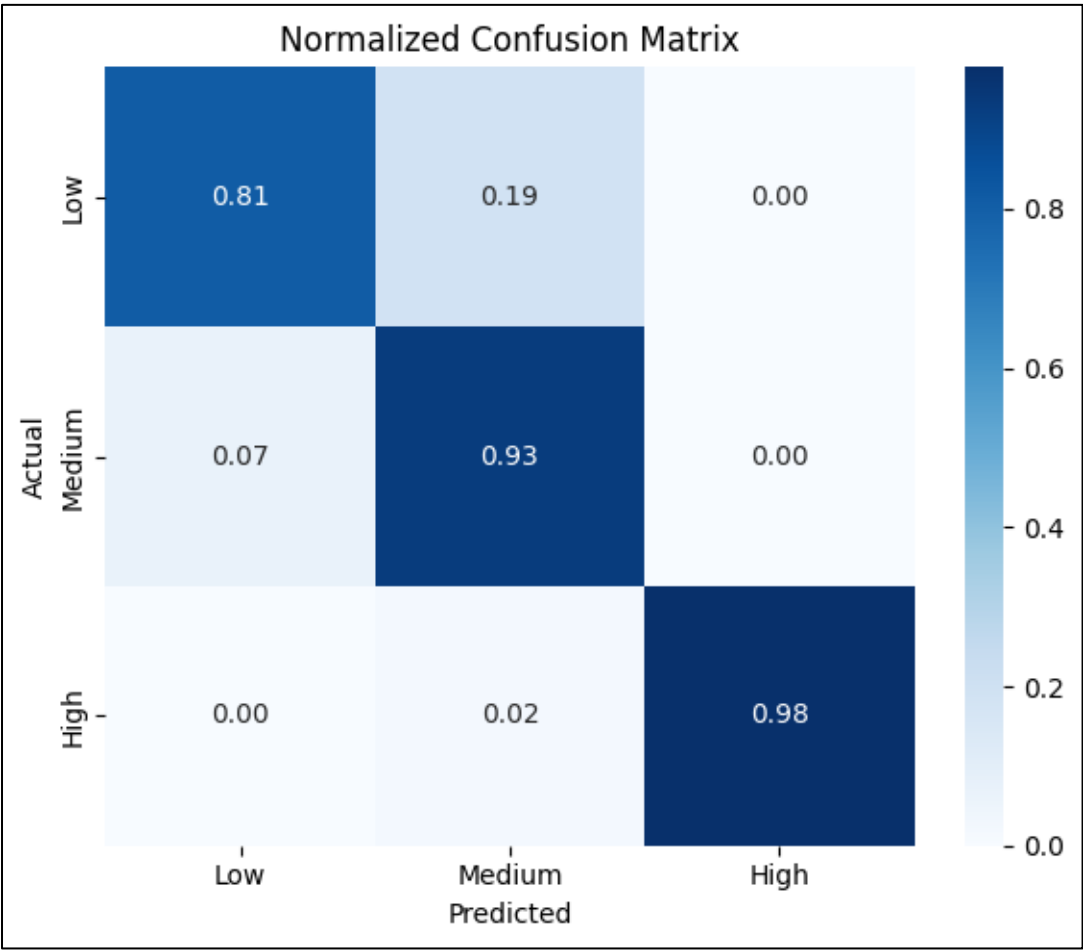


Figure 19. Normalized confusion matrix of XGBoost

The normalized confusion matrix displays the proportion of correct and incorrect predictions across the three inflow categories.

The model achieved a perfect or near-perfect classification rate for high inflow, with approximately 98% of true high inflow events correctly predicted. This high recall rate is essential for flood risk assessment, as it ensures that most potentially hazardous inflow events are successfully identified. The medium inflow category had slightly more variability, with some cases misclassified as low inflow. Nonetheless, the general distribution shows that the XGBoost classifier maintains strong class discrimination, especially where it matters most — in high inflow (flood) conditions.

IV.1.2.6 ROC curve – multi-class

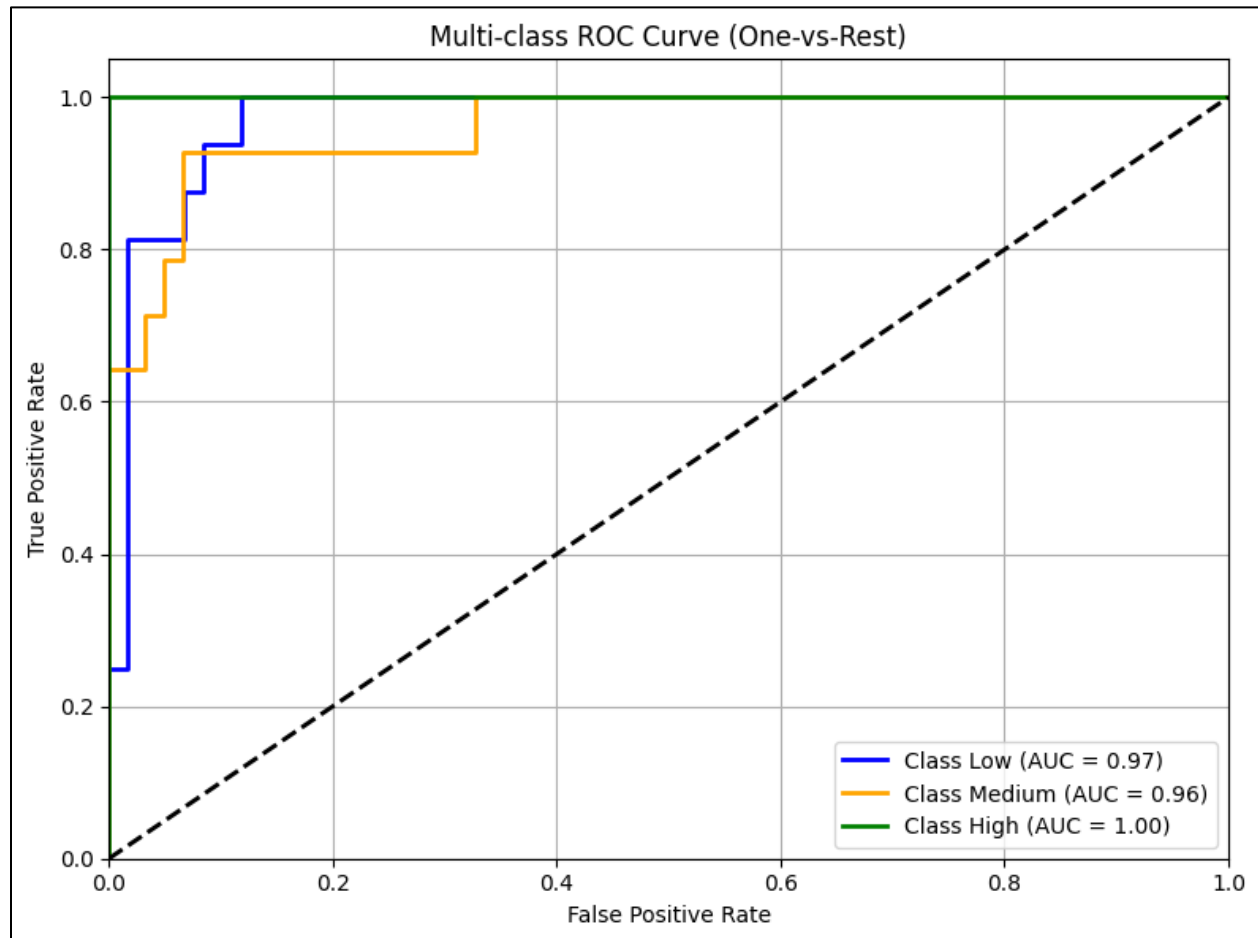


Figure 20. Multi-class ROC curve of XGBoost

The multi-class ROC curve for the XGBoost model provides a comprehensive evaluation of the classifier's ability to distinguish between the three inflow categories: low, medium, and high. This analysis was performed using the One-vs-Rest (OvR) strategy, where the model evaluates each class independently against the other two combined.

The ROC curves produced for each inflow class exhibit excellent shape and spread, with results concentrated near the top-left corner of the plot—indicative of a high true positive rate and a low false positive rate. This reflects a strong ability to differentiate between classes, which is further supported by high AUC (Area Under the Curve) values:

- The high inflow class—which is the most critical for flood risk detection and dam vulnerability—achieved an AUC of 1, confirming XGBoost's exceptional sensitivity and

precision in predicting flood-prone conditions. This means the model is highly effective at identifying high inflow events with minimal misclassification, making it a reliable component for early warning and emergency preparedness systems.

- The medium inflow class also demonstrated a high AUC, though slightly lower than the high inflow class. This suggests the model is competent at detecting transitional inflow events, although there may be minor overlap with low inflow periods—an expected challenge given the subtle boundary between the two.
- The low inflow class exhibited strong classification performance, with an AUC that also reflects high confidence in distinguishing dry or low-risk inflow periods from more hazardous conditions.

Overall, the ROC curve analysis confirms that XGBoost is not only accurate in its classifications but also highly discriminative, which is critical for operational decision-making.

IV.1.2.7 Feature importance

Table 13. Variable and importance of XGBoost

Variable	Importance
Monthly Inflow	0.800303
Volume Change	0.100573
Max Daily Rainfall	0.041564
End of Month Volume	0.032578
Monthly Rainfall	0.024981

XGBoost's internal feature importance analysis revealed that Monthly Inflow was by far the most influential predictor, contributing around 80% to the model's decisions. This was followed by Volume Change with a moderate contribution, while Max Daily Rainfall, End of Month Volume, and Monthly Rainfall had comparatively smaller influence. This distribution of importance is consistent with the physical dynamics of dam operations, where sustained inflows and storage changes are more indicative of impending flood conditions than short-term rainfall variations. The dominance of inflow and volume-based variables reflects the model's alignment with hydrological processes that drive flood risk at Beni Haroun Dam.

