

UNIVERSITE JOSEPH KI-ZERBO

ECOLE DOCTORALE INFORMATIQUE ET
CHANGEMENT CLIMATIQUE (EDICC)



BURKINA FASO

La Patrie ou la Mort, nous Vaincrons

Order N° :

MASTER RESEARCH PROGRAM

SPECIALITY: INFORMATICS FOR CLIMATE CHANGE (ICC)

MASTER THESIS

Subject:

**Impact of climate change on electricity production in the
Kaleta dam, Republic of Guinea.**

Presented July 10, 2025, by:

SYLLA Mohamed Aminata

Examination jury

President: Dr Ousmane COULIBALY, Directeur Adjoint à l'École doctorale
Informatique et Changements Climatiques
(Directeur adjoint ED-ICC)

Members :

Docteur OUEDRAOGO Adama Associate Professor, Joseph KI-ZERBO University,
Burkina Faso: **Reviewer;**

Docteur Mahamadi SAVADOGO Associate Professor, Joseph KI-ZERBO University,
Burkina Faso: **Supervisor;**

Docteur Demba Aissata SAMOURA, Research and Innovation Officer at the University of
N'Zerekore, Guinea: **Co-supervisor;**

Docteur Marcel Bawindsom KEBRE, Associate Professor, Joseph KI-ZERBO University,
Burkina Faso: **Co-supervisor.**

Academic year 2024-2025

WASCAL

West African
Science Service Centre on
Climate Change
and Adapted Land Use



Federal Ministry
of Research, Technology
and Space

Dedication

I dedicate this work to my mother, Aminata BANGOURA, for her unflinching support during my studies.

Acknowledgements

I wish to express my heartfelt gratitude to Almighty God for His unfailing love, wisdom, knowledge, and grace throughout my life. I am deeply thankful for the strength and opportunity He has given me to pursue and complete this work. This work was only made possible by the support and guidance of some. I want to express from the bottom of my heart:

To the West African Scientific Service on Climate Change and Adapted Land Use (WASCAL) for providing this wonderful capacity-building framework, and the German Ministry of Education and Research, Technology and Space (BMBF) for funding this program.

My sincere thanks go to the Director of ED-ICC, Prof. **Amande OUEDRAOGO**, and the Deputy Director, Dr. **Ousmane COULIBALY**, for their continuous support and academic guidance during my time at the school.

I would also like to extend my sincere appreciation to my main supervisor **Mahamadi SAVADOGO**, PhD, and co-supervisor, **Demba Aissata SAMOURA**, PhD, for their invaluable guidance, encouragement, and mentorship throughout the execution of this project.

My sincere thanks to **Marcel Bawindsom KEBRE**, PhD, for his considerable scientific contribution to this project. I would also like to thank him for his invaluable guidance, encouragement, and mentorship.

Thanks to Dr. Fode Kenth **CONDE** and Ibrahim **SOW** for their contributions in providing the data.

Furthermore, I am grateful to the entire staff and student community of ED-ICC for their support and contributions in various ways, all of which contributed to the successful completion of this work.

Lastly, I am profoundly thankful to my family, especially my mother, **Aminata BANGOURA**, for their unwavering love, sacrifices, and constant encouragement throughout my academic journey.

Abstract:

In Guinea, electricity remains one of the primary challenges faced by the population. The situation may become more complicated due to the impacts of climate change, which can directly influence electricity generation from hydropower dams. This study assesses the potential impacts of climate change on electricity generation at the Kaléta hydropower dam in Guinea, utilizing advanced machine learning models. Two gradient boosting algorithms, CatBoost and XGBoost, were employed to predict production flow and effective electricity production based on historical hydro-climatic data from 2016 to 2024. The models incorporated climate variables including precipitation, temperature, potential evapotranspiration, wind speed, and downward longwave radiation flux. Future projections were generated for the period 2026-2034 under two Shared Socioeconomic Pathways (SSP245 and SSP585) using downscaled CMIP6 climate model outputs. The analysis of historical trends showed significant increasing patterns in both production flow and effective production, while water levels exhibited a non-significant declining trend. Future projections indicate positive impacts of climate change on hydropower generation capacity. Under SSP245, CatBoost predicts increases of 10.28% in production flow and 13.12% in effective production, while under SSP585, increases of 9.92% and 13.25% are projected, respectively. XGBoost showed more conservative but still positive projections with increases ranging from 5.54% to 8.32% across scenarios. Anomaly analysis demonstrated that future projections generally maintain seasonal patterns while showing enhanced production capacity, particularly during wet seasons. These findings suggest that the Kaléta hydropower plant may benefit from climate change in the short term (2026-2034), with increased water availability potentially enhancing electricity generation.

Key words: Climate Change, Hydropower, Electricity generation, Machine learning, Kaléta Dam.

Résumé:

En Guinée, l'électricité reste l'un des principaux défis auxquels la population est confrontée. La situation pourrait se compliquer en raison des impacts du changement climatique, qui peuvent directement influencer la production d'électricité des barrages hydroélectriques. Cette étude évalue les impacts potentiels du changement climatique sur la production d'électricité au barrage hydroélectrique de Kaléta en Guinée en utilisant des modèles avancés d'apprentissage automatique. Deux algorithmes de gradient boosting, CatBoost et XGBoost, ont été utilisés pour prédire le flux de production et la production effective d'électricité sur la base des données hydro-climatiques historiques de 2016 à 2024. Les modèles ont intégré des variables climatiques telles que les précipitations, la température, l'évapotranspiration potentielle, la vitesse du vent et le flux de rayonnement de grande longueur d'onde vers le bas. Les projections futures ont été générées pour la période 2026-2034 dans le cadre de deux voies socio-économiques partagées (SSP245 et SSP585) à l'aide des sorties du modèle climatique CMIP6 mises à l'échelle. L'analyse des tendances historiques a montré des schémas d'augmentation significatifs à la fois du débit de production et de la production effective, tandis que les niveaux d'eau présentaient une tendance à la baisse non significative. Les projections futures indiquent des impacts positifs du changement climatique sur la capacité de production hydroélectrique. Dans le cadre du scénario SSP245, CatBoost prévoit des augmentations de 10,28 % du débit de production et de 13,12 % de la production effective, tandis que dans le cadre du scénario SSP585, des augmentations de 9,92 % et de 13,25 % sont prévues, respectivement. XGBoost a montré des projections plus conservatrices mais toujours positives avec des augmentations allant de 5,54% à 8,32% à travers les scénarios. L'analyse des anomalies a démontré que les projections futures maintiennent généralement les schémas saisonniers tout en montrant une capacité de production accrue, en particulier pendant les saisons humides. Ces résultats suggèrent que la centrale hydroélectrique de Kaléta pourrait bénéficier du changement climatique à court terme (2026-2034), une plus grande disponibilité de l'eau pouvant améliorer la production d'électricité.

Mots clés : Changement Climatique, hydroélectricité, production d'électricité, apprentissage automatique, Kaléta Barrage.

Contents

Dedication.....	i
Acknowledgements	ii
Abstract:	iii
Résumé:	iv
Contents	v
List of figures and maps	vii
List of tables	ix
List of Abbreviations	x
I -Introduction.....	1
1.1 BACKGROUND.....	1
1.2 Problem statement	3
Main and specific Research questions.....	5
hypothesis of the research	5
main and specific objective	5
Thesis structure.....	6
<i>Chapter 1 Literature Review.....</i>	<i>7</i>
1.1 Introduction.....	7
1.2 Hydropower in West Africa.....	7
1.3 Hydropower in other regions.....	8
1.5 conclusion	11
<i>Chapter 2 Materials and Methods.....</i>	<i>12</i>
2-1 Study area.....	12
2-1.1 Hydrology.....	13
2-1.2 Climatology	14
2-1.3 Forest.....	15
2-1.4 Energy supply.....	15
2-1.5 Konkouré watershed.....	17
2.2 Datasets	19
2-2.1 Climate data.....	19
2-2.1.1 Observation data	19
2-2.1.2 Reanalysis and satellite products data	19
2-2.1.3 CMIP6 projection data	19

2-2.2 Hydrological and hydropower Data.....	21
2-2.3 Summary of the data used in this research	21
2-2.4 Data correlation	22
2-3 Methodology.....	23
2-3.1 Processing Tools.....	25
2-3.2 Data preprocessing	25
2-3.3 Understanding Machine Learning Models	27
2-3.3.1 CastBoost Model	27
2-3.3.2 XGBoost (Extreme Gradient Boosting).....	29
2-3.3.3 Training and validation	31
Chapter 3 Results and Discussion.....	35
3-1 Variability Of Climatic Parameters	35
3.2. Variability of hydroelectric production parameters	38
3-2.1 Monthly change in production flow, effective production, and water level	38
3-2.2 Annual change in production flow, production effective, and water level	40
3-3 Best climate projection for Konkouré watershed	43
3-4 Variability of Climatic Parameters	45
3-5 SHAP explanation of the variables.....	46
3-6 Hydropower prediction under shared socioeconomic pathways	48
3-6.1 Model performance on historical data (2016-2024)	48
3-6.1.1 Production flow	49
3-6.1.2 Production effective (electricity production).....	50
3-7 The relative change in electricity generation prediction under shared socioeconomic pathways..	52
3-8 Maximum monthly changes comparison of the historical period and the prediction	54
3-9 Anomalies in production flow and production effective	58
Conclusion and Perspective.....	61
Conclusion	61
Perspective.....	63
References	64

List of figures and maps

<i>Figure 2.1: Hydrographic map of the Republic of Guinea(Samoura et al.2011)</i>	14
<i>Figure 2.2: Climatology of precipitation and temperature of the Republic of Guinea</i>	15
<i>Figure 2.3: DEM of Konkouré watershep</i>	18
<i>Figure 2.4: Kaleta power plant</i>	18
<i>Figure 2.5: Correlation heatmap of climate and hydropower variables 2016-2024</i>	23
<i>Figure 2.6: Summary of methodology</i>	24
<i>Figure 2.7: CatBoost model</i>	29
<i>Figure 2.8: XGBoost algorithm explanation(https://www.analytixlabs.co.in/blog/gradient-boosting-algorithm/)</i>	31
<i>Figure 3.1: Anomaly of rainfall bar chart in Koukoure watershed</i>	36
<i>Figure 3.1:Annual Anomaly of minimum temperature (Tasmin) in Koukoure watershed</i> ..	36
<i>Figure 3.2: Annual Anomaly of Tmax bar chart in Koukoure watershed</i>	37
<i>Figure 3.3: Annual Anomaly of Potential evapotranspiration in Koukoure watershed</i>	37
<i>Figure 3.4: Monthly distribution of production flow</i>	38
<i>Figure 3.6: Monthly distribution of production effective</i>	39
<i>Figure 3.7: Monthly distribution of water level</i>	40
<i>Figure 3.8: Annual changes in production flow</i>	41
<i>Figure 3.9: Annual change in effective production</i>	42
<i>Figure 3.10: Annual changes in water level</i>	42
<i>Figure 3.11: Tailor diagrams of CMIP6 comparisons with ERA5 Tasmin</i>	43
<i>Figure 3.12: Tailor diagrams of CMIP6 comparisons with ERA5 Tmax</i>	43
<i>Figure 3.13: Tailor diagrams of CMIP6 comparisons with CHIRPS pr</i>	44
<i>Figure 3.14: Annual Anomaly of Potential evapotranspiration in Konkouré watershed</i> ...45	
<i>Figure 3.15: Annual Anomaly of Potential evapotranspiration bar chart in Konkouré watershed</i>	45
<i>Figure 3.16: SHAP value summary of Hydro-Climatological Parameters on production effective</i>	46
<i>Figure 3.17: SHAP value summary of Hydro-Climatological Parameters on production flow</i>	47
<i>Figure 3.18: Production flow prediction</i>	49
<i>Figure 3.19: Production effective (electricity production) prediction</i>	51

<i>Figure 3.20: Monthly changes in the maximum production flow...</i>	55
<i>Figure 3.21: Monthly changes in the maximum effective production...</i>	55
<i>Figure 3.22: Monthly maximum production flow trends changes...</i>	56
<i>Figure 3.23: Monthly maximum production flow trends changes...</i>	56
<i>Figure 3.24: Monthly anomaly production flow...</i>	58
<i>Figure 3.25: Monthly production effective anomaly...</i>	58

List of tables

Table 2.1: Data sources and variables22

Table 3.1: Scores of validations on production flow.....44

Table 3.2: Scores of validations in effective production.....46

List of Abbreviations

AEP	Africa Energy Portal
AFD	Agence Francaise de Developement
ANM	Agence Nationale de la Météorologie
CatBoost	Categorical Boosting
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station data
CDO	Climate Data Operator
CMIP6	Coupled Model Intercomparison Project Phase 6
CWE	China International Water & Electric Corporation
EDF	Électricité de France
EDG	Électricité de Guinée
ERA5	ECMWF Reanalysis v5
ESGF	Earth System Grid Federation
GHG	Greenhouse Gas
GLEAMv4	Global Land Evaporation Amsterdam Model, version 4
GWh	Gigawatt-hour
IAE / IEA	International Energy Agency
IDW	Inverse Distance Weighted
IPCC	Intergovernmental Panel on Climate Change
ISSA	Improved Sparrow Search Algorithm
KWh	Kilowatt-hour
MAE	Mean Absolute Error
MIROC6	Model for Interdisciplinary Research on Climate, version 6
MPI-ESM1-2-HR	Max Planck Institute Earth System Model, High Resolution
MRI-ESM2-0	Meteorological Research Institute Earth System Model, version 2.0
MSE	Mean Squared Error
MW	Megawatt
NASA_FLDAS	NASA Land Data Assimilation System
NetCDF	Network Common Data Form

NESM3	Nanjing University of Information Science and Technology Earth System Model, version 3
NorESM2-MM	Norwegian Earth System Model, Medium Resolution
R²	Coefficient of Determination
RCP 2.6 / 4.5 / 8.5	Representative Concentration Pathways (low, moderate, high emissions)
RMSE	Root Mean Squared Error
SHAP	SHapley Additive exPlanations
SODEKA SAS	Société de la Gestion de Kaléta
SSP2-4.5 / SSP5-8.5	Shared Socioeconomic Pathways (moderate/high scenarios)
SWAT	Soil and Water Assessment Tool
TaiESM1	Taiwan Earth System Model, version 1
Tasman	Temperature maximum
Tasmix	Temperature minimum
TWh	Terawatt-hour
UNESCO	United Nations Educational, Scientific, and Cultural Organization
UNFCCC	United Nations Framework Convention on Climate Change
WACREEE	West African Regional Center for Renewable Energy and Energy Efficiency
WA	West Africa
WASCAL	West African Science Service Centre on Climate Change and Adapted Land Use
WEAP21	Water Evaluation and Planning System (version 21)

I-INTRODUCTION

1.1 BACKGROUND

Climate and water are connected in complex ways through several variables such as precipitation, temperature, humidity, evaporation, evapotranspiration, and solar radiation; however, these conditions differ from region to region. The impacts of a change in any of these variables alter the hydrological cycle by increasing runoff, the intensity of rainfall, and the evaporation and evapotranspiration rate Lee and Huang. (2014). The changing climate will increase strongly from decades with higher temperatures and extreme precipitation compared to preindustrial levels, resulting in decreased water supplies, water quality, and increased water demand IPCC. (2014), with severe impacts on the southern and northern Saharan communities. One of the key components of water resources is precipitation (runoff), which is highly impacted by climate change. Climate Change has a direct impact on hydropower generation. Many studies can be found across the globe assessing the effects of Climate Change on hydropower production. In the coming decades, climate change is expected to intensify extreme weather events, including floods, droughts, and rising temperatures. These climatic shifts, combined with increasing water demands driven by population growth and rising living standards, are likely to place significant stress on water resources. The situation becomes even more critical when compounded by additional pressures, such as large-scale water transfer projects. These combined factors can have a profound impact on the reliability and efficiency of hydroelectric power generation Da Silva et al. (2021). The production of energy from fossil fuels is harmful to the environment in the long term, which is considered the main source of anthropogenic greenhouse gas (GHG) emissions that contribute to climate change Arvizu et al. (2021). According to the International Energy Agency (*IEA*), fossil fuels remain Africa's largest energy source, with an 80% share of total electricity generation. In comparison, 15.52% comes from hydropower, and the remainder from other sources (nuclear, waste, wind, and solar). In West Africa (WA), hydropower is the main source of renewable energy. In 2017, the total electrical energy production was respectively 32 249 GWh in Nigeria (17.14% from hydropower), 14 068 GWh in Ghana (39.92% from hydropower), 8575 GWh in Ivory Coast (17.83% from hydropower), 232 GWh in Togo (88.4%), and 4777 GWh (7.05%) in Senegal. Hydropower dams play a crucial role in the region's socioeconomic development

by supporting energy production, agriculture, drinking water, fisheries, and livestock. However, their effective utilization is challenged by hydrological variations influenced by rainfall patterns, geological conditions, and climate change, underscoring the need for integrated water resource management and a comprehensive understanding of watershed hydroclimatic dynamics according to Kouadio et al. (2020). Hydropower represents a crucial component of renewable energy strategies in West Africa, offering a sustainable solution to meet the growing energy demands across the region. Despite its promising potential, the development of hydropower in West Africa has been hindered in the past by a lack of comprehensive and reliable data, as well as technologies and competences on river systems and energy generation capacity Kling et al. (2017). The most promising potential hydropower-generating countries in the West Africa region are Guinea, Sierra Leone, Liberia, as well as in the southeastern part of Nigeria. In the Makona River basin in Guinea, warming temperature increases the potential evapotranspiration and thus reduce the runoff. The availability and reliability of hydrometeorological data are the most significant challenges for water resources assessment in the West Africa region (Kling et al. (2016). Guinea has abundant energy resources, including hydroelectric, biomass, solar, and wind. The country's dense hydrographic network gives it a hydroelectric potential estimated at more than 6,000 MW, capable of producing 19,300 GWh annually. However, many micro and small hydro sites remain unidentified and unassessed by researchers and the National Energy Directorate. As a result, most of this potential is still untapped. Currently, only the Konkouré watershed has been explored, and only for about three-quarters of its potential Sakouvogui et al. (2023). According to the Africa Energy Portal (*AEP*), the Republic of Guinea faces significant electricity needs due to its low access rate and high demand growth. As of recent data, only about 36% of Guinea's population had access to electricity in 2020, with this figure being much lower in rural areas. The country aims to achieve universal electricity access by 2030, largely through renewable energy projects, including hydroelectric and solar power installations. Hydropower facilities like the Souapiti and Kaléta dams play a crucial role in meeting national demand, especially in urban centers, but ongoing development is needed to support rural electrification through mini-grids and other initiatives. Moreover, it is essential both to conserve water resources and ensure that the proper functioning of water-related activities is so important, in particular, the hydroelectric production of the Faye dam in the medium and long term Kouadio et al. (2020). The inadequate production of hydroelectric power necessitates the need to assess

climate change on electricity production at Kaléta Dam in the Republic of Guinea and provide useful information on hydropower generation projection in the changing climate.