IV.1.2.8 Temporal flood risk map

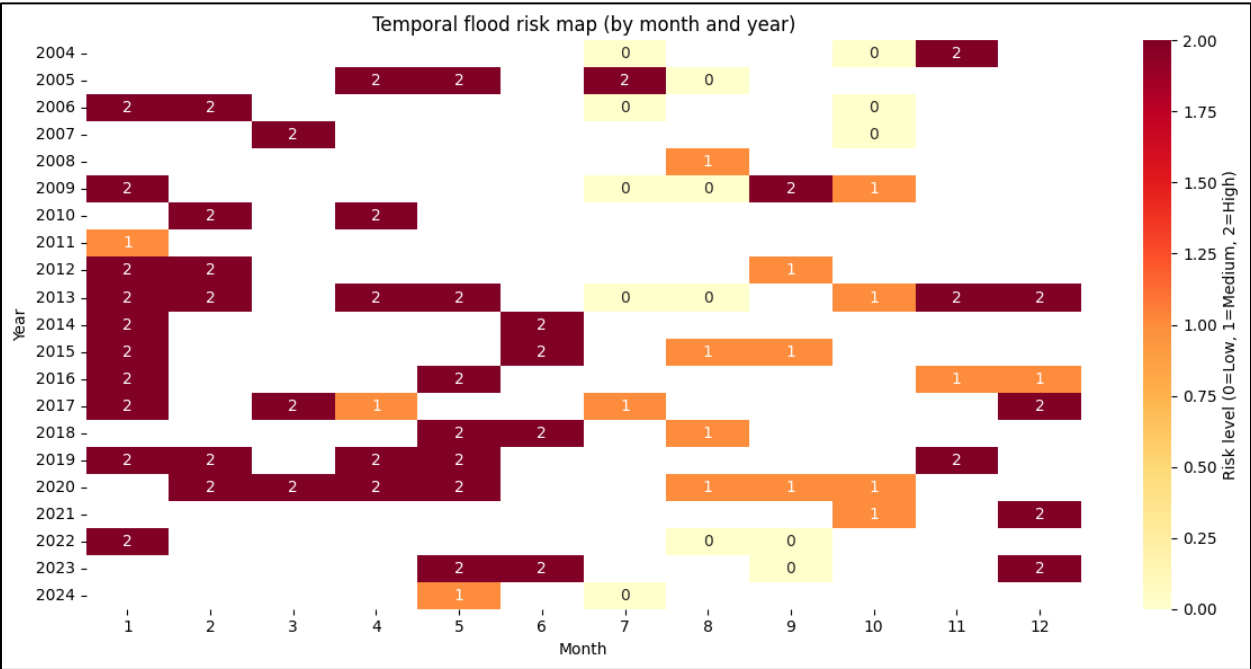


Figure 21. Temporal flood risk map of XGBoost

The temporal risk map produced by the XGBoost model illustrates the distribution of predicted inflow levels over time, allowing for the identification of flood-prone periods and seasonal risk patterns.

The map reveals that high inflow classifications are concentrated during specific months of the year, particularly those corresponding to seasonal rainfall peaks. This consistency demonstrates the model’s ability to capture the temporal behavior of inflow variability, making it a valuable tool for forecasting flood risk windows. The clear identification of high-risk periods supports dam operators in adjusting reservoir operations and preparing mitigation strategies in anticipation of peak inflows.

IV.1.2.9 Annual trend of monthly inflow

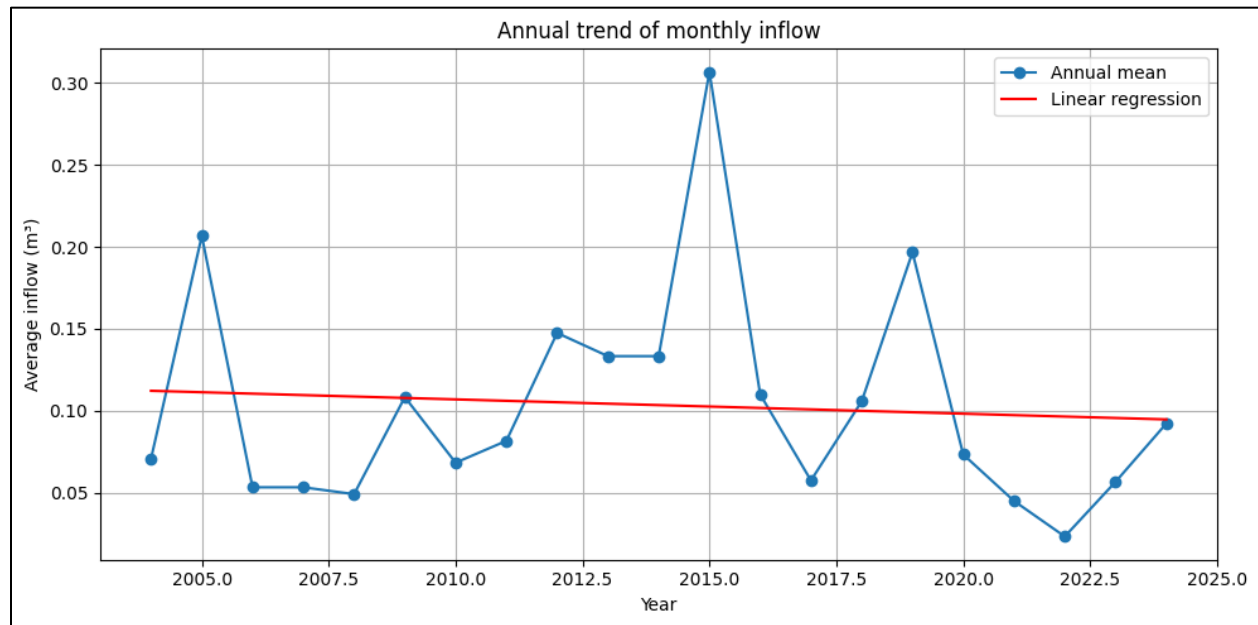


Figure 22. Annual trend of monthly inflow of XGBoost

The annual trend analysis tracks the frequency and distribution of predicted inflow categories across multiple years. It offers insight into how inflow behavior evolves over time, which is crucial for identifying long-term changes in flood risk.

The trend line shows that certain years experienced a greater number of high inflow classifications, reflecting either exceptionally wet years or changes in upstream catchment behavior. Such annual variations are important for strategic reservoir management and may also indicate emerging patterns related to climate variability. For the Beni Haroun Dam, these findings provide a solid foundation for long-term planning and adaptive policy development in response to changing hydrological regimes.

IV.1.4 Gradient Boosting

The Gradient Boosting model was also employed to assess its ability to classify flood risk based on the same set of hydrological indicators. This technique builds models in a sequential manner, where each new tree attempts to correct the prediction errors made by the previous ones. Its strength lies in capturing complex, non-linear relationships and improving performance through a gradual refinement process. Before training the model, it was important to verify that the selected input variables did not suffer from multicollinearity. As with the other models, this was done by calculating Variance Inflation Factor (VIF) and tolerance values to ensure that each variable contributed meaningful and independent information to the learning process.

IV.1.4.1 VIF and tolerance statistics

Table 14. VIF and tolerance statistics of Gradient Boosting

Variable	VIF	Tolerance
Monthly Rainfall	3.830721	0.261047
Max Daily Rainfall	2.976299	0.335988
Monthly Inflow	1.358036	0.736358
Volume Change	1.292619	0.773623
End of Month Volume	1.042016	0.959678

As with the other models, input variables were assessed for multicollinearity using Variance Inflation Factor (VIF) and tolerance values. All variables met acceptable thresholds. Monthly Rainfall exhibited the highest VIF at 3.83, while End of Month Volume had the lowest VIF (1.04) and the highest tolerance value, indicating minimal redundancy and high independence among predictors. This confirms the statistical soundness of the input dataset and its suitability for modeling flood-related inflow behavior.

IV.1.4.2 Model metrics

Accuracy : 0.93
F1-score (macro) : 0.90
Precision (macro): 0.90
Recall (macro) : 0.91
AUC (macro) : 0.99

The Gradient Boosting classifier demonstrated excellent predictive performance, with an overall accuracy of 93%. This high level of precision in classifying inflow levels indicates the model's reliability in operational forecasting scenarios related to flood risk. The macro-averaged F1-score was 0.90, with macro precision and recall values of 0.90 and 0.91, respectively. These metrics show a strong balance between sensitivity and specificity, critical for distinguishing inflow patterns that may indicate flood threats to Beni Haroun Dam. The model also achieved an exceptionally high macro-average Area Under the Curve (AUC) of 0.99, underscoring its superior ability to separate inflow levels across multiple classes.

IV.1.4.3 Classification report

Table 15. Classification report of Gradient Boosting

	Precision	Recall	F1-score	Support
Low	0.93	0.81	0.87	16
Medium	0.76	0.93	0.84	14
High	1.00	0,98	0.99	45

The model exhibited strong classification capability across all inflow categories, with a particular strength in identifying high inflow events—those most associated with flood risk. The high inflow class was predicted with perfect precision and a recall of 0.98, resulting in a near-perfect F1-score of 0.99. For the medium inflow class, the model reached a recall of 0.93 and a precision of 0.76, indicating accurate identification with some degree of misclassification from other categories. The low inflow class achieved a precision of 0.93 and a recall of 0.81, yielding a strong F1-score of 0.87. These results demonstrate the Gradient Boosting model's well-rounded classification strength and its reliability in identifying potentially hazardous flood scenarios.

IV.1.4.4 Confusion matrix

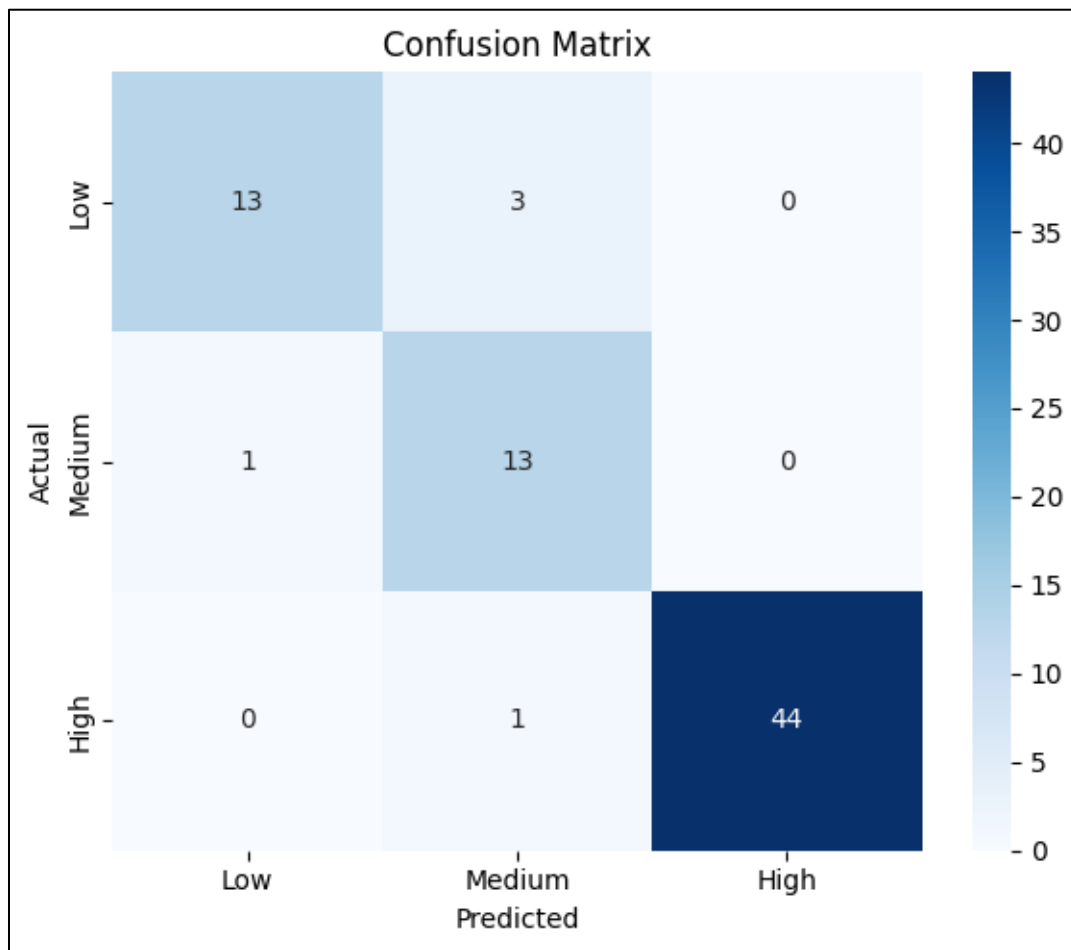


Figure 23. Confusion matrix of Gradient Boosting

The confusion matrix generated by the Gradient Boosting model presents the actual versus predicted inflow classifications. It allows a direct assessment of the model's prediction reliability for each inflow category.

The matrix reveals that the high inflow class was predicted with near-perfect accuracy, reflecting the model's strength in identifying potentially hazardous flood conditions. Some minor misclassifications occurred between medium and low inflow categories, which is expected due to their intermediate and often overlapping characteristics. Nevertheless, the matrix structure confirms that the Gradient Boosting classifier performs exceptionally well in distinguishing the most critical inflow cases related to flood risk and dam vulnerability.

IV.1.4.5 Normalized confusion matrix

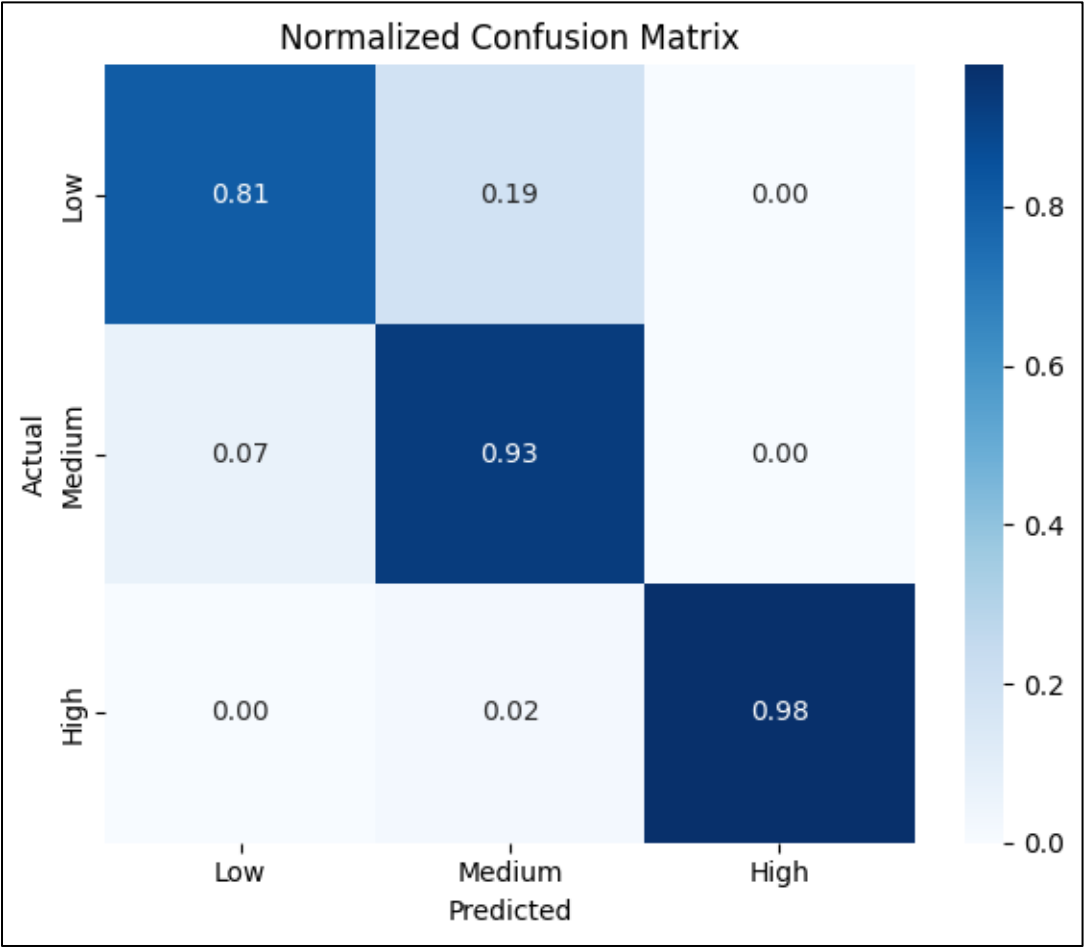


Figure 24. Normalized confusion matrix of Gradient boosting

The normalized confusion matrix highlights the percentage of correct and incorrect predictions for each inflow class, allowing for performance comparison on a proportional basis.

The model demonstrated a very high true positive rate for the high inflow class, with over 98% of high inflow events correctly classified. Medium inflow cases showed a strong recall rate but slightly lower precision due to occasional misclassification from neighboring classes. Low inflow cases were also well detected, with a true positive rate above 80%. These results demonstrate that Gradient Boosting not only identifies flood-prone conditions with great accuracy but also performs reliably across all inflow levels, reinforcing its robustness as a predictive tool in dam risk management.