1.2 PROBLEM STATEMENT

Africa's electricity demand and supply were 621 TWh in 2012 and are projected to rise to 1,258 TWh by 2030 and 1,869 TWh by 2040. However, sub-Saharan Africa continues to struggle with a severe energy crisis, characterized by low electricity access, particularly in rural areas, despite the region's vast energy potential. At the same time, rising temperatures due to global warming make sub-Saharan Africa one of the most climate-sensitive regions, affecting hydropower generation and other sectors across the region. Compared to North Africa and other regions, electricity access remains significantly lower in urban and mostly in rural parts of sub-Saharan Africa, leaving millions without reliable energy sources Barasa et al. (2018). Climate change affects the hydrological regimes of rivers by altering the characteristic amount, intensity, form, and timing of precipitation. Also, it can affect the rate of evapotranspiration by affecting the peak rate, volume, and timing Gleick et al. (1999). Hydropower is a clean and renewable source of electricity generation, primarily generated through run-of-river installations that rely on river flow, which is largely influenced by precipitation. However, rising temperatures and global warming are driving an increased demand for energy in all the sectors of development and a better standard of living, while at the same time, potential hydropower generation may decrease due to the temperature rise. This highlights the need for more accurate production forecasts to enhance planning for green energy generation. In developing countries, achieving this without relying on fossil fuels or nuclear power remains a significant challenge in the electricity sector (Maciejewski et al. (2024). Hydropower is a renewable and low-emission source of greenhouse gas, offering a sustainable alternative to fossil fuels. However, it depends on stable surface water flows, making it more vulnerable to global warming and climate change. The impacts of climate change on hydropower vary by location and can be positive, negative, or neutral, depending on factors like reservoir size, energy market dynamics, and water management priorities on one hand and on the other, by climatic zone, extreme rainfall, drought, etc. Additionally, challenges such as floods, landslides, sediment accumulation, and greenhouse gas emissions from reservoirs should be considered in future hydropower planning. Over the past decade, nations and international agencies have

explored more and more hydropower as a cost-effective and clean energy solution, highlighting the urgent need to assess both its potential and risks under changing climate conditions Wasti et al. (2022).

The seasonal change in climate variability and long-term warming of the global climate have profoundly affected the hydrological cycle. Key impacts include increased atmospheric water vapor, unpredictable climate shifts, more intense and extreme rainfall, changes in hydrological patterns, reduced water availability, higher potential evapotranspiration, and altered soil moisture and runoff dynamics. Climate change increases the levels of water vapor in the atmosphere by evaporation and potential evapotranspiration rate, which can make water availability less predictable on a global and regional scale. When those areas receive the most precipitation, it can most of the time result in flooding; on the other hand, droughts, heat waves Borowski, (2020). The acceleration of the water cycle due to climate variability in the local area by those activities which also contribute to degradation of the forest in the watershed, doing pollution of water quality and due to climate change is indeed driven by increased greenhouse gas (GHG) emissions, which intensify the evaporation, evapotranspiration rate and precipitation patterns. This intensification affects water availability, flood risks, and drought frequencies, posing a significant challenge to global water resource management IPCC.(2021). According to the West African Regional Centre on Renewable Energy and Energy Efficiency (WACREEE), the impact of climate change on West African water resources is well known, but the effects on hydropower generation are not well documented in West Africa due to the lack of available data in some regions. This means that there are few and limited details about the impact of climate change on hydropower generation at the country level. Therefore, it raises the demand to conduct studies related to the impact of climate change on hydropower generation across this region. Especially in the Republic of Guinea, it is crucial to understand how hydrological variability influences the production flow in the Konkouré basin and its impacts on hydropower generation in the context of climate change. Hydropower resources are the key expected solution and will remain a crucial means to meet the growing energy needs of the future in the Guinean and savanna zones, as well as a solar power plant for the Sahelian zones. In the Republic of Guinea, like many other countries in the region, socio-political and economic factors, along with the challenges posed by climate change and the urgent need to reduce greenhouse gas emissions, are driving governments to prioritize investments in renewable energy sources.

These investments are crucial in mitigating human-induced climate change, as renewable energy sources produce minimal greenhouse gas emissions. However, in Guinea, the country shouldn't invest in its hydropower plant due to its ranking of poverty and poor governance. However, the effects of anthropogenic climate change observed in recent years are expected to persist, even if greenhouse gas concentrations had remained stable at the levels outlined in national climate commitments, which, in reality, have not been the case. Additionally, the widespread poverty and lack of electricity in many urban areas further underscore the need for sustainable energy solutions, making the development of renewable energy sources such as hydropower even more urgent Sakouvogui et al. (2023).

MAIN AND SPECIFIC RESEARCH QUESTIONS

The main research question underlying this study is: *How does climate change impact electricity production at Kaléta Dam?*

HYPOTHESIS OF THE RESEARCH

The main hypothesis is: *Climate change, with its rising temperatures and evaporation, is reducing the Kaléta dam's potential for generating electricity.*

❖ The specific hypothesis

H1: Rising temperature and evaporation negatively affect electricity generation at the Kaléta dam.

H2: A decrease in rainfall due to climate change reduces the amount of water available for electricity generation at the Kaléta dam.

MAIN AND SPECIFIC OBJECTIVE

The main objective is to assess the impact of future climate change on hydroelectric power generation at the Kaléta dam.

To achieve the main objective, we need to accomplish the following objectives:

- ✓ Analyze the variability of climatic and hydroelectric production parameters at the Kaléta dam.
- ✓ Evaluate the future changes in climatic parameters at the Kaléta dam under various Shared Socioeconomic Pathways (SSPs) scenarios

- ✓ Predict hydropower production under different SSP2-4.5 and SSP5-8.5 scenarios using machine learning methods.

THESIS STRUCTURE

This thesis comprises a general introduction, followed by three chapters, and then the general conclusions of the study. The content is structured as follows:

1. **The introduction** presents the overall context and justification of the study. The problem statement is followed by research questions, research hypotheses, and research objectives.
2. **Chapter 1: The literature review** presents the main results of related works.
3. **Chapter 2: Materials and Methods**, which describes the Study Area and the Dataset, and the methodology of research.
4. **Chapter 3: Results and discussion**. This chapter presents the results and discussion related to each specific objective.
5. **Conclusion and perspectives**. This last part concludes the thesis while recalling the objectives and how this work can be deepened.

Chapter 1 LITERATURE REVIEW

1.1 INTRODUCTION

Climate change ranks among the most pressing global issues, profoundly affecting sectors such as energy production. The Intergovernmental Panel on Climate Change (IPCC) emphasizes that climate change, primarily driven by human-induced greenhouse gas emissions, leads to rising temperatures, altered precipitation patterns, and increased climate variability (IPCC, 2021). Such changes threaten hydropower generation, which is heavily reliant on stable water resources. In the Republic of Guinea, hydropower is a primary source of electricity, notably through the Kaléta Dam, which is a crucial component of the country's energy production. However, climate change risks disrupting the hydrological cycle, impacting water availability for power generation Kouadio et al. (2020). Likewise, Mailhot et al. (2019) point out that rising evaporation rates from higher temperatures further reduce reservoir water availability. Moreover, Kouadio et al. (2020) suggest incorporating climate predictions into energy planning to prepare for future fluctuations in hydropower capacity. The effects of climate change on hydropower production have been extensively examined in multiple regions, particularly in West Africa. The changes in rainfall, temperature patterns, and shifts in water systems are expected to strongly affect the energy sector. A range of modelling techniques, including machine learning and hydrological models, have been utilized to evaluate these shifts, revealing both regional and temporal disparities in hydropower potential.

1.2 HYDROPOWER IN WEST AFRICA

In West Africa, numerous studies have investigated the effects of climate change on hydropower. Olabanji et al. (2020) highlighted that under the RCP 4.5 and 8.5 scenarios, temperatures are likely to increase by 1°C to 4°C, accompanied by a 5% to 30% decrease in precipitation compared to baseline scenarios, leading to potential reductions in hydropower capacity. Similarly, Obahoundje et al. (2022) employed an ensemble machine learning model, specifically an Artificial Neural Network, to assess climate variability and hydropower generation in West Africa, with a specific focus on the Nagbeto hydropower plant. Their study found that energy production and inflow are projected to decrease by up to 20.5% and 8.5%, respectively, under future climate scenarios.

(Ekwezuo et al. (2018) Also observed that the changes in precipitation over West Africa, particularly during the dry seasons, will have significant implications for hydropower generation. Using climate models (NorESM1-M), they projected an increase in precipitation during the dry months of September to November under RCP 4.5, which could potentially alter seasonal water availability for energy generation. Kouadio et al. (2020) performed an Inverse Distance Weighted (IDW) analysis of the impacts of climate variability on the Kossou Dam in Côte d'Ivoire, showing that projected changes in precipitation and flow from 2030 to 2050 would likely reduce hydropower generation due to altered wet and dry season dynamics.

Further research in the Densu River Basin by (Oti et al (2020) used the WEAP21 model to project that by 2051-2080, temperature could rise by up to 8.23%, while precipitation could decrease by 17%, leading to a significant reduction (58%) in water resources. These findings underscore the need for adaptive water management strategies to mitigate climate-induced risks to hydropower in the region. The UNFCCC (2007) and UNESCO (2020) also emphasized that many African nations are facing escalating water stress, with climate change contributing to more frequent droughts and desertification, especially in the Sahel region.

1.3 HYDROPOWER IN OTHER REGIONS

Beyond West Africa, global studies further support the growing vulnerability of hydropower to climate change. Hydropower remains one of the most important renewable energy sources globally, but its production is highly dependent on climate variables, especially precipitation and temperature. In China, Fan et al. (2020) predicted a reduction of up to 153.29 billion kWh in hydropower generation by 2100 due to changes in precipitation patterns. Hanoon et al. (2023) applied machine learning models, including Artificial Neural Network, AutoRegressive Integrated Moving Average, and Support Vector Machine, to predict hydropower generation in a Chinese reservoir using data from 1979 to 2016. The methodology involved testing three prediction scenarios (daily, monthly, seasonal), statistical pre-processing, model evaluation with five performance metrics, d-factor indicators, and uncertainty analysis using 95PPU. The sensitivity of hydropower to variations in rainfall was also emphasized by (Wang et al (2021), who used the Improved Sparrow Search Algorithm (ISSA) to optimize hydrological modeling. Their research

revealed the high irregularity of precipitation, temperature, and evaporation trends, which affect the stability and predictability of hydropower generation. Singh et al. (2022) indicate that climate-induced shifts in precipitation and streamflow are expected to pose challenges for hydropower generation. Higher temperatures can exacerbate evaporation rates and increase rainfall variability, leading to unpredictable runoff and energy production. This highlights the need for improved forecasting techniques, including machine learning models and hydrological simulations, to assess the future hydropower potential under climate change scenarios. Zhao et al. (2022) analyzed the Yalong River Basin using CMIP6 models and found that power generation may decline by 4% to 6% under future climate scenarios. Similarly, Guo et al. (2021) shows that hydropower generation will decrease under all RCP scenarios while energy demand will significantly increase, leading to a growing supply-demand gap and higher greenhouse gas emissions due to increased reliance on fossil fuels. In Portugal, Teotónio et al. (2017) acknowledged climate change's influence on water availability, suggesting that while hydropower remains cost-effective, it may lose competitiveness to solar and wind energy. Azman et al. (2021) studied Keyir Lake and concluded that higher emissions intensify climate trends, affecting rainfall and power generation, while increased temperatures elevate electricity demand. Furthermore, Clean Technology (2021) emphasized the growing complexity of predicting climate change effects on hydropower due to intricate interdependencies among climate variables, river runoff, and energy systems. Rising temperatures and evaporation rates further challenge electricity production, necessitating advanced modelling and resilience strategies in water resource management.

Climate change is strongly affecting how water moves through the environment, which influences the amount of water available and reduces hydropower production. Hanoon et al. (2023) Quantified these impacts using CMIP6 projections, finding that reduced rainfall and groundwater recharge, coupled with increased evapotranspiration, threaten water sustainability. Similarly, Waheed et al. (2024) highlighted the challenges faced by watersheds in Pakistan due to shifts in precipitation, rising temperatures, and increased evaporation, necessitating continuous assessment and adaptation strategies. In Manitoba, Kim et al. (2022) reported that climate change will cause seasonal inflow variations along the Lower Nelson River (LNR), with winter inflow and hydropower production potentially declining by 35% and 37% under extreme dry scenarios, increasing uncertainty in electricity supply. (Arora, Lima, and Shrestha, 2025) analyzed six major Canadian rivers

and found region-specific hydrological responses to climate warming, with northern basins (Mackenzie and Yukon) experiencing increased runoff, while southern basins (Fraser and Columbia) see declining runoff due to accelerated evaporation. Seasonal streamflow shifts suggest Fraser and Columbia Rivers will transition from snow-dominated to rainfall-dominated regimes by the late 21st century under RCP8.5. Meanwhile, Wei et al. (2020) projected a 12% decline in hydropower production from June to December but a 4% increase in other months under RCP4.5, with overall power generation remaining stable. These findings underscore the need for adaptive water resource management and resilient hydropower strategies in the face of climate change. Climate change poses a serious threat to energy production in West Africa, particularly in a hydropower-dependent country like Guinea. According to the United Nations Framework Convention on Climate Change (UNFCCC, 1992), human-induced warming is the primary driver of climate change, affecting various sectors, including energy generation. However, despite these commitments, climate change continues to have evident effects on hydropower systems. Hydropower, the primary electricity source in Guinea, is highly sensitive to climatic variations. Studies indicate that fluctuations in precipitation, temperature, and evaporation rates significantly influence river flow and reservoir levels. The Dam, a major hydropower facility, is particularly vulnerable to changes in rainfall patterns and increased evaporation due to rising temperatures. These climatic shifts lead to variability in water availability, ultimately impacting electricity generation. The Seimareh Hydropower Plant is vulnerable to climate change, as rising temperatures and reduced precipitation are expected to decrease dam inflow and energy production, emphasizing the need to incorporate climate change considerations into hydropower planning and design Goodarzi et al. (2020). Higher temperatures contribute to an increase in evaporation rates, reducing reservoir storage capacity (IPCC, 2021). To mitigate climate change impacts, adaptation strategies such as improved reservoir management and integration of alternative energy sources are necessary Turner et al. (2017). While climate change poses significant risks to hydropower, it also calls for innovative adaptation strategies. Future hydropower planning must incorporate robust climate models and flexible water management systems to ensure resilience in the face of uncertainty. The growing body of literature on hydropower's sensitivity to climate change, particularly in regions such as West Africa, underscores the urgent need for sustainable and climate-resilient energy solutions. As climate variability continues to

challenge hydropower generation, future studies and models will play a crucial role in mitigating its impacts and securing long-term energy availability.

1.5 CONCLUSION

This chapter reviewed studies on the impact of climate change on hydropower production, highlighting methods like hydrological modeling, climate projections, and historical data analysis. While ANNs have been used in hydropower forecasting, especially by Obahoundje et al. (2022) in Nagbeto, their application in Guinea remains limited. This research addresses that gap by applying Extreme Gradient Boosting and CatBoost to model climate-hydropower relationships at Kaléta Dam. The next chapter presents the study area and datasets.

Based on the literature review, there is little or no identified research conducted on assessing the impact of climate change on hydropower across the country, nor is there explicit analytical research on forecasting hydropower within common socio-economic pathways. It is also essential to understand how the impact of climate change affects the hydrological regime when developing projects, so that they do not overlook the consequences of climate change in forecasts, hydropower planning, and projections.

Across West Africa, several studies have been conducted using the SWAT analysis, the WEAP21 model, and machine learning models such as Artificial Neural Networks. Extreme Gradient Boosting and CasBoost offer superior performance, and they are particularly adept at understanding the intricate relationships between climate variables. Extreme Gradient Boosting and CatBoost demonstrate superior performance and can effectively handle complex relationships between climate variables and hydrological behavior. Although these models have been mostly used in other regions, their strong results led me to choose them for this research.

Chapter 2 **MATERIALS AND METHODS**

Introduction:

This chapter describes the materials and methods used in this research. To meet the study's goals, test the hypotheses, and answer the research questions, various materials and techniques were employed. First, the Kaléta hydropower plant in Guinea was chosen as the study area. This was followed by data collection and analysis processes. The study concludes with a prediction of hydropower output using two machine learning algorithms: CatBoost and XGBoost. This chapter will detail the study area and the datasets supporting this research.

2-1 STUDY AREA

The geographical focus of this study is the Republic of Guinea, a country located in the western part of Africa. Guinea is positioned between 7° and 13° North latitude and 7° and 15° West longitude, with its western border adjacent to the Atlantic Ocean. The country has a total area of approximately 245,857 square kilometres, making it a relatively medium-sized country in West Africa World Bank. (2022). Its geographical location provides a diverse landscape that plays an important role in its ecological and economic structure.

Guinea is characterized by a diverse range of landscapes, including coastal plains that stretch along its western edge, a mountainous region dominated by Mount Nimba to the southeast, and the Fouta Djallon Plateau in the central part of the country Geological Survey. (2020). The Fouta Djallon Plateau, in particular, is known for its rich biodiversity and is considered a major watershed for many rivers in West Africa. The Guinea Highlands, located in the southeastern part of the country, are the highest elevation areas, contributing to the formation of significant rivers that flow into neighboring countries.

As of 2019, the country's population was estimated at approximately 12 million people, a figure that continues to grow rapidly, exerting pressure on the country's infrastructure, economy, and natural resources World Bank. (2020). The population is predominantly rural, with many people living in the coastal and river areas; however, though urbanization is steadily increasing, particularly in the capital, Conakry. The population distribution is closely linked to the country's geographical features, with urban development concentrated around major river systems, particularly in the lowland regions.

2-1.1 Hydrology

The country has more than 1,166 rivers and is divided into 19 river basins, 13 of which are shared with neighboring countries. Additionally, six coastal watersheds comprise 24 major rivers, all of which originate from the western slopes of the Guinean ridge. Guinea is known as the "Water Tower of West Africa" because it is the source of several major transboundary river basins that flow into neighboring countries, each with its own geographical significance and hydrological role in West Africa. The several transboundary rivers took their sources in Guinea, their most important being in West Africa. The Niger River Basin is the largest, with the Niger River stretching 4,180 km. Its major tributaries from Guinea include the Tinkisso, Milo, and Niandan rivers. The Senegal River Basin originates from the Bafing and Bakoye rivers in the Guinea Highlands. The Senegal River spans 1,086 km. The Gambia River Basin is defined by the 1,120 km-long Gambia River, which originates in the Fouta Djallon region of Guinea. The Kaba River Basin is shared between Guinea and Sierra Leone. The Mano River Basin is characterized by the Mano River, which flows through Guinea, Liberia, and Sierra Leone, ultimately discharging into the Atlantic. The Konkouré River Basin is mainly located in Guinea. The Konkouré River drains directly into the Atlantic Ocean.

These rivers provide essential water resources for agriculture, hydroelectric power, and biodiversity across the region; however, unpredictable rainfall also affects agriculture and water supply. Issues such as soil degradation, water pollution due to industry, and irregular mining are faced.

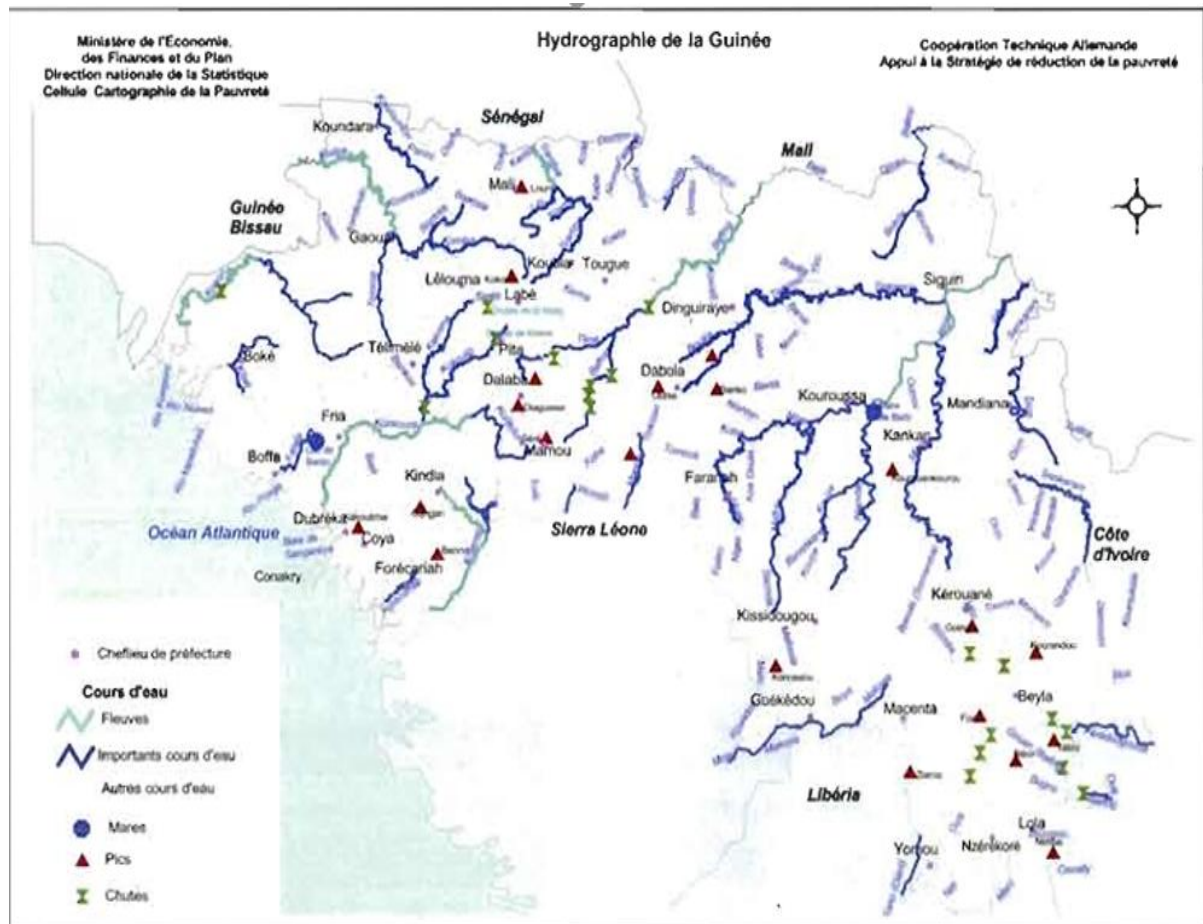


Figure 2.1: Hydrographic map of the Republic of Guinea (Samoura et al. 2011)

2-1.2 Climatology

The country experiences a tropical climate characterized by distinct wet and dry seasons of unequal duration. The dry season, during which the Harmattan wind blows, contrasts with the rainy season, which is governed by the West African monsoon. The rainy season lasts between five and seven months (April to October) from north to south. Annual precipitation averages around 1,988 mm, though it varies significantly based on latitude, longitude, topography, and continentality. For example, Conakry receives approximately 4,000 mm of rainfall, while Koundara records around 930 mm annually. The temperature varied disproportionately with precipitation, which reached its maximum in the upper Guinea (Kankan), and the northern part of lower Guinea Koundara due to its border in Senegal and also due to the exploitation of natural resources by the destruction of the natural habitat to the bauxite profile in the middle of April during which there was almost no precipitation or little precipitation during the dry season it varies also from region to region in the country. It can be observed that the maximum in the Upper region, where the climate is drier and the region is more continental almost 38 °C, and the maximum temperature in other regions

is 35 °C. While the minimum temperature in Upper Guinea is observed around 25 °C, the minimum temperature in Middle Guinea, like Dalaba, can vary between 0 to 5 °C.

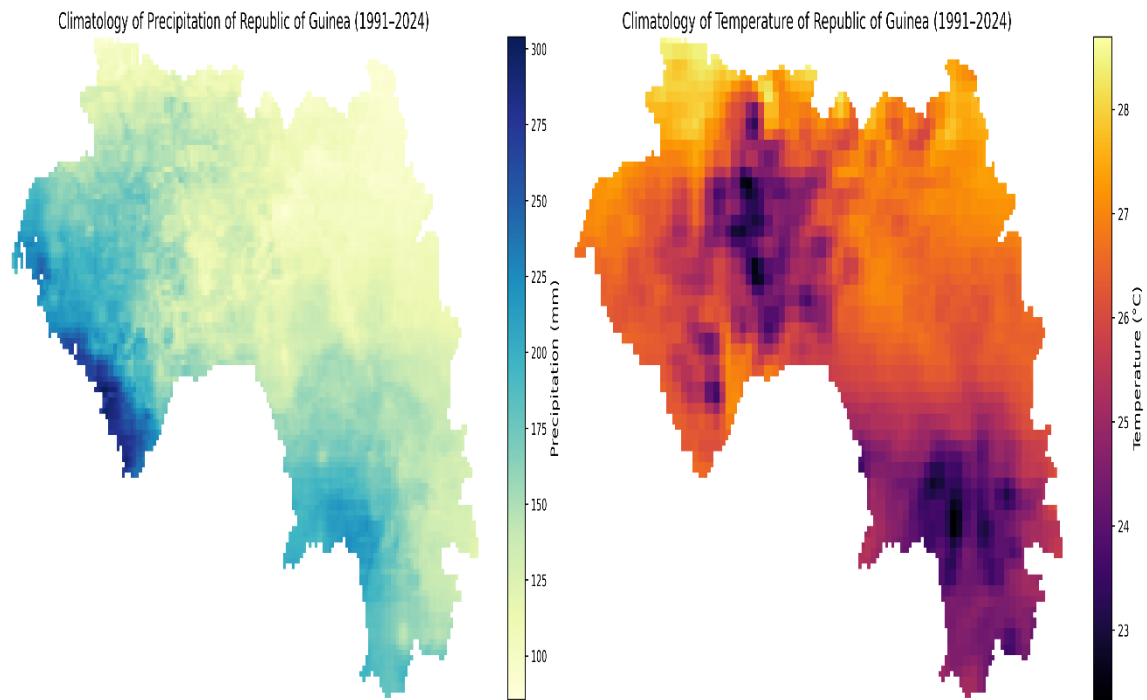


Figure 2.2: Climatology of precipitation and temperature of the Republic of Guinea

2-1.3 Forest

The dominant intact vegetation types of Guineas are woodland and grassland. The country's native vegetation is severely reduced and fragmented, and agriculture is extensive. By the end of the 20th century, 96% of the original forest had been recorded as lost, and Guinea burns from end to end during the dry season due to the practice of fresh burning to promote forage for livestock. Wetland areas have been extended and modified for the cultivation of West African rice Agnoun et al. (2012) the traditional staple crop. Anthropogenic activities such as irregular agricultural practices, illegal logging, and mining are the primary causes of forest loss and degradation.

2-1.4 Energy supply

The Republic of Guinea's energy sector is predominant hydropower/hydroelectricity, leveraging its extensive river systems. The Kaléta Dam, completed in 2015 with a capacity of 240 MW, significantly boosted the nation's electricity supply. Following this, the

Souapiti Dam, with a 450 MW capacity, began operations in 2021, aiming to further stabilize and expand Guinea's energy production. Despite these developments, the Garafiri Dam, operational since 1999 with a 75 MW capacity, often underperforms due to various operational challenges Dagne et al. (2020). However, Guinea continues to experience frequent power outages and uneven energy distribution, particularly disadvantaging rural regions where electrification rates are as low as 10-15%. The national grid, managed by Electricité de Guinée (EDG), primarily serves greater Conakry, with isolated networks providing limited service elsewhere Aljounaidi et al. (2018). To supplement hydropower, Guinea operates thermal power plants such as the Tombo Thermal Plant in Conakry. These facilities rely on imported heavy fuel oil and diesel, leading to high operational costs and environmental concerns. This dependency underscores the urgency for Guinea to diversify its energy sources and reduce reliance on imported fuels AFD (2021). In response, Guinea is exploring renewable energy alternatives, with a focus on solar and wind projects. Initiatives include the development of solar mini-grids in rural areas and assessing wind energy potential in coastal and highland regions. These efforts aim to enhance energy access and sustainability nationwide. Despite its energy potential, Guinea faces significant challenges. The national grid suffers from transmission losses and outdated equipment, hindering efficient energy distribution. Additionally, limited private sector involvement and financial constraints impede the expansion and modernization of energy infrastructure. Guinea aspires to become a regional energy hub, exporting electricity to its neighboring countries, including Mali and Senegal. The Souapiti Dam is central to this vision, aiming to double the nation's electric capacity and attract further investment. However, ensuring domestic energy security remains a critical challenge that must be addressed concurrently Dagn et al. (2020). Large-scale hydropower projects have also led to social and environmental impacts, including the displacement of communities and alterations to local ecosystems. For instance, the construction of the Souapiti Dam resulted in the displacement of thousands, raising concerns about adequate resettlement and compensation. Financing these energy projects heavily relies on foreign investment, particularly from Chinese entities. The Souapiti Dam, for example, was funded through a \$1.175 billion loan from the Export-Import Bank of China. While such investments facilitate infrastructure development, they also contribute to national debt and necessitate careful financial planning Dagne et al. (2020). Looking ahead, Guinea's energy strategy focuses on expanding hydropower capacity, integrating renewable energy sources, and improving rural

electrification. By enhancing energy efficiency, attracting private investment, and strengthening regional partnerships, Guinea aims to achieve a more stable and sustainable electricity supply in the years to come.

2-1.5 Konkouré watershed

The Kaléta Hydroelectric Dam, located 115 km northeast of Conakry and 130 km upstream from the mouth of the Konkouré River, is positioned just downstream of the valley's widening. It features a 700-meter-wide weir on a riverbed with a natural 40-meter drop. The upstream elevation ranges from 100 to 130 meters, with water release occurring at 60 meters Comité Français des Barrages et Réservoirs, (1999). The Konkouré River basin receives an annual rainfall of 1173.90 mm and 2016.06 mm, respectively, in 2018 and 2024, with 80% concentrated between July and August. The mean annual temperature remains relatively stable Coyne et al. (1999). The watershed contains three other major hydropower plants: Garafiri (built in 1999), Souapiti (built in 2020), and Amaria (currently under construction), all located upstream of Kaléta, which significantly influences the hydrological regime, as it fluctuates according to the wet and dry seasons. Environmental challenges within the watershed include land degradation, deforestation, and bushfires caused by agricultural expansion, fuelwood collection, and livestock grazing. Studies indicate that 29.50% of forested areas have been converted into built-up land Bah et al. 2024). A feasibility study by Coyne et al. (1999) proposed two potential operating configurations for Kaléta: (1) a 125-meter operational level with a 950-hectare reservoir, storing 142 hm³ of water and generating 2,183 GWh annually with an installed capacity of 356 MW, or (2) a 110-meter operational level with a 200-hectare reservoir, storing 30 hm³ of water, yielding 957 GWh annually with an installed capacity of 157 MW. The Kaléta Hydroelectric Dam officially became operational in September 2015 as a run-of-river facility with an installed capacity of 240 MW, designed to generate approximately 965 GWh of electricity per year Sakouvogui et al. (2023). Its reservoir capacity of 23 million cubic meters minimizes environmental impacts due to its relatively small inundated area (CWE). The dam's construction, led by China International Water & Electric Corporation (CWE), provided significant economic and social benefits, including improved schools, healthcare facilities, and roads. Funded through a partnership between the Guinean government and Chinese entities, the project played a crucial role in addressing regional electricity shortages and fostering local development.

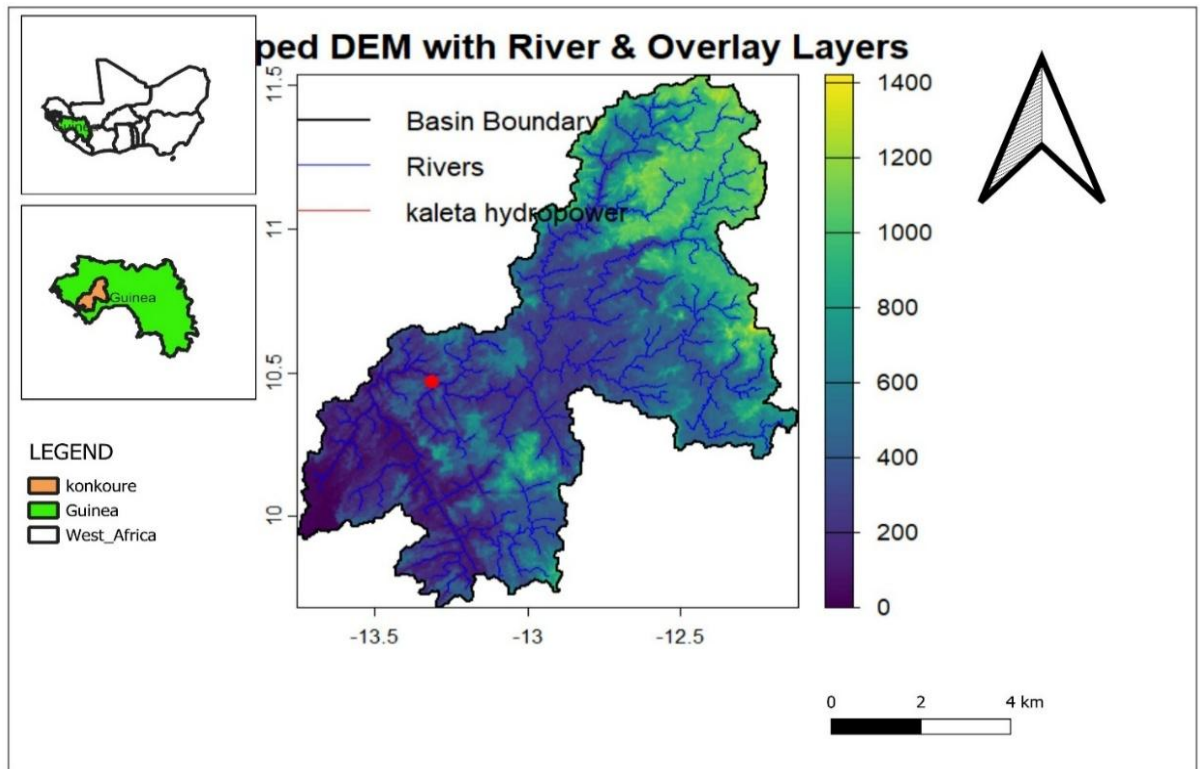


Figure 2.3: DEM of Konkoure watershed



Figure 2.4: Kaleta power plant

2.2 DATASETS

In this research, several datasets were used. Historical observation of climate data as variables, precipitation, and temperature from an observation weather station inside the watershed.

2-2.1 Climate data

2-2.1.1 Observation data

Observed precipitation and temperature data were collected from the weather station upstream of the watershed. We collected data from five stations; four stations were provided by the Agence Nationale de la Meteorologie (ANM) of Guinea: Kindia, Dubreka, Mamou, and Dalaba in the Koukoure watershed. Due to a lack of daily scale data, we just received monthly data from January 1991 to December 2023.

Additionally, we received observed precipitation data from the rain gauge station located at the Kaléta dam and operated by Société de Gestion de Kaléta (SOGEKA SA). Monthly data from September 2017 to December 2024 were collected.

2-2.1.2 Reanalysis and satellite products data

To complete the ground observation data, we downloaded:

- Monthly maximum and minimum temperature from ERA5 land reanalysis data
- Monthly precipitation from CHIRPS
- Monthly downward longwave radiation flux and near-surface wind speed from NASA_FLDAS_NOAH01),
- Monthly potential evapotranspiration of GLEAMv4.

2-2.1.3 CMIP6 projection data

Instead of CMIP5, which deals with RCP data projection, we used in this study the new Shared Socioeconomic Pathways (SSPs) climate projections. These projections are part of the Coupled Model Intercomparison Project Phase 6 (CMIP6), which brings together advanced climate models from research institutions worldwide.