IV.1.4.6 ROC curve – multi-class

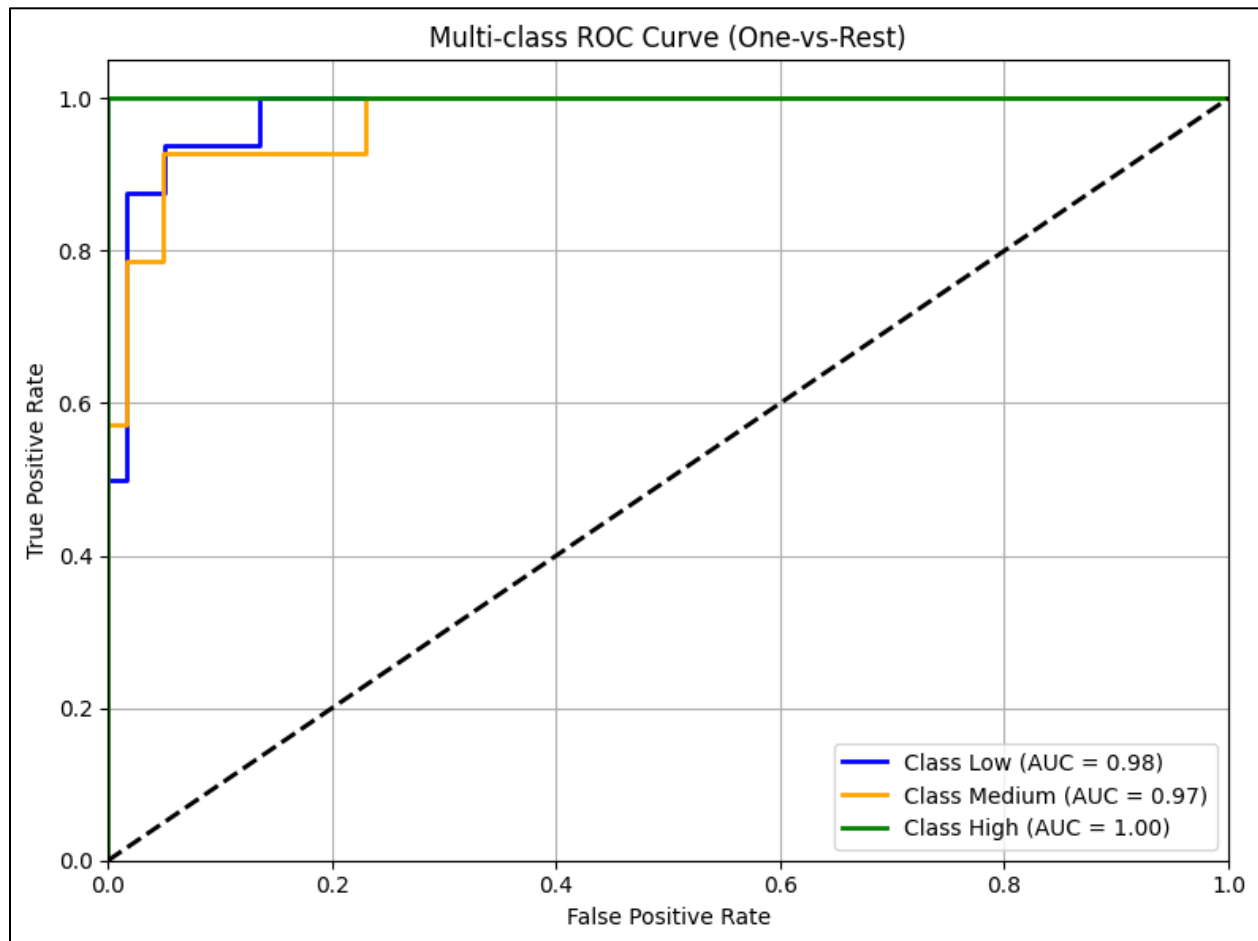


Figure 25. Multi-class ROC curve of Gradient Boosting

The Receiver Operating Characteristic (ROC) curve in a multi-class classification context offers a visual and quantitative evaluation of how well the model distinguishes between different inflow levels: low, medium, and high. For the Gradient Boosting model, the ROC curve was generated using the One-vs-Rest (OvR) approach, where each class is compared against the combination of the other two.

The results show that the Gradient Boosting model exhibits excellent class separability, particularly for the most critical category—high inflow—which is directly associated with flood risk and dam vulnerability:

- The AUC (Area Under the Curve) for the high inflow class reached an exceptional value of 1, confirming the model's perfect ability to detect flood-prone events with minimal

misclassification. This performance aligns with the model's high recall and precision values reported in the classification report and confusion matrices.

- The medium inflow class also achieved a high AUC, although slightly lower than that of the high inflow class. This indicates strong, though slightly less consistent, discriminative power, which is understandable given the transitional nature of medium inflow conditions.
- The low inflow class also demonstrated high AUC performance, confirming that the model can reliably distinguish low inflow periods from riskier conditions.

The ROC curves themselves are skewed toward the top-left corner of the plot for all three classes, signifying high sensitivity (true positive rate) and low false positive rates. This visual behavior supports the conclusion that Gradient Boosting is highly capable of correctly classifying inflow levels across all risk categories.

IV.1.4.7 Feature importance

Table 16. Variable and importance of Gradient Boosting

Variable	Importance
Monthly Inflow	0.160000
Volume Change	0.109333
Max Daily Rainfall	0.037333
End of Month Volume	0.026667
Monthly Rainfall	0.013333

An evaluation of feature importance revealed that Monthly Inflow had the strongest influence on the model's predictions, followed by Volume Change. The remaining features—Max Daily Rainfall, End of Month Volume, and Monthly Rainfall—played smaller but still relevant roles. The importance pattern is consistent with hydrological understanding: sustained inflows and volume variations serve as more reliable indicators of flood potential than isolated rainfall events. This feature hierarchy aligns with dam operations, where cumulative and stored water levels are crucial to risk assessment.

IV.1.4.8 Temporal flood risk map

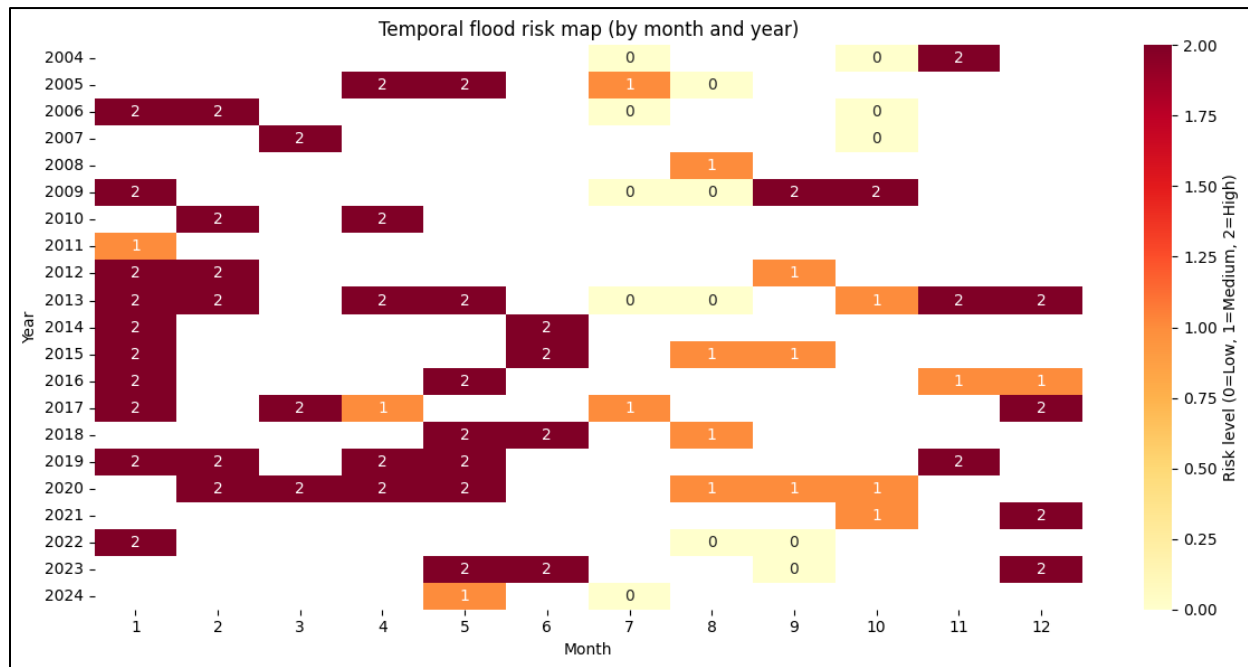


Figure 26. Temporal flood risk map of Gradient Boosting

The temporal risk map created using the Gradient Boosting model shows the distribution of predicted inflow categories over time, providing valuable insight into seasonal and monthly flood risk variations.

The visual pattern confirms that high inflow events are concentrated around the known rainy seasons, reflecting seasonal hydrological behavior. This consistency supports the model's credibility and practical utility in helping dam operators at Beni Haroun anticipate high-risk periods. The risk map can serve as an operational planning tool, enabling early interventions, such as adjusting reservoir levels or preparing emergency protocols during months with high inflow predictions.

IV.1.4.9 Annual trend of monthly inflow

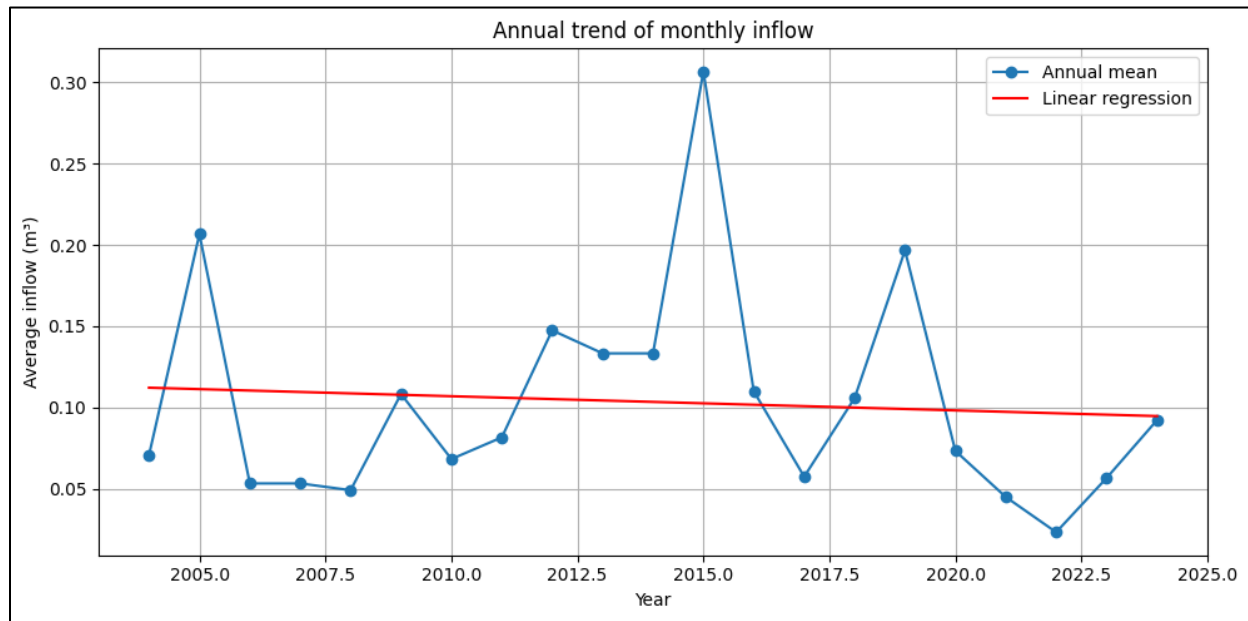


Figure 27. Annual trend of monthly inflow

The annual trend graph offers a broader view of how the frequency of each inflow class changes over multiple years. This long-term perspective is essential for understanding flood vulnerability in the context of evolving climatic and catchment conditions.

The trend line indicates that some years experienced an increased number of high inflow events, suggesting periods of intensified flood risk. These peaks may correspond to extreme weather patterns or above-average seasonal rainfall. For dam safety and water resource management, recognizing such trends is critical to designing adaptive strategies. This trend analysis allows Beni Haroun Dam managers to refine reservoir operations and supports forward-looking flood preparedness policies.