The SSPs used in these projections are SSP2-4.5 and SSP5-8.5, which represent medium or intermediate emissions (business as usual, moderate policies) with 4.5W/m² radiation forcing. Additionally, there are high-emission or very high-emission (fossil fuels intensive

worst-case scenario) with $8.5\text{W}/\text{m}^2$ radiation forcing resulting in approximately $2.7\text{ }^\circ\text{C}$ and $4.4\text{ }^\circ\text{C}$, respectively. Each of the two (2) scenarios outlines a different future based on economic growth, energy usage, policy decisions, and environmental sustainability efforts. SSP2-4.5 assumes moderate mitigation efforts, resulting in a moderate emissions trajectory. SSP5-8.5 represents the worst-case scenario, depicting a world with rapid fossil fuel expansion, high emissions, and minimal climate policy intervention, resulting in severe global warming impacts at the end of the century.

We considered six CMIP6 models, which were generated according to climate projections under these scenarios, each offering distinct capabilities for climate projections and impact assessments:

- MIROC6 (*Model for Interdisciplinary Research on Climate, version 6*) developed by the Japan Agency for Marine-Earth Science and Technology (JAMSTEC). MIROC6 features a spatial resolution of $0.9^\circ \times 0.9^\circ$ (100 km grid spacing). This model is widely used in climate impact assessments due to its advanced simulation of atmospheric and oceanic processes.
- MPI-ESM1-2-HR (*Max Planck Institute Earth System Model, High Resolution*) created by the Max Planck Institute for Meteorology. MPI-ESM1-2-HR operates at a resolution of $0.9^\circ \times 0.9^\circ$ (100 km grid spacing).
- MRI-ESM2-0 (Meteorological Research Institute Earth System Model, version 2.0) from Japan's Meteorological Research Institute, which has a resolution of $0.9^\circ \times 0.9^\circ$ (100 km grid spacing).
- NESM3 (*Nanjing University of Information Science and Technology Earth System Model, version 3*) was developed by the Nanjing University of Information Science and Technology (China), with a $0.9^\circ \times 0.9^\circ$ grid spacing (approximately 100 km).
- NorESM2-MM (*Norwegian Earth System Model, Medium Resolution*), developed in Norway, also features a resolution of $0.9^\circ \times 0.9^\circ$ (100 km grid spacing).
- TaiESM1 (*Taiwan Earth System Model, version 1*) was developed by Taiwan's Research Center for Environmental Changes, with a $0.9^\circ \times 0.9^\circ$ grid spacing (approximately 100 km).

These models collectively provide high resolution and accuracy, enabling researchers to analyze temperature variations, precipitation trends, and hydrological changes in most regions of the world. They deliver critical data on temperature, precipitation, and extreme

weather events across various future periods, including the near-future (2031-2061) and far-future (2071-2100).

Data from CMIP6 projections are extensively used in several studies, such as those by Kim et al. (2020) for climate risk assessments, Mohammed et al. (2024) for agricultural forecasting, and Zhao et al. (2022) climate change impact assessment on hydropower generation worldwide. The accuracy of these models is continually enhanced through data assimilation techniques, historical validation, and ensemble modeling approaches that compare multiple model outputs to improve reliability.

Data were downloaded through the WRSB website (<https://esgf-metagrid.cloud.dkrz.de/esgf-idp/openid/balkoro>).

2-2.2 Hydrological and hydropower Data

Monthly hydrological hydropower data for Kaléta dam were collected from SODEKA from April 2015 to December 2024:

- water level (m),
- production flow (m³/s),
- hydroelectricity effective production (10⁴ kWh)
- outflow (m³/s),
- volume of production (10⁶ m³),

Volume-based metrics, measured in millions of cubic meters, track total monthly released discharge, water used for electricity generation, and total outflow, offering a cumulative perspective on water management. The dataset also includes long-term average discharge at the design sites, providing historical context for water availability trends.

2-2.3 Summary of the data used in this research

To analyze the impact of climate variability on hydropower production at Kaléta Dam, this study relies on a combination of climatic and hydrological datasets. These datasets cover variables such as precipitation, temperature, evapotranspiration, water inflow, and energy production. They have been collected from various reliable sources, including global

climate databases and national agencies. The table below summarizes the key variables used in this research, along with their respective data sources.

Table 2.1: Data sources and variables

N	data	description	source
1	Pr	precipitation	ANM, NASA_FLDAS, ESGF, 25 km
2	Tasmin	Temperature minimum 2m above the ground	ANM, ERA_monthly2. (25 km), ESGF (100 km),
3	Tasmax	temperature maximum at 2m above the ground	ANM, ERA_monthly (25 km), ESGF (100 km)
4	Pet	Potential evapotranspiration	GLEAMv4_monthly (25 km),
5	rlds	Surface Downwelling Longwave Radiation	ESGF (100 km), NASA_FLDAS (25 km)
6	sfwind	near-surface (usually, 10 meters) wind speed.	ESGF (100km), NASA_FLDAS (25 km)
7	Production flow, effective production, water level, Volume of production, outflow	Hydroelectric variables	SODEKA

2-2.4 Data correlation

The correlation heatmap below shows how climate and hydropower variables are related from 2016 to 2024. Red indicates strong positive correlations, while blue represents negative ones. The focus is on identifying which factors most influence hydropower performance, especially effective production and production flow. Effective Production has the highest positive correlations with volume of production, outflow, and water level, with values of 1.00, 0.82, and 0.53, respectively, suggesting that greater water production, higher outflow, and elevated reservoir levels lead to increased power generation. Conversely, temperature 2m max and production flow show negative correlations, with values of -0.48 and -0.47, respectively, indicating that higher temperatures and excessive flows may reduce production efficiency, possibly due to evaporation or system losses.

Likewise, production flow is positively correlated with outflow at 0.79 and volume of production at 0.53, indicating that more available water leads to stronger flows. However, it has strong negative correlations with temperature 2m max at -0.64, potential evapotranspiration at -0.47, and "downward longwave radiation flux" at -0.64. These findings suggest that heat and evaporative factors restrict water availability, which negatively affects both flow and energy output.

Figure 2.3 shows the correlation heatmap between climate and hydropower variables from 2016 to 2024.

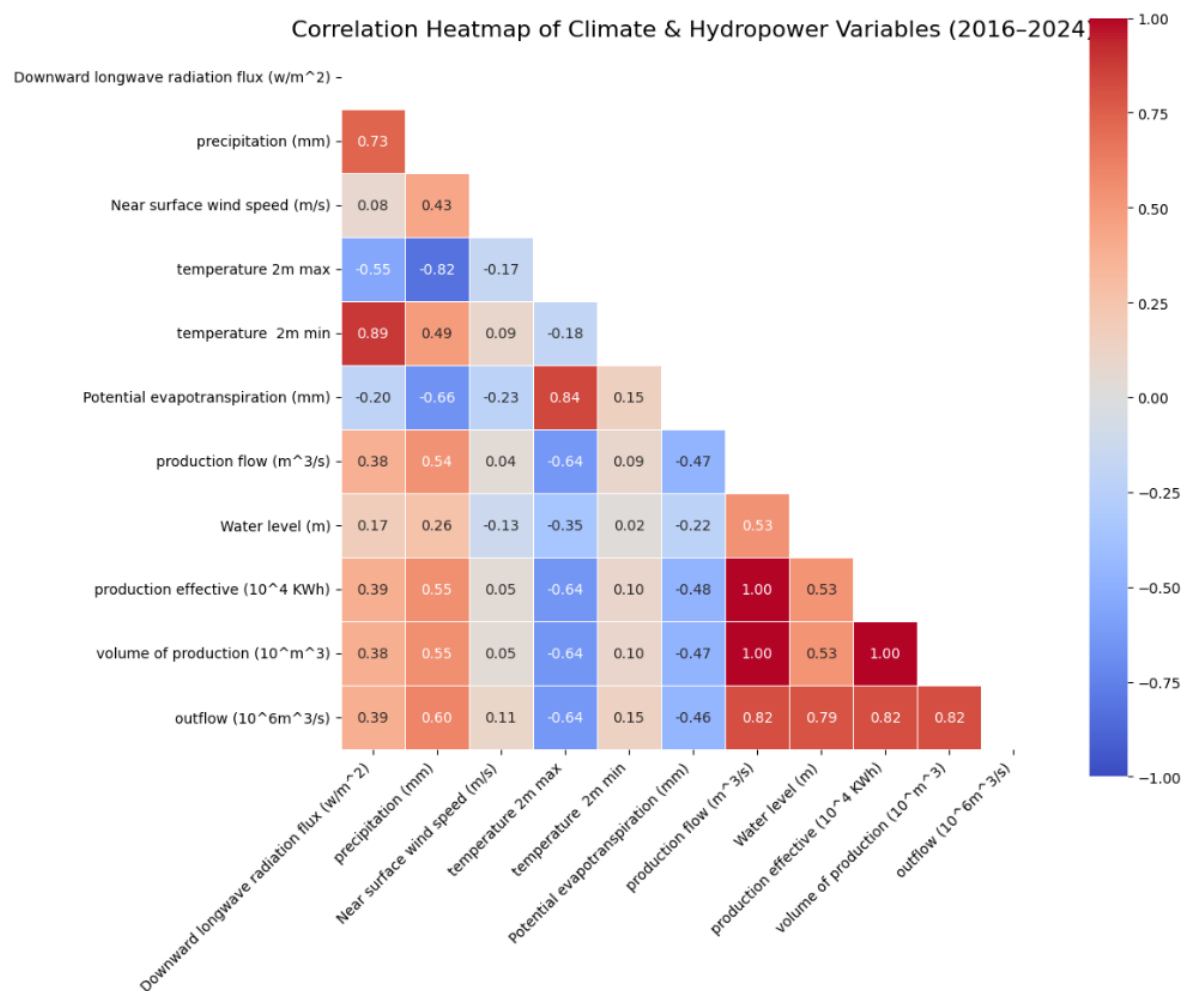


Figure 2.5: Correlation heatmap of climate and hydropower variables 2016-2024

2-3 METHODOLOGY

This section describes the steps followed to meet the research objectives. The first step is data preprocessing, where raw data is cleaned and prepared for analysis. The next phase

focuses on analyzing anomaly changes from normal patterns in climate and hydrological variables, as well as identifying long-term trends using time series analysis. A Taylor Diagram is also used to visually compare how well different models or datasets match real observations. Additional plots of anomalies are included to support the analysis. Finally, the study applies machine learning methods, specifically CatBoost and Extreme Gradient Boosting, to better understand and predict key patterns, where the optimized set of hyperparameters is selected. Finally, the last phase focuses on validating the models and conducting tests. Then comes the second phase, machine learning algorithm, CatBoost, and Extreme Gradient Boosting calibration, which involves the machine learning algorithm CatBoost and an Extreme Gradient Boosting calibration, where an optimized set of hyperparameters is selected. Finally, the last phase deals with validation of the models and testing. The detailed workflow for this objective is further elaborated below.

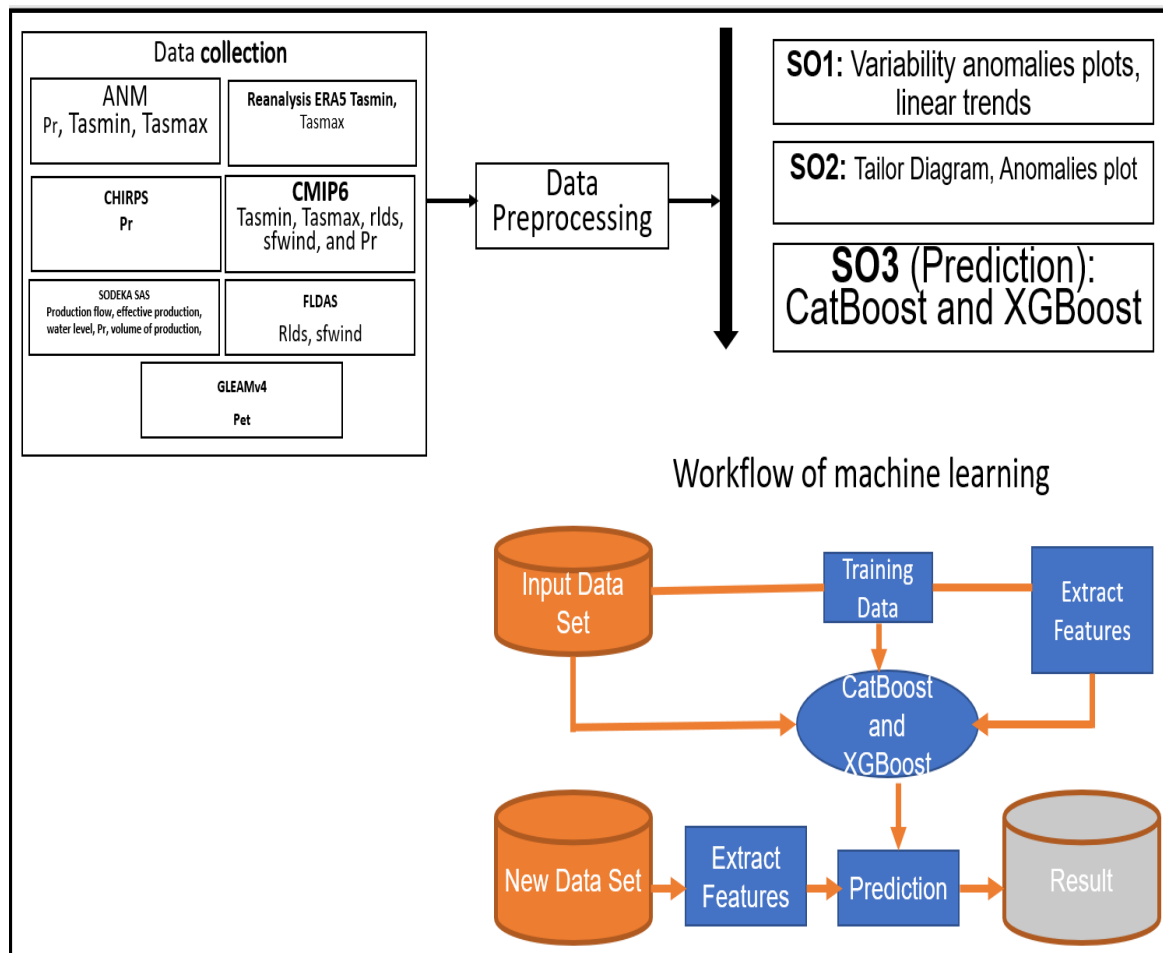


Figure 2.6: Summary of methodology

2-3.1 Processing Tools

The processing of these data required the utilization of software programs that run in a Linux environment (CDO, FERRET) and a Windows environment, as well as Python, and Google Earth Engine.

- **The Climate Data Operator (CDO)** is a software package developed by the Institute Für Meteorologie (IFM) designed for the standardized processing of outputs from climate models and weather forecasts. This software is a powerful tool with over 400 operators, including functions for statistical analysis and arithmetic operations, data selection, as well as tools for subsampling and spatial interpolation. Best of all, it is free software and open source.
- **Ferret** is an interactive tool for data visualization and analysis, developed by the **National Oceanic and Atmospheric Administration (NOAA)**. It is specifically designed to support oceanographers and meteorologists in analyzing large and complex gridded datasets. Ferret is especially well-suited for working with data in **NetCDF format** and is available as free software.
- **Google Earth Engine** is an attractive tool which were used to download climate data from ERA5, CHIRPS, and FLDAS.
- **Python** was used as the programming language for data analysis and model development on the Windows operating system. The programming interface was provided through **Jupyter Notebook**, an open-source web application that helps has to extract the NetCDF files to CSV files, supports interactive computing, and allows for the creation and sharing of live code, equations, visualizations, and narrative text. For machine learning applications, focusing on two widely used machine learning algorithms: **CastBoost** and **Extreme Gradient Boosting (XGBoost)**. Keras, built on top of TensorFlow, offers a straightforward and flexible approach to defining and training models. It was initially developed for research purposes and to evaluate models for predicting outcomes based on meteorological data, leveraging their robustness, interpretability, and strong performance on complex datasets.

2-3.2 Data preprocessing

Since this research focuses on the impact of climate change on hydropower and its

corresponding change in electricity production, the hydrological data collected were in CSV format. Therefore, we need to obtain the NetCDF dataset in CSV (ideally in a CSV file).

We used Ensemble mean with the best CMIP6 models for our study area with Climate Data Operator (CDO). We first resample daily data to monthly data from NorESM2-MM and CHIRPS. After that, we converted them to a CSV file for both scenarios, SSP2-4.5, SSP5-8.5, and finally selected data raw from 2026 to 2034

The potential evapotranspiration was computed from the minimum and maximum temperatures of the two scenarios, SSP2-4.5 and SSP5-8.5, using the Thornthwaite equation (Eq.1). In several studies, we find that more variables can affect potential evapotranspiration, such as relative humidity, solar radiation, soil moisture, and so on. But because of the lack of observation data on these variables from the observation station, we considered this simple equation based on temperature to estimate the potential evapotranspiration for the future (2026 to 2034) with the two scenarios, SSP2-4.5 and SSP5-8.5.

The Thornthwaite equation below is an empirical formula that estimates potential evapotranspiration (PET) based on temperature and latitude. It integrates climate and physical parameters. It was used in various research domains, including agriculture, evolutionary trends analysis, and to evaluate water loss evaporation in Ethiopia's reservoir under a warming scenario Yohannes et al. (2021).

$$PET = 16 \left(\frac{10T}{I} \right)^a \quad (1)$$

where:

- *PET* is the potential evapotranspiration (mm/month);
- *T*, the mean monthly temperature (°C) and
- *I*, the annual heat index (°C), calculated with each monthly mean temperature *T_i*(°C) as:

$$I = \sum_{i=1}^{12} \left(\frac{T_i}{5} \right)^{1.514} \quad (2).$$

2-3.3 Understanding Machine Learning Models

A machine learning model is a computational algorithm that enables a system to identify patterns in data and make decisions or predictions without requiring explicit programming Mitchell, (1997). The algorithms of this nature work by training them on past examples, thus enabling them to generalize and discover relationships when dealing with new data Bishop (2006). Machine Learning models are the fundamental parts of AI, the domain concerned with mimicking human intelligence, unlike rule-based systems Goodfellow, (2016). Machine learning has demonstrated its power in various areas, most notably in hydrology, meteorology, and energy forecasting, as it can process large datasets and represent complex and nonlinear trends Shrestha and Solomatine, 2006). Machine learning models can be more accurate than traditional statistical models because they can continuously adjust to the incoming data patterns Abrahart et al. (2012). The application of machine learning in environmental sciences has led to improved water management, more accurate weather forecasts, and more efficient energy strategies for production Chen and Guestrin, 2016). Two (2) algorithms in machine learning were performed to reach these objectives. It has several algorithms of machine learning algorithms, but we choose these two (2) because of their robustness, accuracy. There are CastBoost and Extreme Gradient Boosting models.

2-3.3.1 CastBoost Model

A Gradient Boosting Approach Optimized for Categorical Features

CatBoost is a supervised machine learning algorithm based on gradient boosting over decision trees, employed in tools such as Train Using AutoML for both classification and regression tasks. As its name implies, *CatBoost* is optimized explicitly for *categorical* data ("Cat") and leverages *boosting* techniques ("Boost") to enhance predictive performance.

In gradient boosting, an ensemble of decision trees is constructed sequentially, where each subsequent tree attempts to correct the residual errors of the previous one. This iterative refinement improves model accuracy over time. CatBoost builds upon traditional gradient boosting frameworks by introducing algorithmic optimizations that lead to improved speed and stability.

One of CatBoost's key innovations lies in its native support for categorical variables. Unlike conventional tree-based algorithms, which often require categorical features to be transformed through one-hot encoding or label encoding, CatBoost inherently processes both categorical and numerical features without manual preprocessing. This is achieved via a technique known as ordered target encoding, wherein each categorical value is encoded using statistics derived from preceding observations, thereby minimizing target leakage and overfitting.

Another distinctive feature of CatBoost is its use of symmetric trees, where a consistent split condition is applied across all nodes at the same depth. This structure contributes to computational efficiency and facilitates faster inference.

Compared to other gradient boosting frameworks such as XGBoost or LightGBM, CatBoost often demonstrates competitive or superior performance, particularly when dealing with datasets that contain a high proportion of categorical variables. Its robust handling of categorical data and algorithmic innovations make it a strong candidate for a wide range of predictive modeling tasks. Şebcioğlu et al. (2025) and Kumar et al. (2023) performed this machine learning model for his studies, respectively, the comparative traditional hydrological forecasting model with the CatBoost algorithm and advanced machine learning techniques to improve hydrological prediction, a case of comparative analysis of streamflow prediction models Shahani et al. (2022). Figure 2.4 below shows the CastBoost model explanation.

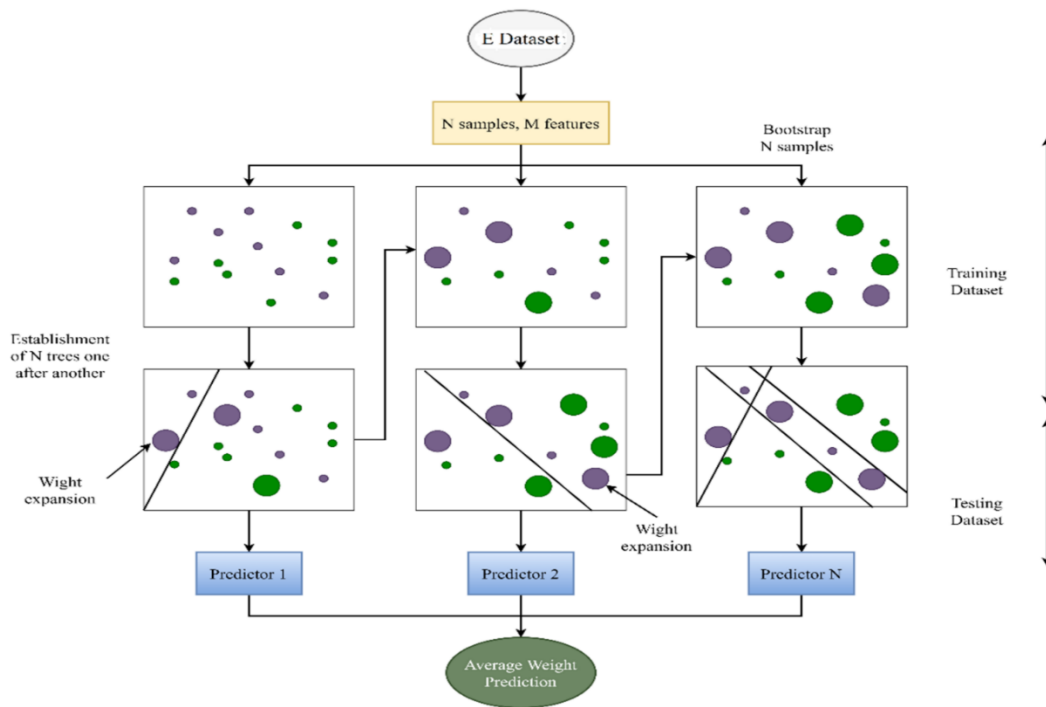


Figure 2.7: CatBoost model

2-3.3.2 XGBoost (Extreme Gradient Boosting)

XGBoost (Extreme Gradient Boosting) is a fast and powerful machine learning algorithm that is widely used for predictive tasks. It is based on the gradient boosting technique and was developed by Tianqi Chen and Carlos Guestrin to improve speed, accuracy, and scalability Chen and Guestrin (2016). XGBoost is especially popular for tasks such as regression (predicting continuous values), classification (predicting categories), and ranking. It performs well even when working with large datasets, many variables (high dimensionality), and missing values Friedman (2001).

One of the primary improvements in XGBoost over earlier boosting algorithms is the incorporation of regularization techniques, including L1 and L2 penalties. These help to prevent overfitting and allow the model to generalize better to new data Chen and Guestrin (2016). Another important feature of XGBoost is its use of a second-order Taylor expansion to more accurately estimate the loss function. This leads to better gradient updates and more efficient tree building Ke et al. (2017).

XGBoost builds a strong model by combining many weak learners, usually small decision trees, in a sequence. Each new tree is trained to correct the errors made by the previous ones. This step-by-step process helps reduce the overall error. While this idea is similar to how neural networks reduce errors (for example, using mean squared error for regression), XGBoost does not use backpropagation. Instead, it employs a gradient-based method to enhance each tree.

Its development is connected to the broader field of ensemble learning, where multiple models are combined to enhance performance. A significant milestone in this area was the work by Friedman (2001) , who introduced gradient boosting, a method that builds a strong model from weak learners such as decision trees.

Since then, it has evolved into one of the most effective and widely used tools in machine learning. It has been successfully applied in numerous fields, including hydrology, climate forecasting, and energy prediction Zhang et al. (2018). For instance, in hydropower modeling, it has excelled at predicting water inflow, evaporation, and electricity generation by learning complex, nonlinear patterns in climate data Sun et al. (2021).

It also features many practical tools for addressing real-world data problems. These include the ability to handle missing values automatically, deal with unbalanced datasets by applying weights, and efficiently process large datasets. Because of its flexibility, high accuracy, and computational speed, it often outperforms older boosting methods and other ensemble techniques in predictive modeling tasks. As is shown in Figure 2.5 XGBoost model explanation

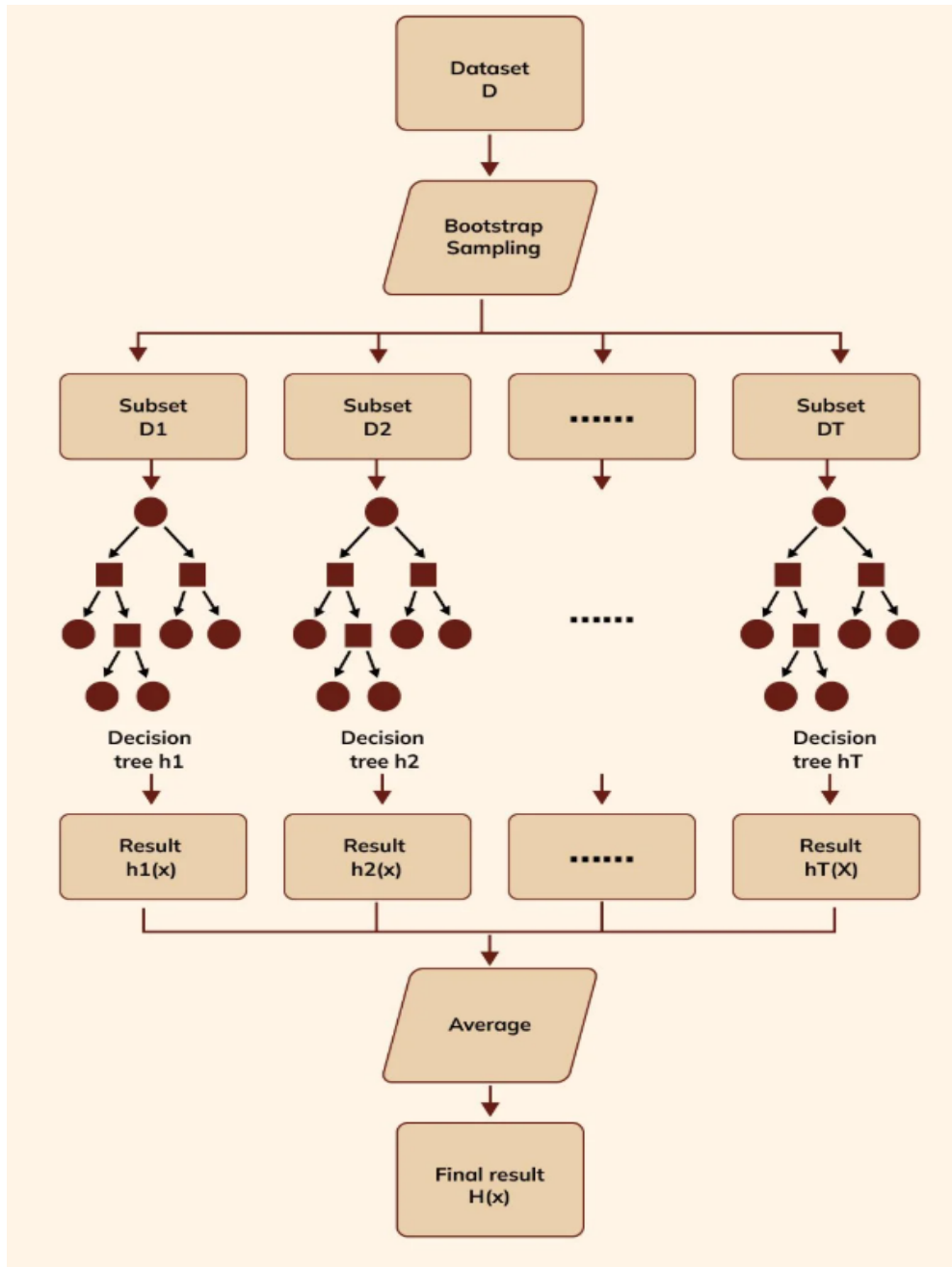


Figure 2.8: XGBoost algorithm explanation(<https://www.analytixlabs.co.in/blog/gradient-boosting-algorithm/>)

2-3.3.3 Training and validation

Model training and validation are a crucial step in the machine learning process. At this stage, data are modeled for training, and the model’s effectiveness is assessed on a separate dataset for validation. The goal is to develop a model that accurately predicts the futures and generalizes well to new inputs. To assess the impact of climate variability on hydropower generation, a data-driven modeling approach was employed. The historical

dataset comprising climatic and hydrological variables was partitioned into training (80%) and testing (20%) subsets to ensure robust model evaluation. Several advanced machine learning algorithms, Extreme Gradient Boosting (XGBoost) and CatBoost, were trained using the historical observations of inflow, evaporation, and electricity production. The model training and validation procedure is summarized as follows:

Data split: a training set, a validation set, and a test set are each provided with a separate dataset. The models are trained using the training set. The validation set is used to fine-tune the model and assess model performance throughout training, whereas the test set is used to measure the trained model's final performance on unseen data.

Model training: Develop the selected model using the training dataset. During the training phase, the model's parameters are frequently improved to minimize a chosen loss error function. To do this, training data are fed into the model, predictions are generated, and compared to actual values, the model's parameters are then updated based on the computed errors. This procedure continues until the convergence requirement is satisfied, after a certain number of epochs.

Model validation: Using the validation datasets, evaluate the performance trained model. The validation data is used to generate predictions, which are then compared to actual results. Several assessment measures are employed, including mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and R-squared (R^2). Model performance was evaluated using R^2 , RMSE, and MAE metrics. Hyperparameters were optimized through grid search and random search methods.

Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N (Y_j - \hat{Y}_j)^2} \quad (3)$$

Normalized Root Mean Squared Error (nRMSE)

$$nRMSE = \frac{RMSE}{\sigma} \quad (4)$$

Pearson correlation

$$corr = \frac{n \left(\sum_{j=1}^N y_j \hat{y}_j \right) - \left(\sum_{j=1}^N y_j \right) \left(\sum_{j=1}^N \hat{y}_j \right)}{\sqrt{\left[n \sum_{j=1}^N y_j^2 - \left(\sum_{j=1}^N y_j \right)^2 \right] \left[n \sum_{j=1}^N \hat{y}_j^2 - \left(\sum_{j=1}^N \hat{y}_j \right)^2 \right]}} \quad (5)$$

R-square

$$Rsquared = Corr^2 \quad (6)$$

Mean Absolute Error (MAE)

$$MAE = \frac{1}{N} \sum_{j=1}^N |y_j - \hat{y}_j| \quad (7)$$

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{N} \sum_{j=1}^N \frac{|y_j - \hat{y}_j|}{y_j} \quad (8)$$

Where the N refers to the sample number, y_j to the observation, and \hat{y}_j is the simulation or model for the j th date, and σ is the standard deviation. Where the overall number of data points is n . The dependent variable's actual (observed)

The value for the i th data point is represented by y_i . The predicted value of the dependent variable for the i th data point is represented by \hat{y}_i . S stands for the total sum, or the sum of the squared differences for each data point. The dependent variable's mean is represented by the symbol \bar{y} .

Iterative Refinement: to enhance performance, modify the model architecture or data preparation stages based on the evaluation findings. Until a suitable performance is attained, this iterative procedure is continued.

Final Assessment: After the model has been adjusted, its performance is evaluated using the test dataset, which simulates unseen data. This offers a neutral assessment of how well the model performs in realistic situations. To ensure accurate and trustworthy model training and assessment, it is essential to remember that correct data preparation, including managing missing values, feature scaling, and addressing class imbalance, must be

performed during the training and validation processes. These processes may be efficiently used to train, validate, and assess machine learning models, to create reliable and accurate prediction models.

Conclusion

In this chapter, the Study Area and the Datasets were described. The approach adopted for this research is two machine learning methods and their descriptions. The next chapter presents the results and discussion.

Chapter 3 RESULTS AND DISCUSSION

Introduction

This chapter presents and discusses our findings. The anomalies of precipitation (rainfall), temperature minimum, temperature maximum, and potential evapotranspiration. It further assesses the CatBoost and XGBoost within prediction under two shared socioeconomic pathways (245 and 585). All models were developed using appropriate input variables. The models were trained and tested using data from 2016 to 2024. The prediction performance of models was evaluated by comparing observations and predictions of production flow and effective production.

The structure of this chapter is as follows. Feature selection variability of climate variables and variability of hydroelectric production anomalies of precipitation, temperature minimum, temperature maximum, and potential evapotranspiration are shown in section 1. The Tailor Diagram and the coefficient of correlation in section 2. The monthly and annual changes in the production flow, effective production, and water level variables are presented in Section 3. The assessment of variables on effective production and production flow by the SHAP model in section 4. Finally, the training, testing, and prediction from the historical period (2016, 2024) to the future period 2026 to 2034 in section 5.

3-1 *Variability Of Climatic Parameters*

Among the climatic parameters, we propose to analyze the variability of precipitation, temperature (maximum and minimum), and potential evapotranspiration at the Kaléta dam. These are the parameters most directly correlated with the hydroelectric production potential of the dam's facilities.

The analysis focuses on average monthly values applied to the Konkouré watershed, where the Kaléta dam is located. Figures 3.1 to 3.4 show, respectively, climatic anomalies in cumulative annual precipitation values, mean annual maximum and minimum temperatures and cumulative annual potential evapotranspiration (PET). The analysis period is from 1990 to 2024. As a reminder, temperature data are from ERA5; GLEAM for PET and CHIRPS for precipitation.

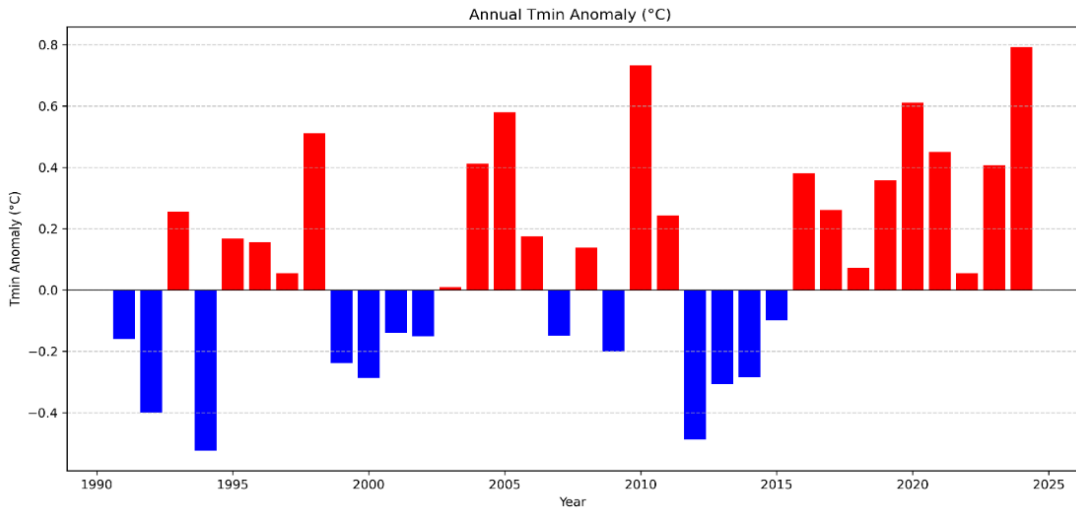


Figure 3.1: Anomaly of rainfall bar chart in Koukoure watershed

An analysis of the monthly rainfall shows significant interannual variability. Positive anomalies occurred in several years, while negative anomalies indicate drier-than-normal years. Notably, 1994 recorded the highest positive anomalies, suggesting above-normal precipitation, while 2002 experienced pronounced negative deviations, indicating potential drought conditions.

In Figures 3.2 and 3.3, we can observe the changes in both Tasmin and Tasmax.