IV.1.5 CatBoost

CatBoost was the final model applied in this study, chosen for its ability to handle categorical features efficiently and its strong performance with relatively small or structured datasets. Although all input variables in this case were numerical, CatBoost still offered advantages in terms of accuracy, speed, and reduced overfitting. Like the other models, CatBoost relies on decision trees and boosting techniques, but with enhanced processing of feature distributions. To ensure a reliable foundation for training, the dataset was first examined for multicollinearity by calculating VIF and tolerance values—confirming that the variables provided distinct and meaningful contributions to the model.

IV.1.5.1 VIF and tolerance statistics

Table 17. VIF and tolerance statistics of CatBoost

Variable	VIF	Tolerance
Monthly Rainfall	3.830721	0.261047
Max Daily Rainfall	2.976299	0.335988
Monthly Inflow	1.358036	0.736358
Volume Change	1.292619	0.773623
End of Month Volume	1.042016	0.959678

To ensure the reliability of the model inputs, multicollinearity among predictors was assessed using Variance Inflation Factor (VIF) and tolerance values. All five variables demonstrated acceptable levels of multicollinearity, with VIF values well below the conventional threshold of 5. Monthly Rainfall had the highest VIF at 3.83, while End of Month Volume recorded the lowest (1.04), with a correspondingly high tolerance. These results confirm that the selected predictors are sufficiently independent and suitable for use in the CatBoost model for flood-related inflow classification.

IV.1.5.2 Model metrics

Accuracy : 0.91
F1-score (macro) : 0.86
Precision (macro): 0.86
Recall (macro) : 0.88
AUC (macro) : 0.97

The CatBoost classifier achieved a high overall accuracy of 91%, indicating that the model correctly predicted inflow levels in the majority of cases. The macro-averaged precision and F1-score were both 0.86, and the macro recall was 0.88, reflecting strong balance in detecting and classifying inflow categories. A macro-average Area Under the Curve (AUC) of 0.97 further confirmed the model's capability to accurately distinguish between inflow conditions that may signal different levels of flood risk at Beni Haroun Dam.

IV.1.5.3 Classification report

Table 18. Classification report of CatBoost

	Precision	Recall	F1-score	Support
Low	0.87	0.81	0.84	16
Medium	0.71	0.86	0.77	14
High	1.00	0,96	0.98	45

Class-specific evaluation revealed that CatBoost performed especially well in identifying high inflow events—those most critical to flood risk monitoring and dam vulnerability assessment. The model achieved perfect precision (1.00) and a recall of 0.96 for the high inflow class, resulting in an F1-score of 0.98. For the medium inflow class, the model reached a recall of 0.86 and a precision of 0.71, suggesting that while most medium inflow cases were correctly identified, some misclassifications occurred from other classes. The low inflow class was predicted with a precision of 0.87 and recall of 0.81. These results demonstrate CatBoost's strong and balanced classification performance, particularly in high inflow conditions that pose the greatest risk to dam safety.

IV.1.5.4 Confusion matrix

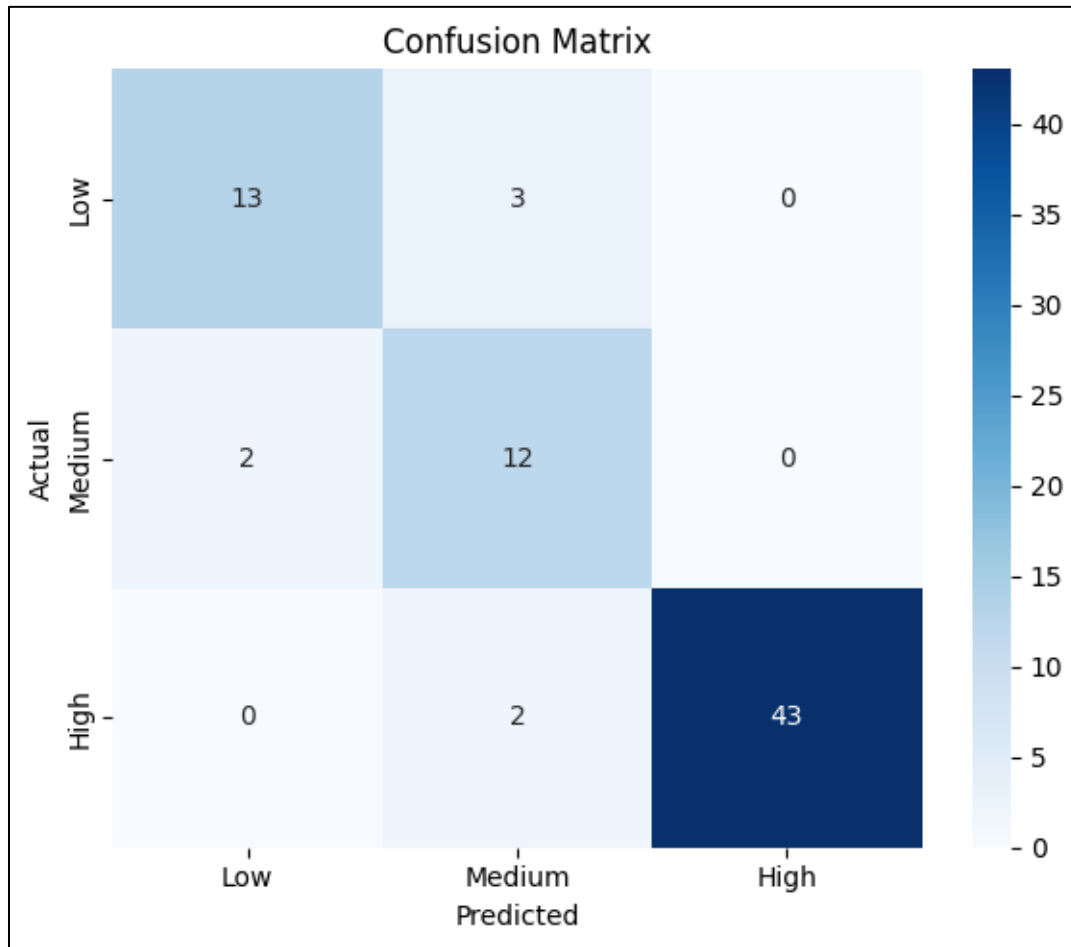


Figure 28. Confusion matrix of CatBoost

The confusion matrix produced by the CatBoost model offers a detailed view of how well the model distinguishes between actual and predicted inflow levels. It highlights the model's classification effectiveness for each of the three inflow categories: low, medium, and high.

The matrix reveals exceptional performance in detecting high inflow events, with nearly all high inflow instances accurately predicted. A small number of misclassifications occurred between the medium and low inflow categories, which is consistent with the transitional nature of those inflow conditions. This strong performance in identifying high inflow cases is especially valuable, as these are directly associated with flood risk and operational vulnerability for the dam.

IV.1.5.5 Normalized confusion matrix

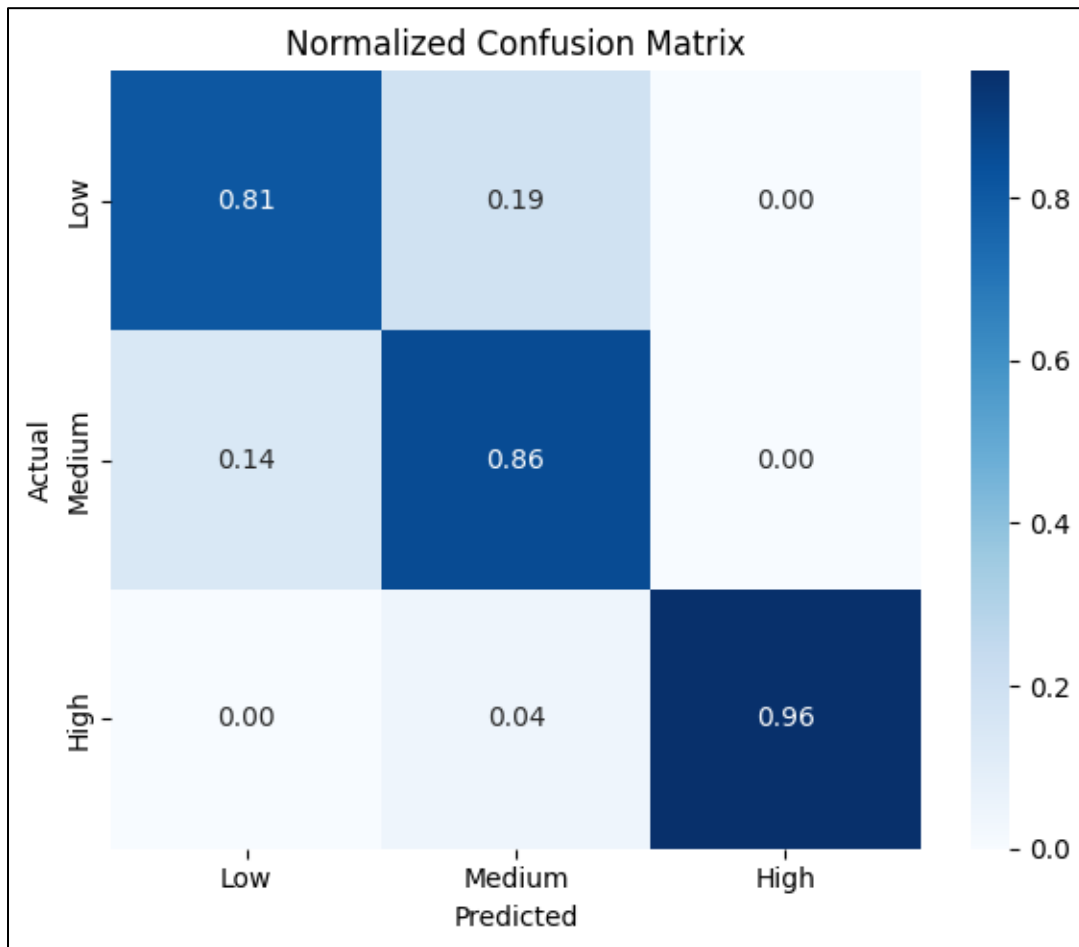


Figure 29. Normalized confusion matrix of CatBoost

The normalized confusion matrix expresses classification results as percentages, providing a clearer picture of how each class was handled by the model.