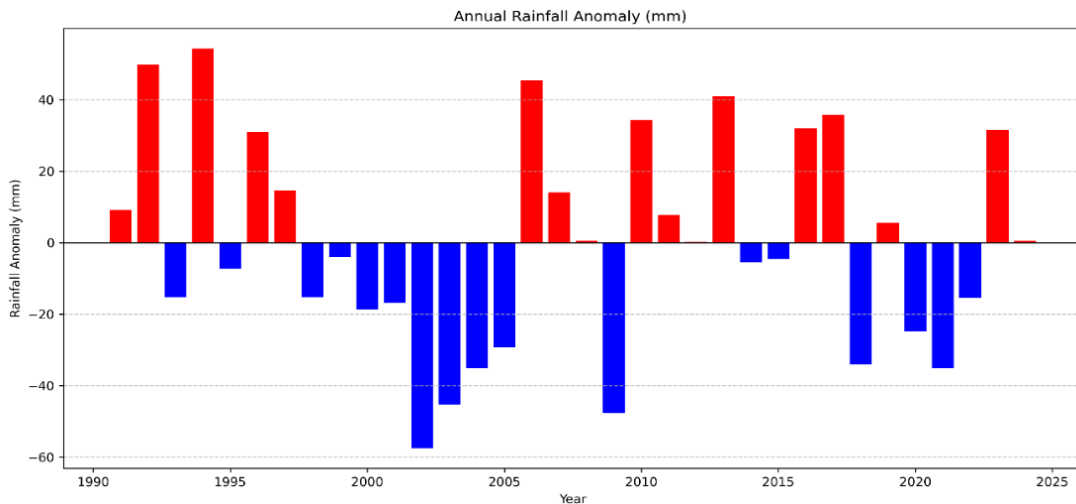


Figure 3.2: Annual Anomaly of minimum temperature (Tasmin) in Koukoure watershed

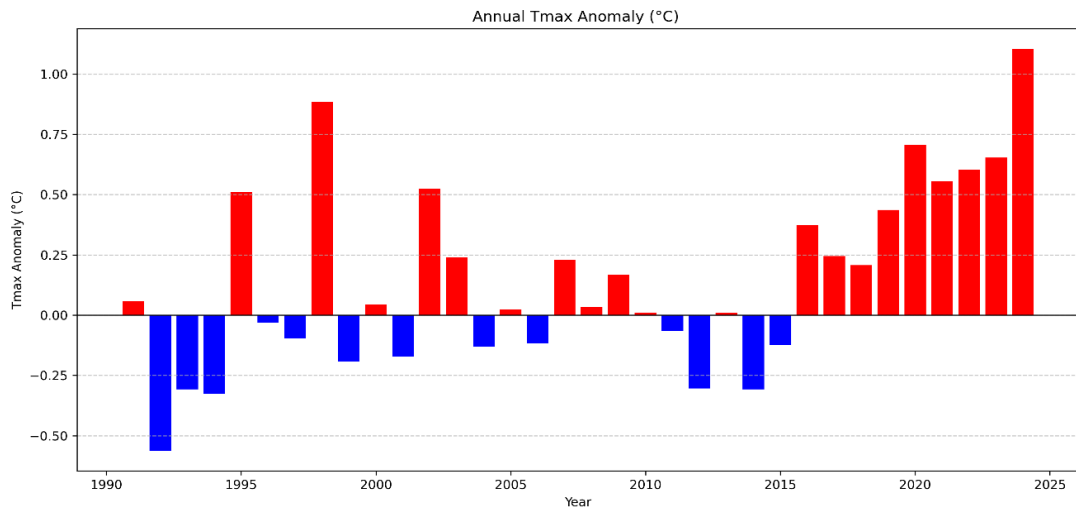


Figure 3.3: Annual Anomaly of Tmax in Koukoure watershed

Both maximum and minimum temperatures exhibited a general upward trend in their anomalies over time. Years showed notable increases in temperature maximum and minimum at 2 meters above the soil, indicating warming conditions. This trend is consistent with broader regional and global patterns of climate change. The persistent positive anomalies in recent years may reflect ongoing global warming and its impact on local temperature regimes.

Figure 3.4 illustrates the potential evapotranspiration in the Koukoure River basin.

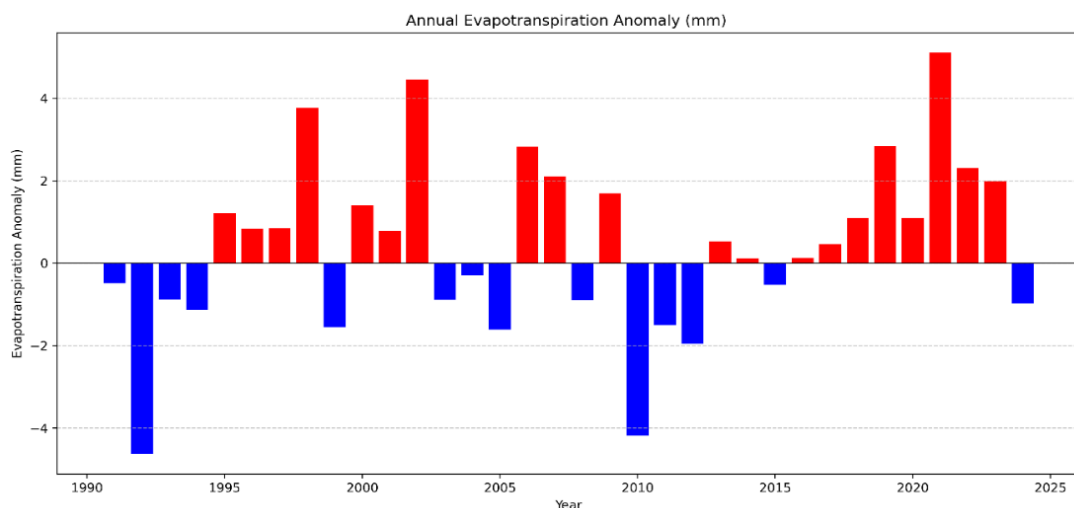


Figure 3.4: Annual Anomaly of Potential evapotranspiration in Koukoure watershed

Evapotranspiration anomalies largely mirrored the temperature trend, with positive

anomalies becoming more frequent after the baseline period. Higher evapotranspiration rates may result from increased temperatures and changes in atmospheric demand, potentially intensifying water stress, especially during dry years.

3.2. Variability of hydroelectric production parameters

An analysis of the variability in production flow, effective production, and water level was conducted on a monthly and annual basis from 2016 to 2024. The results are presented with both annual and monthly variations, reflecting the dam’s hydroelectric production capacity.

3-2.1 Monthly change in production flow, effective production, and water level

The monthly analysis of production flow, effective production (electricity production), and water level at the Kaléta Hydropower Plant for the period 2016 to 2024 was conducted using box plots. Each boxplot represents the distribution of monthly production flow, energy production, and water level values over the years. The central tendency of the data was highlighted using two markers: the mean (presented by a red diamond) and the median (presented by a blue circle).

Figure 3.5 shows the monthly distribution of production flow throughout the study period.

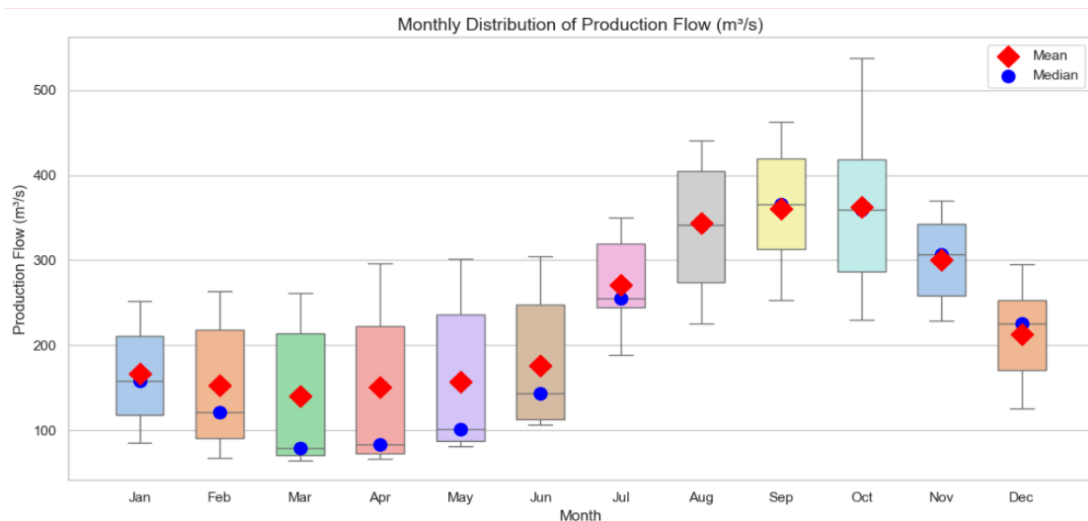


Figure 3.5 : Monthly distribution of production flow

Monthly boxplot analysis of production flow at the Kaléta Hydropower Plant reveals that the highest average production flow occurs during the rainy season, specifically from August to October, with an average of around 400 to 430 m³/s. The lowest flow values are recorded during the dry months of January to March, ranging from 210 to 220m³/s. The mean and median values are similar in most months, indicating that the production flow is generally well-balanced. However, in some months, the difference between the mean and median suggests that the production flow data is not evenly distributed. This may be caused by unusual rainfall or changes in the plant’s operation. The variability in production flow is higher in the wet season, reflecting changes in river discharge and inflow into the reservoir. Dry months show less variability and more stable production flow patterns. These results confirm that water availability has a significant impact on power generation capacity.

As illustrated in Figure 3.6, effective production varies significantly from month to month.

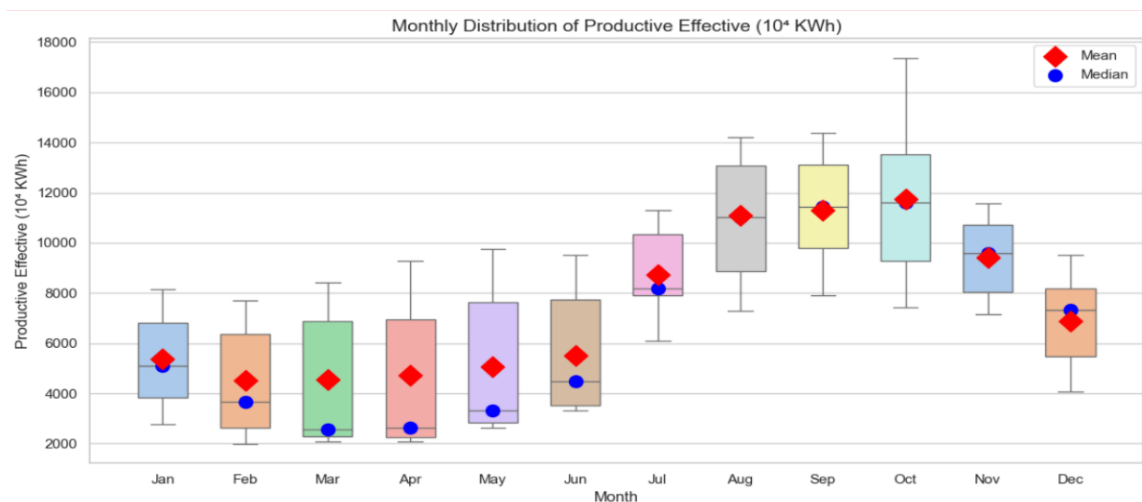


Figure 3.6: Monthly distribution of production effective

The plots show a clear seasonal variation in production, specifically in effective production/electricity production. The highest average production occurred in August and September, which corresponds with the rainy season and high-water inflow, while January and February showed the lowest production, reflecting the dry season and limited water availability.

The mean and median values are close in most months (July to January), which means the production effective is usually stable. However, in some months (March to May), the gap

between the mean and median shows that the flow is irregular. This could be due to unusual rainfall or changes in the operation of the Kaléta hydropower plant. The range of the data, indicated by the size of the boxes and lines, was bigger during the wet season months, particularly from June to October. This is probably because of changes in rainfall and water levels in the reservoir. The most minor changes were seen in the dry season, from December to March.

Figure 3.7 highlights the seasonal trends in monthly water level at Kaléta Dam.

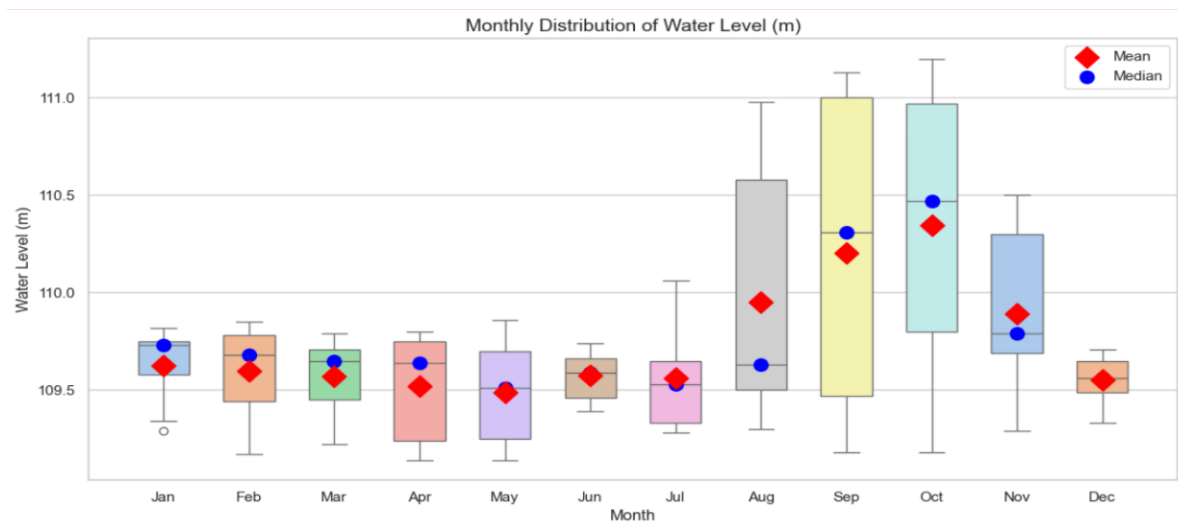


Figure 3.7: Monthly distribution of water level

The reservoir water level also follows a seasonal pattern. High water levels are observed from August to November, which is the period of maximum rainfall in the region upstream of the Konkouré watershed. Conversely, the lowest levels are found in December and June, which may be due to water management practices, just before the onset of the rainy season. The monthly box plots reveal that water levels are more stable in the dry season but show greater fluctuation in the rainy season due to heavy and sometimes unpredictable rainfall.

In some months, the range of water level values can be extensive, indicating variability driven by either climatic factors or operational changes in reservoir management, as well as for ecological reasons in the downstream area of Konkouré. For instance, maintenance work or the planned release of water for purposes may affect water level stability.

3-2.2 Annual change in production flow, production effective, and water

level

Our analysis of the annual linear trends in production flow, water level, and effective production from 2016 to 2024 reveals distinct patterns, as identified through linear regression. These trends illustrate the changes in these key variables over time.

Figure 3.8 indicates the annual changes in production flow.

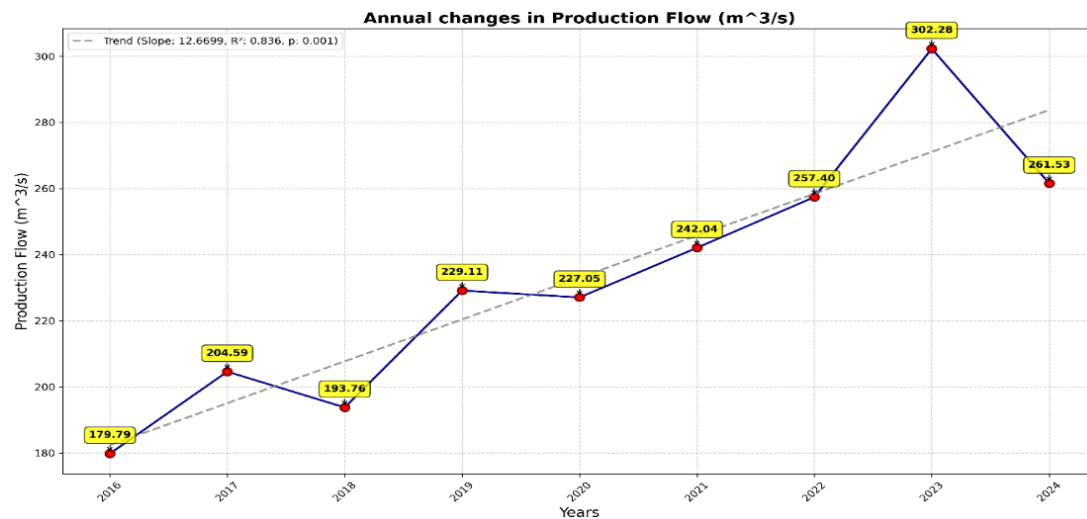


Figure 3.8: Annual changes in production flow

The production flow exhibits a notable increasing trend over the analyzed period 2016-2024. The calculated slope of 12.67 units per year signifies a consistent annual rise in production flow. This upward trajectory is highly statistically significant, with a P-value of 0.001. The R-squared value of 0.836 further indicates that approximately 83.6% of the variability in annual production flow can be explained by the linear trend, suggesting a strong and consistent increase year-over-year due to water availability and management of reservoir level in the Kaleta Dam.

Figure 3.9 displays the annual variations in production effective.