The high inflow class achieved a precision of 100% and a recall of 96%, demonstrating that the model correctly identified nearly all flood-prone events, with no false positives. The medium inflow class showed a recall of 86% but a lower precision (71%) due to misclassifications from adjacent classes. The low inflow class also showed strong detection, with over 80% correct identification. Overall, the normalized results reinforce CatBoost's reliability in high-stakes inflow prediction, making it a solid tool for flood warning applications.

IV.1.5.6 ROC curve – multi-class

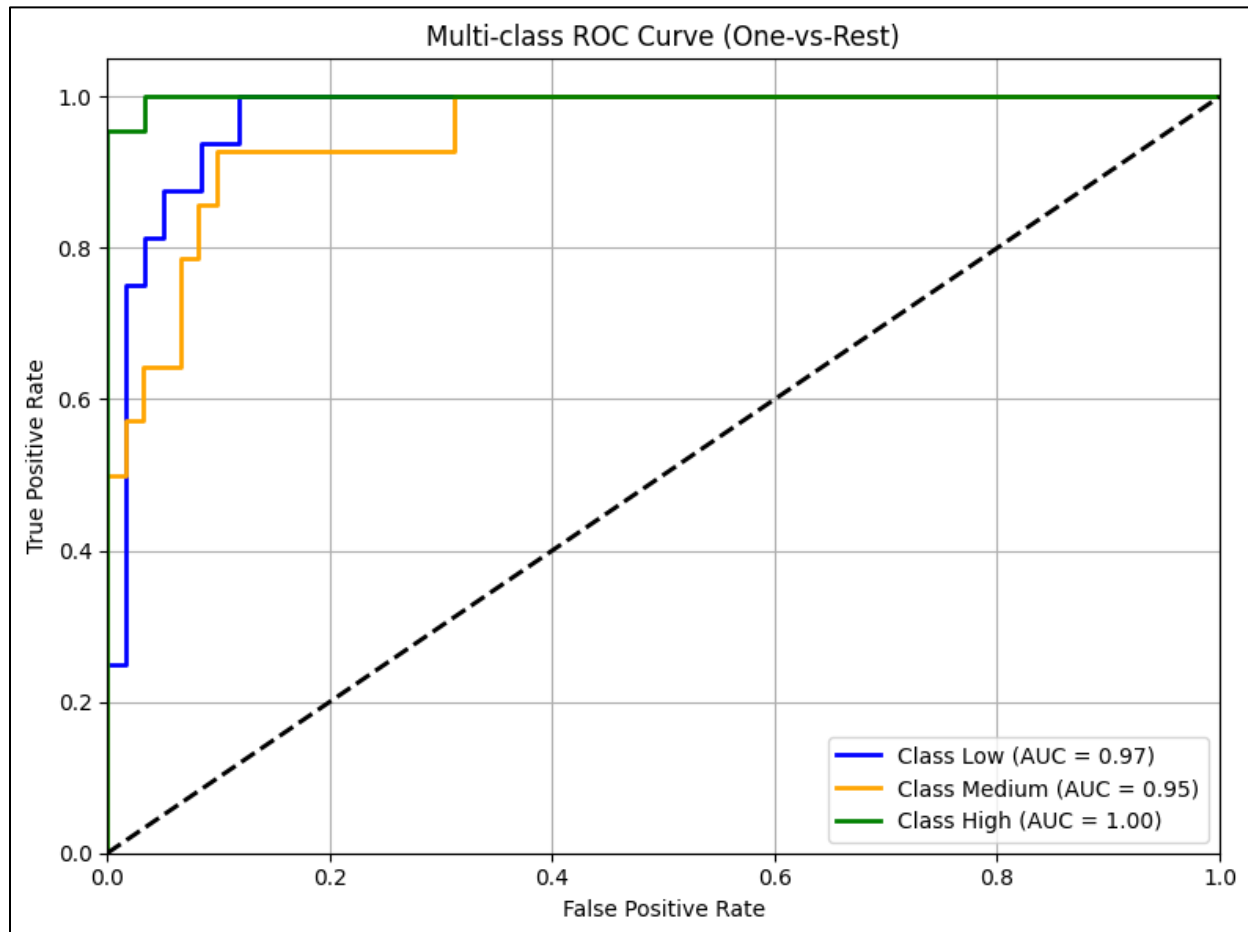


Figure 30. Multi-class ROC curve of CatBoost

The Receiver Operating Characteristic (ROC) curve for the CatBoost model in a multi-class setting provides an important visual and quantitative assessment of the model's ability to distinguish between the three inflow categories: low, medium, and high. In this case, CatBoost was evaluated using a One-vs-Rest (OvR) approach, where a separate ROC curve is plotted for each inflow class against the combination of the remaining classes.

The ROC curves for all three inflow classes demonstrate excellent separation:

- The high inflow class achieved a near-perfect AUC (Area Under the Curve) value of 1, indicating that the model is perfectly capable of correctly identifying high inflow events—critical for flood risk detection and dam safety planning. This reflects the model's high precision and recall for the flood-prone category.

- The medium inflow class also showed a strong AUC, although slightly lower than the high inflow class. This is consistent with earlier classification results, where some overlap was observed between medium and low inflow levels.
- The low inflow class demonstrated a similarly high AUC, confirming the model's reliability in distinguishing between dry, moderate, and wet hydrological conditions.

The shape of the ROC curves—concentrated toward the top-left corner—indicates a high true positive rate and a low false positive rate across all classes. This behavior confirms the robustness of the CatBoost classifier in a multi-class environment, where accurate classification across all inflow levels is essential.

IV.1.5.7 Feature importance

Table 19. Variable and importance of CatBoost

Variable	Importance
Monthly Inflow	0.59424227
Volume Change	0.17111426
Max Daily Rainfall	0.10869699
End of Month Volume	0.06503956
Monthly Rainfall	0.06090693

The CatBoost model's feature importance ranking showed that Monthly Inflow had the greatest impact on classification outcomes, contributing approximately 60% of the total decision-making weight. Volume Change followed with about 17%, while Max Daily Rainfall, End of Month Volume, and Monthly Rainfall had smaller but still relevant contributions. This distribution aligns with hydrological reasoning, where sustained inflows and changes in reservoir volume are primary indicators of flood potential. CatBoost's ability to extract meaningful patterns from these core variables reinforces its suitability for dam safety assessments under flood risk conditions.

IV.1.5.8 Temporal flood risk map

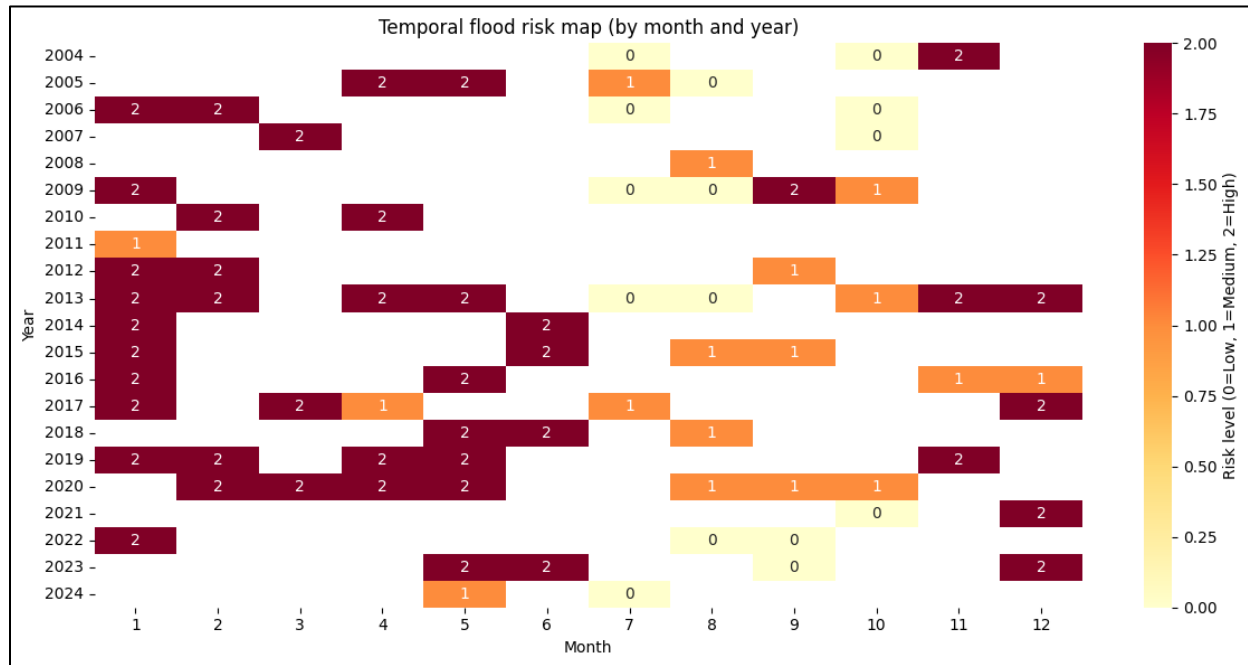


Figure 31. Temporal flood risk map of CatBoost

The temporal risk map created using the CatBoost model visualizes the monthly classification of inflow levels over multiple years. This allows for the identification of recurring flood risk periods and helps reveal the seasonal nature of high inflow conditions.

The map clearly shows clusters of high inflow predictions during the wettest months, particularly in the late winter and spring seasons. These patterns confirm the model's sensitivity to seasonal flood dynamics and offer an evidence-based guide for timing operational adjustments at Beni Haroun Dam. With this tool, dam managers can plan reservoir levels and water releases in advance of expected flood events.