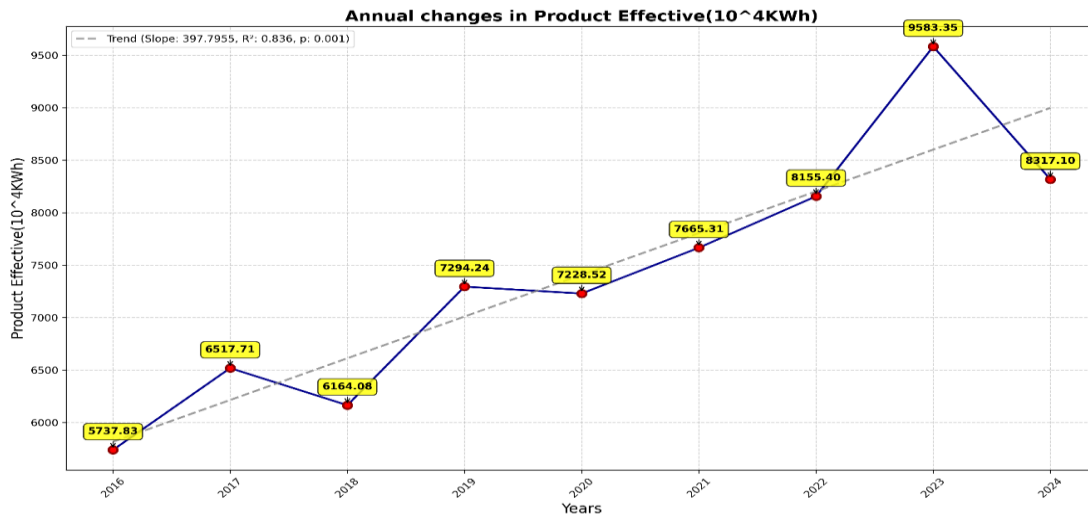


Figure 3.9: Annual change in effective production

Similar to production flow, the effective production demonstrates a significant increasing trend. With a slope of 397.80 units per year, this indicates a substantial annual growth in production effective (electricity production) in the Kaléta hydropower plant. This trend is also, highly statistically significant, mirroring the P-value of 0.001 seen in production flow. The high R-squared value of 0.836 indicates that the linear trend accounts for a substantial proportion (83.6%) of the variation in annual electricity production, suggesting a strong and consistent upward trend.

Figure 3.10 illustrates the annual fluctuations in water level at the reservoir.

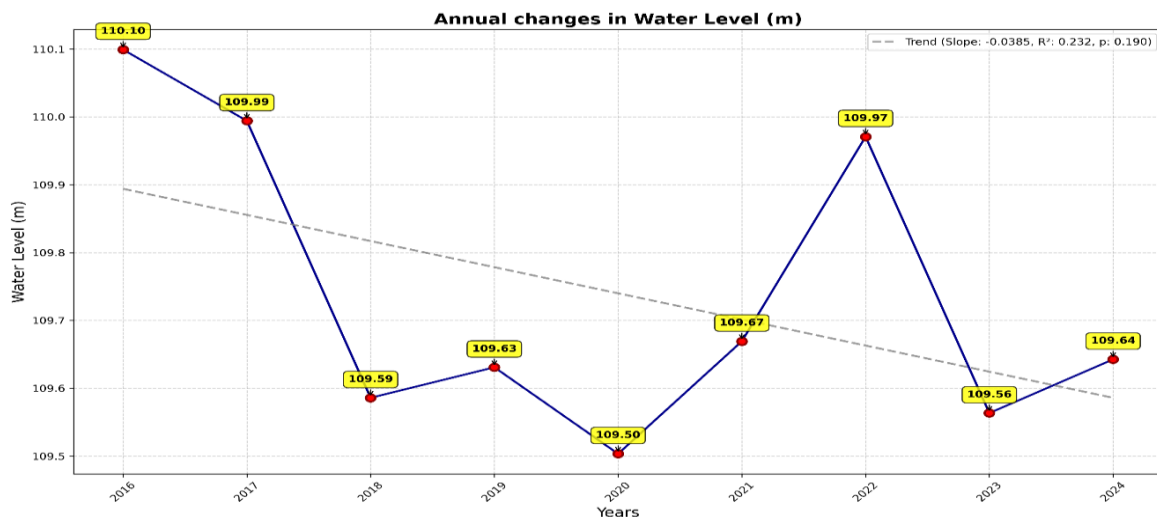


Figure 3.10: Annual changes in water level

Conversely, the water level shows a decreasing trend, with a slope of -0.0385 meters per year in the reservoir. This suggests a gradual decline in the average annual water level.

However, this observed trend is not statistically significant (P-value = 0.190). The R-squared value of 0.232 implies that the linear trend accounts for only a relatively small portion (23.2%) of the variability in water levels, indicating that other factors or natural fluctuations may be more influential in the annual changes of water level due to water management in the Kaléta reservoir.

3-3 Best climate projection for Konkouré watershed

The figures below show the Tailor diagram of ERA5 reanalysis, CHIRPS, and CMIP6 models datasets.

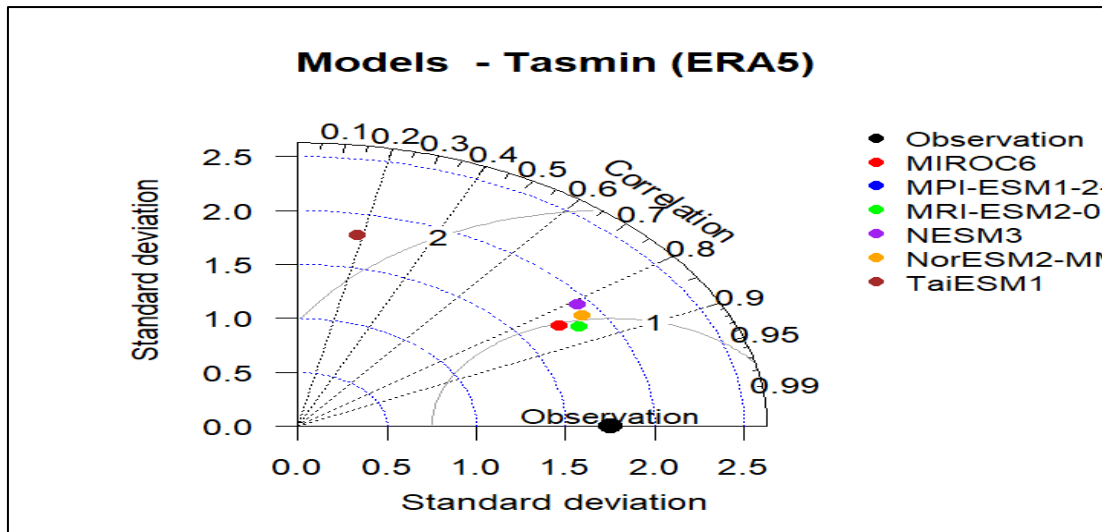


Figure 3.11: Tailor diagrams of CMIP6 comparisons with ERA5 Tasmin

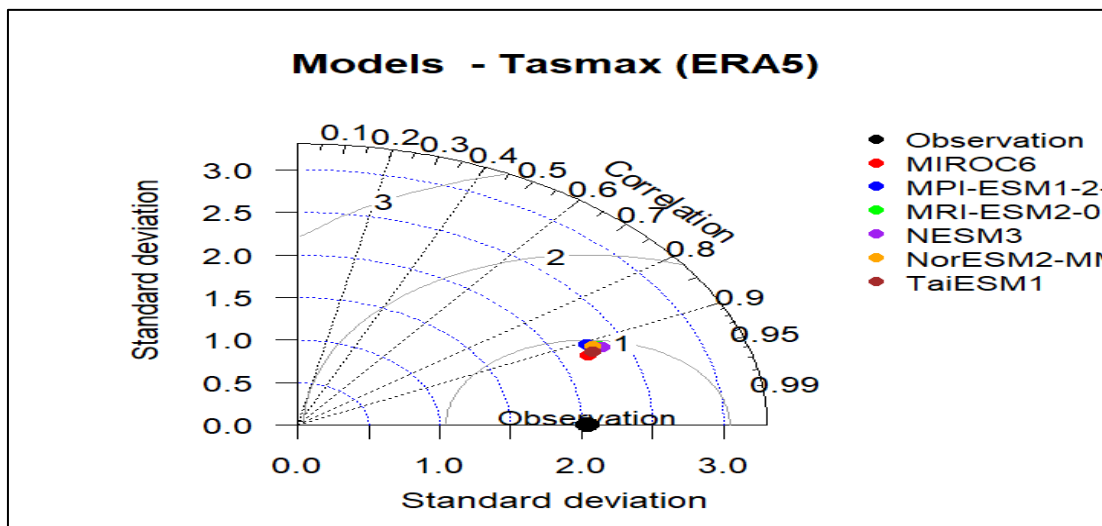


Figure 3.12: Tailor diagrams of CMIP6 comparisons with ERA5 Tasmx

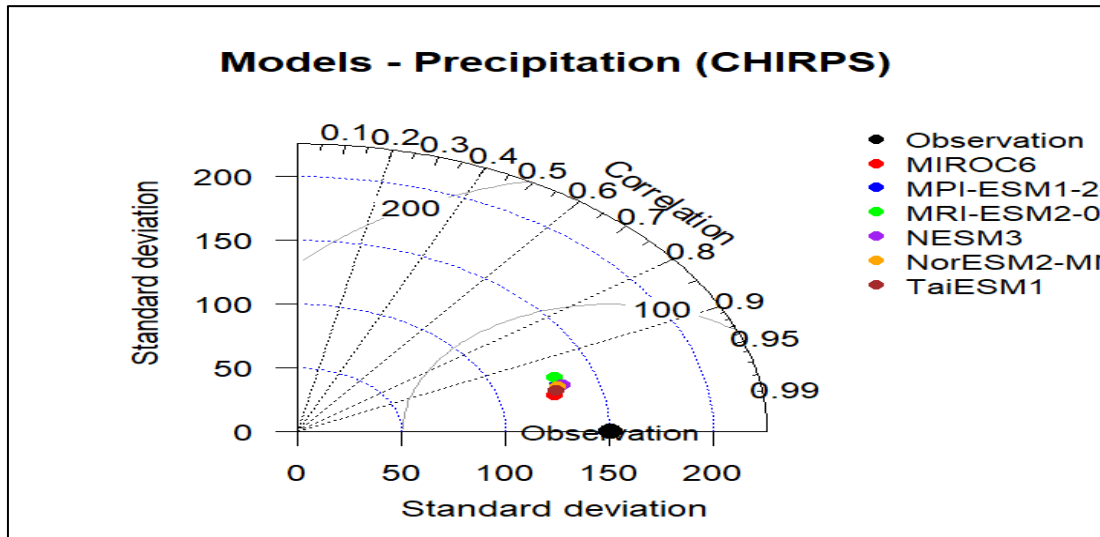


Figure 3.13: Taylor diagrams of CMIP6 comparisons with CHIRPS pr

To evaluate the performance of different climate models, we used Taylor diagrams. We employed the following climate datasets: CHIRPS for rainfall, ERA5 monthly data for maximum and minimum temperatures. These were compared with six CMIP6 climate models: MIROC6, MPI-ESM1-2-HR, MRI-ESM2-0, NESM3, NorESM2-MM, and TaiESM1.

For precipitation, it showed how closely each model's results matched the CHIRPS. The models closest to the CHIRPS point on the diagram performed better in terms of correlation and standard deviation. The plots for maximum and minimum temperature compare the models' results with ERA5. Once again, the models that appeared closer to the observation point were considered more accurate.

Models were then selected based on the Taylor diagrams, choosing those closest to the CHIRPS and ERA reanalysis data. We found that MPI-ESM1-2-HR, MRI-ESM2-0, and NorESM2-MM are the best models for future climate projection in our study area. These models were used in the stages of our research, as they are more reliable for simulating conditions. We used the CDO toolbox to run the ensemble mean to obtain a single, improved projection dataset.

Ensemble mean climate models area technique used in climate science where multiple different climate models are run, and their results are averaged to produce a single, more robust projection of future climate scenarios. This approach helps to quantify uncertainty and variability in climate predictions by considering the range of possible outcomes from

various models

3-4 Variability of Climatic Parameters

Figures 3.14 and 3.15 below show the analysis on climate projections for the Konkouré Basin from 2026 to 2034 under the SSP245 and SSP585 emission scenarios. It calculates and visualizes annual anomalies of minimum and maximum temperature, precipitation, and potential evapotranspiration, all on their respective means over the selected period.

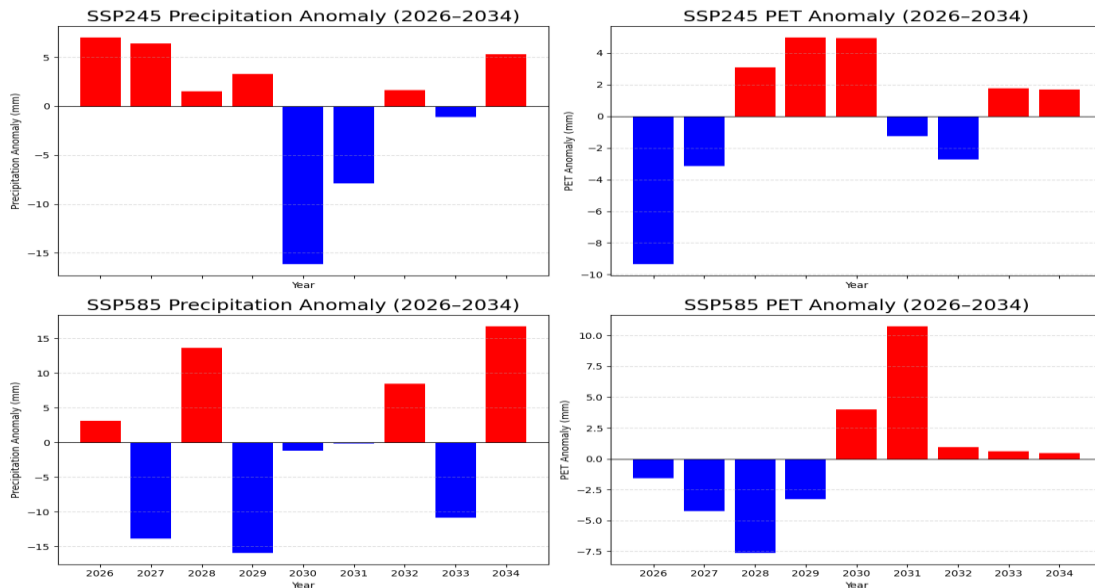


Figure 3.14: Annual Anomaly of Potential evapotranspiration in Konkouré watershed

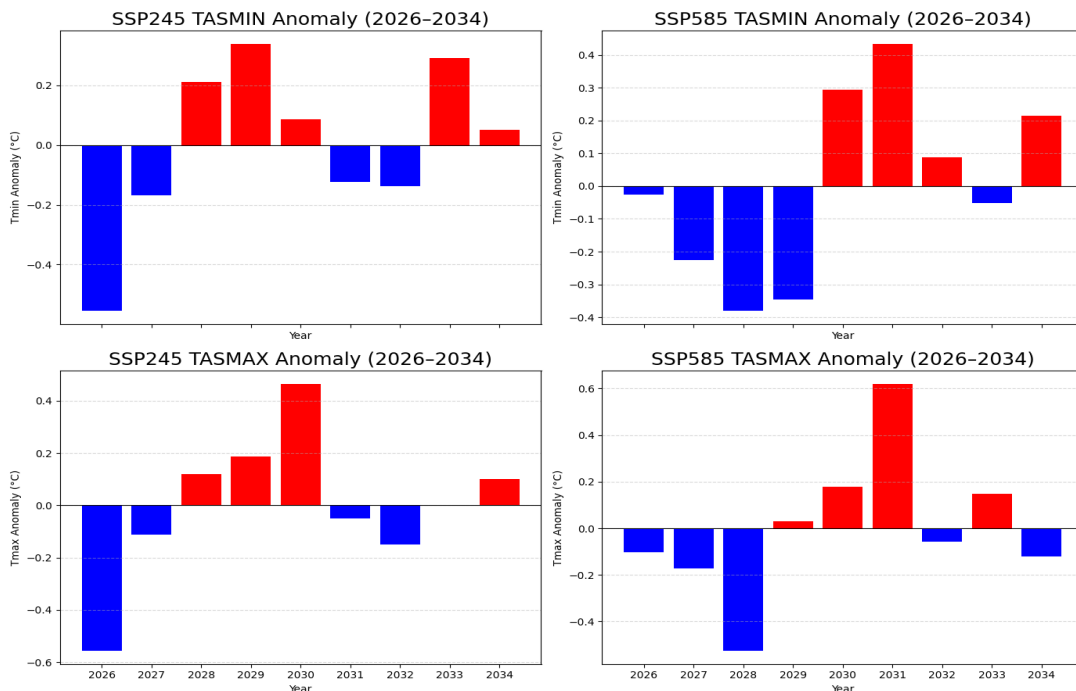


Figure 3.15: Annual Anomaly of Potential evapotranspiration bar chart in Konkouré watershed

The temperature anomaly charts show a clear warming trend in both SSP scenarios, with higher positive anomalies in the SSP5-8.5 scenario. The maximum temperature anomalies are almost greater than the temperature minimum, especially under SSP5-8.5. From 2026 to 2034, precipitation anomalies vary from year to year, but there is no clear trend of increase or decrease in either scenario. On the other hand, PET anomalies show a clear increase, especially in SSP5-8.5, which suggests that the need for water in the atmosphere is rising. These results suggest that the region will likely become warmer and drier due to higher temperatures and more evapotranspiration. This could reduce water availability for nature, farming, and hydropower, which would strongly affect the overall water supply.

3-5 SHAP explanation of the variables

Figures 3.16 and 3.17 provide a comprehensive overview of the influence of hydro-climatological variables, as interpreted through SHAP analysis, on production flow and effective energy production.