IV.1.5.9 Annual trend of monthly inflow

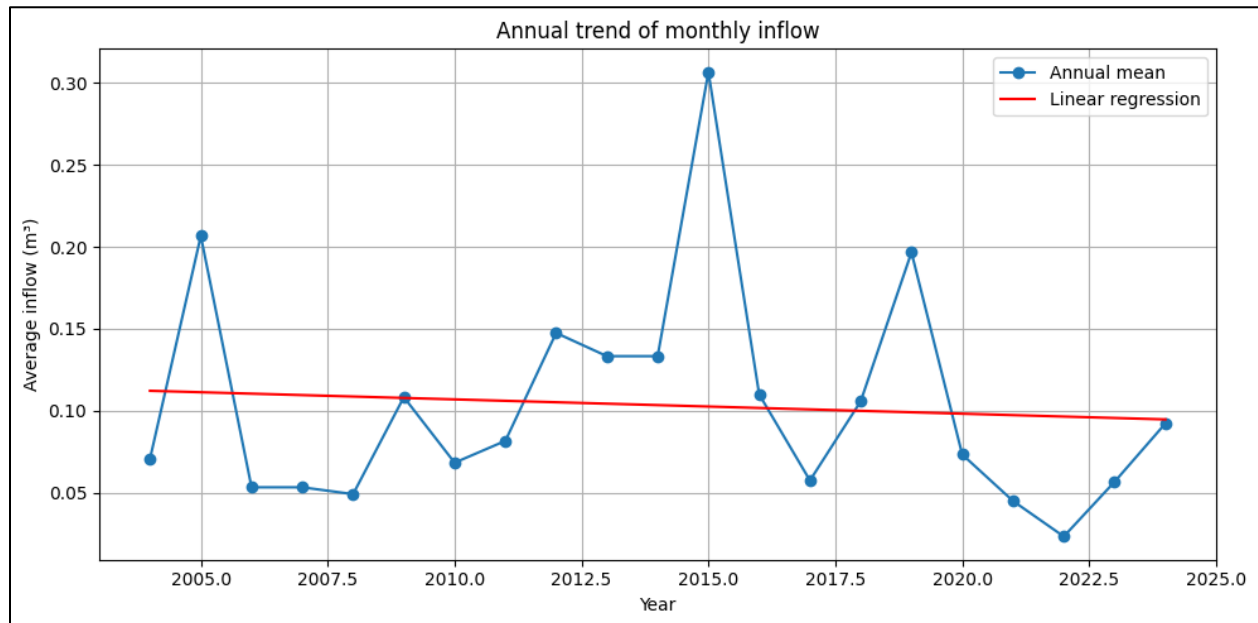


Figure 32. Annual trend of monthly inflow of CatBoost

The annual trend analysis illustrates how the frequency of low, medium, and high inflow classifications varies from year to year. This long-term perspective supports strategic planning and the assessment of whether flood risk is increasing, decreasing, or remaining stable over time.

The trend line indicates interannual variability, with some years showing a higher number of high inflow predictions, possibly linked to exceptional rainfall or hydrological anomalies. This insight is important for dam operators and policymakers, as it supports the development of climate-resilient reservoir management strategies. Understanding how inflow behavior evolves across years enables more proactive and adaptive dam safety planning.

IV.2 Comparative analysis of Machine Learning models for flood risk classification

To support the analysis of the vulnerability of Beni Haroun Dam to flood risks, a comparative analysis was conducted between four top-performing machine learning models: XGBoost, Gradient Boosting, Random Forest, and CatBoost. These models were chosen due to their outstanding classification accuracy, their consistent performance in detecting high inflow events, their interpretability through feature importance, and their practical potential for informing flood risk prediction and dam safety strategies.

IV.2.1 Overall model performance

A review of classification metrics across the four models shows strong and consistent performance. XGBoost, Gradient Boosting, and Random Forest each achieved an accuracy of 93%, while CatBoost followed closely with 91%. These figures highlight the models' ability to correctly classify inflow levels that contribute to different levels of flood risk.

In addition to accuracy, macro-averaged evaluation metrics confirmed the robustness of these models. XGBoost and Gradient Boosting both reached macro F1-scores of 0.90, indicating a high level of performance across all classes. Random Forest achieved a macro F1-score of 0.88, and CatBoost followed with 0.86. The macro-averaged Area Under the Curve (AUC) was above 0.97 for all models, underlining their strong capability to distinguish between flood-related inflow classes.

IV.2.2 Detection of high inflow events

Given the objective of assessing dam vulnerability to flooding, the detection of high inflow events is of critical importance. These events represent potential flood scenarios that demand early warning and timely intervention.

All four models demonstrated excellent performance in identifying high inflow conditions. XGBoost, Gradient Boosting, and CatBoost achieved perfect or near-perfect precision (1.00) and recall values above 0.96, while Random Forest achieved a precision of 1.00 and recall of 0.98. These metrics confirm the effectiveness of the selected models in recognizing periods of elevated hydrological risk, supporting their use in flood preparedness and dam management operations at Beni Haroun.

While some challenges remained in classifying medium and low inflow levels, all models maintained their strength in detecting the most critical class — high inflow — which directly aligns with flood risk scenarios.

IV.2.3 Feature importance and model transparency

Another key advantage of these models is their ability to provide clear insights into the importance of predictor variables, which enhances their interpretability and practical usefulness.

In all cases, Monthly Inflow was the dominant feature, contributing up to 80% in XGBoost and approximately 60% in CatBoost. This was followed by Volume Change and End of Month Volume, both of which reflect reservoir storage dynamics—critical indicators of dam vulnerability. Rainfall variables, though less dominant, added value in identifying variations linked to storm-driven inflow.

The ability to interpret these relationships ensures that the models are not just accurate, but also actionable. Decision-makers can understand why a certain classification is made, allowing greater trust in model outputs and more confident flood risk planning.

IV.2.4 Temporal risk mapping and operational utility

Each model was also used to generate temporal risk maps and examine the annual trend of inflow levels. These outputs are essential for visualizing seasonal flood patterns, assessing long-term inflow behavior, and identifying high-risk periods.

Among the four, Gradient Boosting and CatBoost produced particularly consistent and high-resolution trend outputs. These visualizations help dam managers anticipate inflow surges, optimize reservoir operations, and plan emergency actions more effectively. By integrating these models into real-time monitoring systems, Beni Haroun Dam's flood risk management framework can be strengthened significantly.

IV.3 Conclusion

The comparison between XGBoost, Gradient Boosting, Random Forest, and CatBoost demonstrates that these four models each offer strong potential for classifying flood risk in the context of dam vulnerability analysis. Their high accuracy, ability to detect high inflow events. Additionally, the interpretability of feature importance across models provided valuable insights into the most influential hydrological factors driving flood risk. While each model has its own strengths, the results suggest that tree-based ensemble methods can serve as reliable tools for supporting early warning systems and improving dam safety planning.

Of these, XGBoost and Gradient Boosting stand out as the most balanced in terms of performance and interpretability. Their application can significantly improve early warning systems, support flood response planning, and enhance the overall safety and resilience of Beni Haroun Dam.

IV.4 Recommendations

Based on the outcomes of the machine learning analysis for flood risk classification and dam vulnerability assessment, a set of recommendations is proposed to strengthen the flood preparedness and operational safety of Beni Haroun Dam. These recommendations address both short-term risk mitigation measures and long-term improvements to the dam's monitoring, forecasting, and emergency response systems. Emphasis is placed on leveraging the insights generated by the machine learning models, particularly the high-performing ones such as XGBoost, Gradient Boosting, Random Forest, and CatBoost, to enhance data-driven decision-making.

IV.4.1 Flood risk mitigation strategies

Effective flood risk mitigation begins with the accurate identification of potential flood events and the timely implementation of response measures. Based on the classification results of the selected models, the following strategies are recommended:

- **Implementation of a flood early warning system (FEWS):** The high classification accuracy of the models—especially in detecting high inflow events—can be used to trigger automated warnings when inflow thresholds are exceeded. Integrating these predictions with real-time monitoring of rainfall, reservoir level, and river inflow will improve response times.
- **Seasonal flood forecasting and preparedness planning:** The temporal trend analyses provided by the models offer a basis for forecasting flood-prone periods on a seasonal or monthly basis. These forecasts should be incorporated into reservoir operation plans, guiding pre-emptive storage adjustments during the rainy season to create buffer capacity for anticipated inflows.
- **Strengthening communication protocols:** Alerts generated by predictive models should be linked to local authorities, dam management teams, and downstream communities. Communication protocols should be standardized to ensure clear and timely dissemination of flood warnings.
- **Monitoring of contributing catchment areas:** Given the role of upstream rainfall and runoff in generating high inflow events, it is important to enhance monitoring of contributing sub-basins. This includes installing additional rain gauges and integrating remote sensing data to track changes in catchment hydrology that influence inflow to Beni Haroun Dam.

IV.4.2 Proposed improvements for dam flood management

In addition to risk mitigation, structural and operational improvements can be made to support long-term dam resilience in the face of future flood events. Based on the findings of this study, the following enhancements are proposed:

- **Integration of machine learning models into reservoir decision support systems:** The predictive models used in this research have demonstrated strong performance and interpretability. Embedding them into the dam's management software will allow operators to simulate scenarios, optimize release schedules, and prioritize safety-driven decisions based on inflow predictions.
- **Development of a dynamic reservoir operation policy:** Rather than relying on static rules for water release, a dynamic policy informed by real-time inflow classifications and forecasts would allow more flexible and adaptive responses. This could reduce the risk of overtopping while maintaining water supply and power generation objectives.
- **Regular updating and retraining of models with new data:** As hydrological conditions evolve due to climate variability and land use changes, it is important to update the machine learning models with the latest data. Continuous learning will help maintain the accuracy and relevance of the predictions over time.
- **Investment in dam infrastructure and instrumentation upgrades:** Physical improvements, such as increasing spillway capacity or modernizing gate control systems, can enhance the dam's ability to respond to extreme flood events. In parallel, upgrading instrumentation—such as installing automated sensors for flow, rainfall, and sediment—will provide more accurate inputs to the forecasting models.
- **Capacity-building and training for dam operators:** Effective use of predictive models depends not only on the technology but also on the people using it. Training programs should be conducted to ensure dam personnel understand how to interpret model outputs and integrate them into operational protocols.

General conclusion

The present study examined the application of machine learning techniques to assess the vulnerability of dams to flood risks, using Beni Haroun Dam as a case study. In light of increasing hydrological extremes and the operational challenges faced by large hydraulic infrastructures, the development of reliable, data-driven prediction tools is essential for enhancing dam safety and flood preparedness. This research addressed that need by evaluating the effectiveness of machine learning models in classifying inflow levels that correspond to different degrees of flood risk.

Four high-performing models were selected for detailed analysis: Random Forest, XGBoost, Gradient Boosting, and CatBoost. Among these, ensemble-based methods such as XGBoost and Gradient Boosting demonstrated superior performance, achieving classification accuracies above 90% and displaying high sensitivity in detecting high inflow events — the most critical category for flood risk. These models also offered strong interpretability through feature importance rankings, with variables such as Monthly Inflow and Volume Change consistently emerging as the most influential in predicting flood-prone conditions.

In addition to static classification, the models enabled the generation of temporal risk maps and trend analyses, which added significant operational value. These outputs support dam operators in identifying seasonal patterns and anticipating critical periods, contributing to more proactive and informed reservoir management.

Building on these findings, a series of recommendations were proposed, including the integration of machine learning models into real-time decision support systems, the development of dynamic reservoir operation policies, and the implementation of early warning protocols based on inflow classification results. These strategies aim to improve flood preparedness and reduce the vulnerability of Beni Haroun Dam under both current and future hydrological conditions.

In conclusion, the study demonstrates the potential of advanced machine learning models to support more accurate flood risk assessment and strengthen dam management practices. Continued model refinement, periodic retraining with updated data, and closer integration with on-ground monitoring systems are recommended to sustain and enhance the practical impact of these tools in safeguarding critical water infrastructure.

Perspectives

While this study has demonstrated the effectiveness of machine learning in classifying flood-related inflow levels and assessing dam vulnerability, several promising directions remain open for future exploration and development.

One important perspective is the integration of real-time data streams into the predictive models. Incorporating live measurements from rainfall sensors, river gauges, and satellite data would enable near-instant model updates and improve the responsiveness of early warning systems. Such real-time integration could allow the models not only to predict but also to adapt dynamically to changing hydrological conditions.

Another avenue is the extension of the modeling framework to include hydrodynamic simulations or coupling with physically based models. While machine learning offers high accuracy, it functions as a data-driven black box. Combining it with traditional hydrological models could improve physical interpretability and enhance its credibility among engineers and water resource managers.

Additionally, expanding the spatial scale of analysis beyond a single dam—such as applying the methodology across a network of dams or within an entire river basin—would make the approach more comprehensive. This would also allow decision-makers to prioritize interventions at a regional level and allocate resources more efficiently.

From a technical standpoint, there is room to explore deep learning architectures, such as recurrent neural networks (RNNs) or long short-term memory (LSTM) networks, which are well-suited for time series prediction. These could offer enhanced capabilities in capturing temporal dependencies in inflow data, particularly in multi-step ahead flood forecasting.

Finally, future work should emphasize institutional adoption and operational implementation. Building user-friendly interfaces, integrating the models into dam management platforms, and training local operators in their use will be crucial for transitioning this research from academic insight to practical application.

Overall, this work lays the foundation for more intelligent, adaptive, and data-informed dam risk management systems. Continued collaboration between researchers, engineers, and decision-makers will be key to realizing the full potential of machine learning in protecting critical infrastructure from the growing threat of floods.

References

- [1] UN-Water. (n.d.). Water and climate change. UN-Water. Retrieved June 5, 2025, Retrieved from <https://www.unwater.org/water-facts/water-and-climate-change>
- [2] Fluixá-Sanmartín, J., Altarejos-García, L., Morales-Torres, A., & Escuder-Bueno, I. (2018). Review article: Climate change impacts on dam safety. *Natural Hazards and Earth System Sciences*, 18(9), 2471–2488. Retrieved from <https://doi.org/10.5194/nhess-18-2471-2018>
- [3] Nguyen, D. H., Nguyen, T. H., & Tran, D. H. (2023). Flood risk assessment using machine learning, hydrodynamic modeling, and GIS: A case study in the Vu Gia–Thu Bon basin in Vietnam. *Journal of Hydroinformatics*, 26(8), 1852–1867. Retrieved from <https://doi.org/10.2166/hydro.2023.123>
- [4] Fluixá-Sanmartín, J., Altarejos-García, L., Morales-Torres, A., and Escuder-Bueno, I. Review article: Climate change impacts on dam safety, *Nat. Hazards Earth Syst. Sci.*, 18, 2471–2488,. Retrieved from <https://doi.org/10.5194/nhess-18-2471-2018>, 2018.
- [5] UK Government. (2022, August 25). Flood risk and coastal change. GOV.UK. Retrieved from <https://www.gov.uk/guidance/flood-risk-and-coastal-change>
- [6] Internet Geography. (n.d.). What are the human and physical factors that affect flood risk?. Retrieved from <https://www.internetgeography.net/topics/what-are-the-human-and-physical-factors-that-affect-flood-risk/>
- [7] Walls, M. A., Pesek, S., & Peterson, D. (2023, September 26). Flooding in the United States 101: Causes, Trends, and Impacts. Resources for the Future. Retrieved from <https://www.rff.org/publications/explainers/flooding-in-the-united-states-101-causes-trends-and-impacts/>
- [8] Souza, Leonardo; Sanjay Pandit, Grishma; Prakash Chanekar, Tanvi. "Case Study and Forensic Investigation of Failure of Dam Above Kedarnath" (PDF). International Society for Soil Mechanics and Geotechnical Engineering. Retrieved from <https://www.issmge.org/uploads/publications/46/47/29.pdf>
- [9] Central Water Commission. (n.d.). Guidelines for Mapping Flood Risks Associated with Dams. Retrieved from https://damsafety.cwc.gov.in/ecm-includes/PDFs/Guidelines_for_Mapping_Flood_Risks_Associated_with_Dams.pdf
- [10] Ingram, E. (2024, October 4). Climate change challenges to dams: A wave is building. Renewable Energy World. Retrieved from <https://www.renewableenergyworld.com/news/climate-change-challenges-to-dams-a-wave-is-building/>

- [11] Water Power Magazine. (n.d.). Enhancing flood resilience: Adapting dams and levees to a changing climate. Retrieved from <https://www.waterpowermagazine.com/analysis/enhancing-flood-resilience-adapting-dams-and-levees-to-a-changing-climate/>
- [12] Association of State Dam Safety Officials. (n.d.). Living With Dams: Extreme Rainfall Events. Retrieved from https://damsafety.org/sites/default/files/ASDSO-LivingWithDams-ExtremeRainfallEvents-NO%202-WEB_0.pdf
- [13] ScienceDirect. (n.d.). [Title of the article]. Retrieved from <https://www.sciencedirect.com/science/article/pii/S1364815223001639>
- [14] New York State Department of Environmental Conservation. (n.d.). Dam Safety. Retrieved from <https://dec.ny.gov/environmental-protection/water/dam-safety-coastal-flood-protection/dam-safety>
- [15] Dam Safety Toolbox. (n.d.). Comprehensive Evaluation. Retrieved from https://damtoolbox.org/wiki/Comprehensive_Evaluation
- [16] University of Illinois. (n.d.). i-GUIDE scientists take on dam vulnerability. Retrieved from <https://sustainability.illinois.edu/i-guide-scientists-take-on-dam-vulnerability/>
- [17] Federal Emergency Management Agency. (2012, March). Assessing the Consequences of Dam Failure. Retrieved from <https://damsafety.org/sites/default/files/files/FEMA%20TM%20AssessingtheConsequencesofDamFailure%20March2012.pdf>
- [18] U.S. Army Corps of Engineers. (n.d.). Managing Risk. Retrieved from <https://www.nwp.usace.army.mil/Missions/Flood-Risk-Management/Dam-Safety/Managing-Risk/>
- [19] IGI Global. (n.d.). Introduction. Retrieved from <https://www.igi-global.com/pdf.aspx?tid=341708&ptid=338398&ctid=15&t=Introduction&isxn=9798369342848>
- [20] LEVES BATHYMETRIQUES DE ONZE BARRAGES EN EXPLOITATION - ANBT – PDF
- [21] BARRAGE DE BENI HAROUN SUR L'OUED KEBIR - MONOGRAPHIE - ANBT – PDF
- [22] University of California, Berkeley. (n.d.). What is Machine Learning?. Retrieved from <https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/>
- [23] IBM. (n.d.). Random Forest. Retrieved from <https://www.ibm.com/think/topics/random-forest>
- [24] GeeksforGeeks. (n.d.). XGBoost. Retrieved from <https://www.geeksforgeeks.org/xgboost/>

[25] IBM. (n.d.). Gradient Boosting. Retrieved from <https://www.ibm.com/think/topics/gradient-boosting>

[26] Kolli, A. (n.d.). Understanding CatBoost: The Gradient Boosting Algorithm for Categorical Data. Medium. Retrieved from <https://aravindkolli.medium.com/understanding-catboost-the-gradient-boosting-algorithm-for-categorical-data-73ddb200895d>

[27] Esri. (n.d.). How CatBoost Works. Retrieved from <https://pro.arcgis.com/en/pro-app/latest/tool-reference/geoai/how-catboost-works.htm>

[28] ResearchGate. (n.d.). Structure of CatBoost Algorithm. Retrieved from https://www.researchgate.net/figure/Structure-of-CatBoost-algorithm_fig3_372827215

[29] Données d'exploitation barrage Béni Haroun - ANBT – Excel