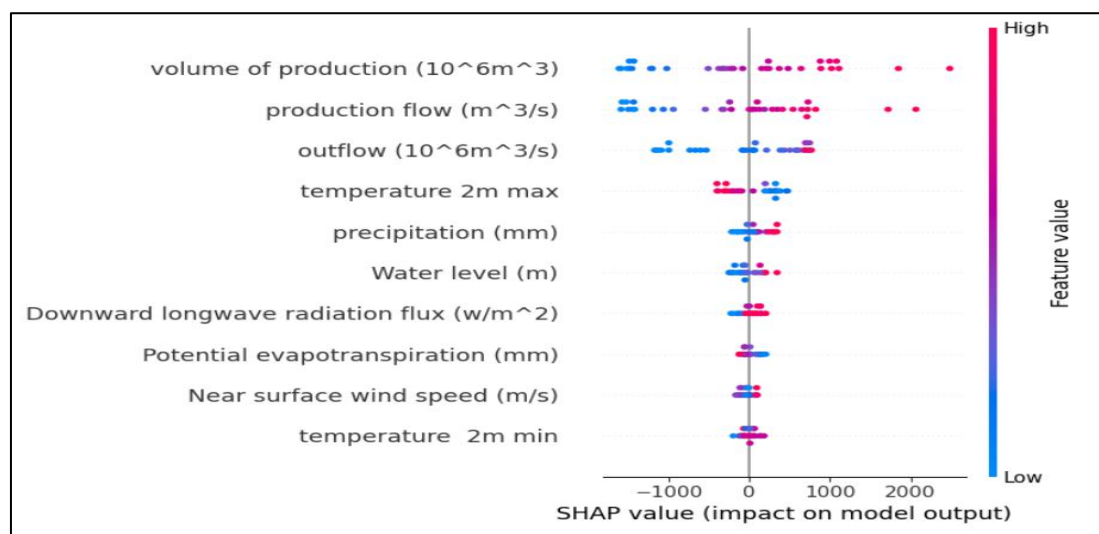


Figure 3.16: SHAP value summary of Hydro-Climatological Parameters on production effective

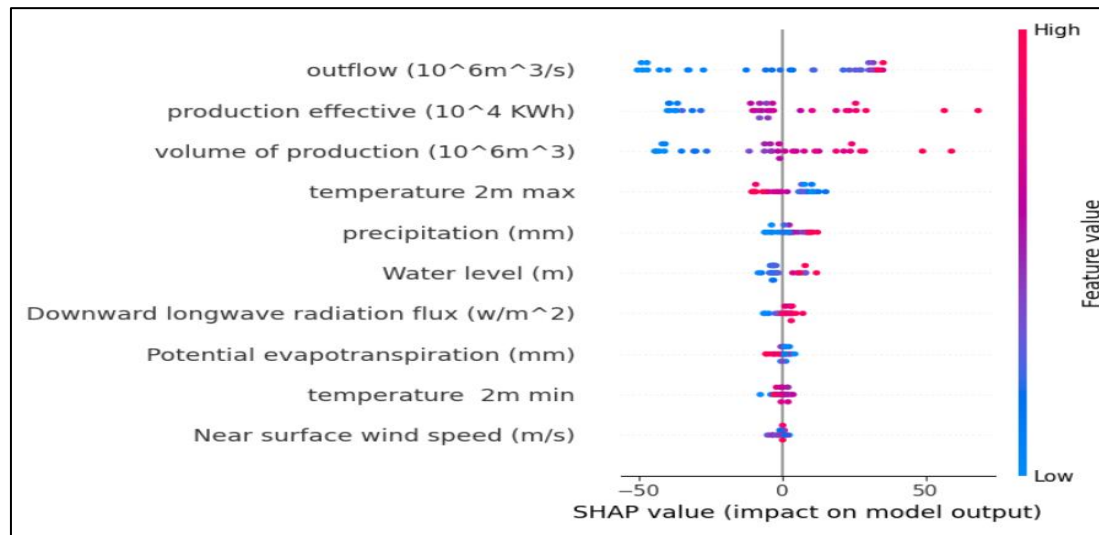


Figure 3.17: SHAP value summary of Hydro-Climatological Parameters on production flow

These analyses help us to understand how the most variable to the lowest variable have an impact on the production flow and electricity production. We used a computer program called CatBoost. Each variable learning different patterns from the historical data. It tells us that from the input data, from the input data, the volume of production is a more important variable that has the most influence on production flow and electricity production in the Kaléta Hydropower Dam. However, SHAP gave some fascinating insight into the cause-and-effect relationship between input (Downward longwave radiation flux, Relative humidity, precipitation, near-surface wind speed, temperature maximum at 2 meters, temperature minimum at 2 meters, Potential evapotranspiration, production flow, Water level, production effective, volume of production, outflow) and target variables production flow and effective production. The magnitude of feature importance is visualized in the SHAP feature importance charts given in Figures 3.14 and 3.15. It is evident that the volume of production is the most important feature in production flow and effective production, and the precipitation is the least important feature.

The model utilized climate-related variables such as downward longwave radiation, wind speed, precipitation, minimum and maximum temperature at 2 meters, and potential evapotranspiration. It demonstrated strong predictive performance, achieving an R² of 0.967 and an RMSE of 19.41 for production flow, and an R² of 0.963 with an RMSE of 661.64 for effective energy production. In comparison, Ishfaque et al. (2022) used the CatBoost model to analyze the Tarbela Dam and reported R² values of 0.978, 0.805, and 0.773 for training, validation, and testing phases, respectively. Similarly, (He et al. 2023)

applied ensemble machine learning methods, including Random Forest (ELRF) and Gradient Boosting Decision Tree (ELGBDT), for discharge prediction at the Huitangou Sluice Station in China. Their model achieved an R^2 of 0.963, RMSE of 31.27, and a correlation coefficient of 0.984. Compared to the Tarbela study, the Kaléta model showed consistently high R^2 values throughout the entire prediction period, indicating means it may have an even better ability to make accurate predictions over time. While the Tarbela study focused more on physical factors such as water level and sediment inflow, our study highlights the importance of climate-related factors, particularly temperature extremes and evapotranspiration, in influencing hydropower production. The CatBoost model used in this study not only gives similar accuracy for flow-related outputs but also provides a better understanding of how each input affects the results through SHAP values. These findings demonstrate that CatBoost is a robust and reliable method for predicting hydropower production under different climate conditions.

3-6 Hydropower prediction under shared socioeconomic pathways

3-6.1 Model performance on historical data (2016-2024)

We began our study by loading historical hydro-climatic data to train our machine learning models. Two models were used: CatBoost and XGBoost. These models were trained to predict production flow based on the historical data.

Next, we performed hyperparameter tuning to improve the performance of both models. For the CatBoost model, the best performance was achieved with a tree depth of 6, 1000 iterations, L2 regularization of 3, and a learning rate of 0.05. For the XGBoost model, the optimal settings included a tree depth of 6, 500 estimators, a learning rate of 0.05, and using 80% of the data for each tree (`colsample_bytree`) and for each boosting round (`subsample`).

After tuning, we evaluated the models' ability to predict both production flow and effective production using hydro-climatic data from the period 2016 to 2024. The dataset was split into training and testing sets, with 20% of the data reserved for testing the model's generalization ability.

To assess model performance, we used several evaluation metrics: coefficient of determination (R^2), root mean square error (RMSE), normalized RMSE (nRMSE), mean

absolute error (MAE), mean absolute percentage error (MAPE), and Pearson correlation coefficient.

3-6.1.1 Production flow

For production flow, a comparison plot shows that the predicted values are closely aligned with the observed values. The evaluation results are summarized in Table 3.1, which presents the performance metrics for both models.

Table 3.1: Scores of validations

Model	CatBoost	XGBoost
R ² (%)	0.7193	0.6586
RMSE (m ³ /s)	51.0083	56.2499
nRMSE (m ³ /s)	0.6345	0.6997
MAE (m ³ /s)	40.0971	43.5214
MAPE (m ³ /s)	16.5211	17.0887
Corr (m ³ /s)	0.8377	0.8273

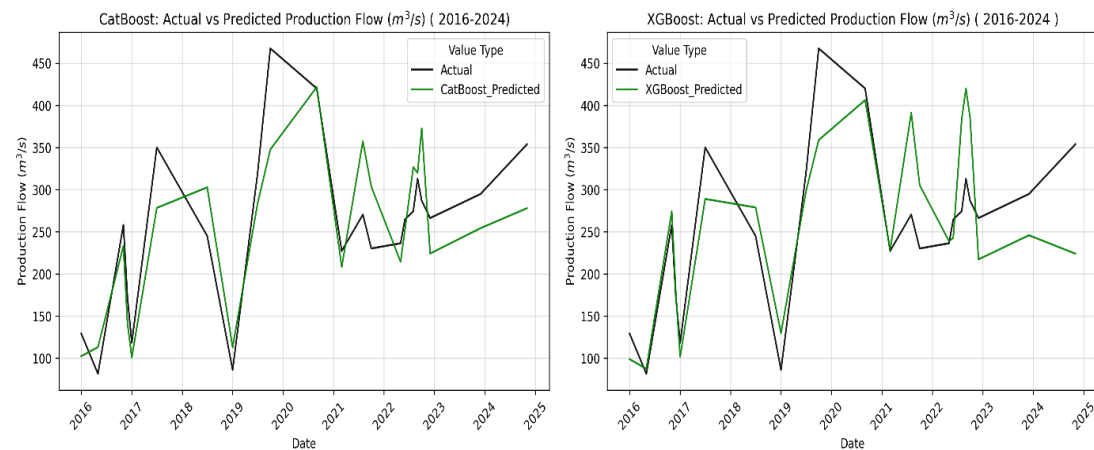


Figure 3.18: Production flow prediction

Finally, we created a plot to compare the actual values of production flow with the predicted values from our models. These metrics indicate the proportion of the variance in the dependent variable that can be predicted from the independent variables

3-6.1.2 Production effective (electricity production)

In this study, we focused on predicting effective production (electricity production) using machine learning techniques. We applied two advanced models: CatBoost and XGBoost, both known for their ability to identify complex patterns in data.

To achieve the best possible performance, we conducted a hyperparameter tuning process for each model. This involved adjusting various internal parameters such as tree depth, learning rate, number of iterations, and regularization factors to find the optimal settings. The final configurations that produced the best results are summarized under “*Best CatBoost Parameters*” and “*Best XGBoost Parameters*”, tree depth of 6, 1000 iterations, L2 regularization of 3, and a learning rate of 0.05. For the XGBoost model, the optimal settings included a tree depth of 6, 500 estimators, a learning rate of 0.05, and using 80% of the data for each tree (colsample_bytree) and for each boosting round (subsample). Once the models were tuned, we used them to predict effective electricity production. To assess the quality of the predictions, we employed several evaluation metrics. These included the coefficient of determination (R^2) and the Pearson correlation coefficient, which indicate how closely the predicted values match the observed values; the higher the coefficient, the better. Additionally, we used RMSE (root mean square error), nRMSE (normalized RMSE), MAE (mean absolute error), and MAPE (mean absolute percentage error) to evaluate the average error of the predictions, where lower values indicate higher accuracy.

The evaluation results indicate that the CatBoost model generally outperforms XGBoost across most metrics. The detailed validation results are presented in Table 3.2.

Table 3.2: Scores of validations in effective production

Model	CatBoost	XGBoost
$R^2(\%)$	0.6776	0.6515
RMSE (kWh)	1736.2761	1805.1255
nRMSE(kWh)	0.7385	0.7678
MAE (kWh)	1380.8216	1424.5757
MAPE (kWh)	17.8719	17.4957

<i>Corr</i> (kWh)	0.8377	0.8273

To further illustrate model performance, we created *actual vs. predicted* plots. These visualizations compare the real values of effective electricity production with the values predicted by each model. They provide a clear and intuitive way to assess model accuracy and to identify any deviations between observed and predicted outputs.

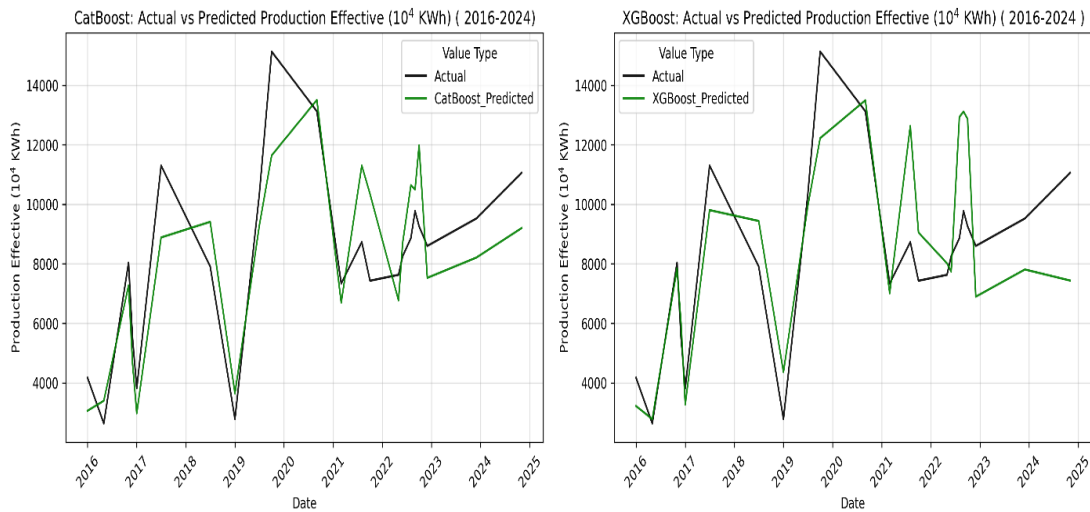


Figure 3.19: Production effective (electricity production) prediction

To assess how future climate scenarios might affect hydropower generation, we used two supervised machine learning models: CatBoost and XGBoost. These models were trained on historical hydro-climatic data from 2016 to 2024, including variables such as temperature, precipitation, and potential evapotranspiration. The target outputs for prediction were Production Flow (in cubic meters per second, m³/s) and Effective Production (in 10⁴ kilowatt-hours, kWh), which represent water discharge and energy output, respectively.

Once trained, the models were applied to future climate data corresponding to the SSP2-4.5 and SSP5-8.5 scenarios for the 2026-2034 period. These climate projections were obtained from downscaled CMIP6 outputs. By feeding these projections into the models, we generated estimates for future production flow and energy production. The machine learning models effectively learned the relationships between climate variables and

hydropower performance, allowing us to simulate future values under different climate pathways.

To measure the expected changes, we computed the mean values of production flow and effective production for both the historical (2016-2024) and future (2026-2034) periods. The percentage change was calculated. This allowed us to quantify the projected increase or decrease under each climate scenario and model combination.

3-7 The relative change in electricity generation prediction under shared socioeconomic pathways

This study assessed the potential impact of climate change on electricity generation at the Kaléta hydropower Dam using machine learning models, specifically CatBoost and XGBoost (Tables 3.1 and 3.2). Similar to the approach employed by Fukš et al (2024) in their assessment of river systems in the Outer Western Carpathians (Ropa and Biała rivers) along the Polish-Slovak border, the models in this study were applied to forecast production flow and analyze its effects on hydropower generation (Effective Production). The models were trained using historical hydroclimatic data from 2016 to 2024 and applied to predict future conditions from 2026 to 2034. These predictions were based on climate projections under two shared socioeconomic pathways (SSP2-4.5 and SSP5-8.5). The study area, the Konkouré watershed, provides the inflow to the Kaléta reservoir, making it essential for consistent power generation.

The predictive capacity of the CatBoost and XGBoost models was found to be robust. CatBoost achieved a higher R^2 of 0.7193. Wang et al. (2025) obtained $R^2 = 0.8219$ outperforms for CatBoost and $R^2 = 0.6586$ for XGBoost in simulating production flow during their observed period. Both models captured seasonal hydroclimatic variations, particularly the strong linkage between production flow, effective production, precipitation, temperature minimum, temperature maximum, potential evapotranspiration, water level, and wind speed. However, the models occasionally overestimated production flow during months with low but intense rainfall events. This behavior highlights the limitations of static input features in detecting sub-daily rainfall extremes, which can significantly contribute to hydrological responses in tropical catchments.

Production flow predictions were directly used to estimate hydropower production (effective production). Again, CatBoost outperformed XGBoost, with a slightly better predictive fit ($R^2 = 0.6776$ vs. 0.6515). Errors were within acceptable thresholds, with CatBoost producing a lower RMSE of 1736.3 MW-months. The strong Pearson correlations ($r > 0.83$) in both models confirmed a consistent relationship between model predictions and observed electricity output. This suggests that machine learning methods can serve as viable tools for hydropower planning under changing climatic conditions, especially when process-based models are limited by data availability in the increase of production flow in the Kaléta hydropower plant with 10.28% and 9.28% for the CatBoost model and 6.28% and 5.54% for the XGBoost model respectively for SSP2-4.5 and SSP5-8.5. The increase in annual streamflow was observed by (Arora, Lima and Shrestha, 2025) for the Mackenzie River (+16 %, +39 %), and Yukon rivers (+17 %, +53 %) for the RCP 4.5 and 8.5 scenarios, due to an increase of precipitation in these river basins, while they observed a decrease of (-7%) Columbia River in the RCP8.5.

The results indicate a consistent increase in effective production across both machine learning models and climate scenarios. According to the CatBoost model, production flow is projected to rise by 13.12% under SSP2-4.5 and 13.25% under SSP5-8.5. The XGBoost model also shows a positive trend, although with slightly lower increases of 8.32% under SSP2-4.5 and 7.91% under SSP5-8.5. These findings are comparable to those of Khaniya et al. (2020), who reported increases in power generation of 7.29% and 10.22% under RCP4.5 and RCP8.5, respectively, for the Samanalawewa hydropower plant in Sri Lanka using an artificial neural network.

Such trends suggest that, under both moderate and high-emission scenarios, the Konkouré river discharge is expected to increase from 2026-2034. This could be linked to intensified rainfall or shifts in seasonal patterns, as reflected in the climate model inputs.

In terms of estimated energy output (Effective Production), both models predict notable increases. CatBoost estimates a rise of 9.75 GWh under SSP2-4.5 and 9.84 GWh under SSP5-8.5 for the Kaléta hydropower plant during the 2026-2038 period. XGBoost forecasts a smaller but positive growth of 6.21 GWh under SSP2-4.5 and 5.90 GWh under SSP5-8.5. The similar projections between SSP2-4.5 and SSP5-8.5 suggest that, in the short term, electricity production may not be susceptible to differences between moderate and high-emission pathways.

However, these findings contrast with the results of Guo et al. (2021), who projected a decline in hydropower generation at the Manji Dam due to rising temperatures. Their study estimates that hydropower production could decrease by 10.98 MW, 12.93 MW, and 14.77 MW under RCP2.6, RCP4.5, and RCP8.5, respectively, between 2021 and 2040. This contrast highlights the importance of local climate, catchment characteristics, and model inputs in shaping hydropower projections. Yalcin et al. (2023) find the decrease of approximately 10.1% and 21.5% in annual hydropower production at Dipni Dam. The hydropower will decrease in the 50 years approximately 3.36, 4.62, and 6.64 MW for the RCP2.6, RCP4.5, and RCP8.5 over the Aras River flows. Additionally, Da Silva et al. (2021) found that under SSP2-4.5 and SSP5-8.5, the hydropower potential may decrease by 30 to 50 % compared to the baseline period 1901-2000 in the São Francisco River Basin, Brazil.

Overall, the projections indicate a positive impact of climate change on hydropower potential short term, assuming current climate model projections remain accurate. However, differences between the models (CatBoost and XGBoost) emphasize the significant importance of model selection and potential uncertainties in data-driven predictions. These outcomes can guide water resource managers and energy planners in preparing for an increase in hydropower generation capacity. While also highlighting the necessity to assess other risks, such as interannual variability or extremes, that are not reflected in average trends.

3-8 Maximum monthly changes comparison of the historical period and the prediction

Figure 3.20 presents the maximum monthly changes in production flow.

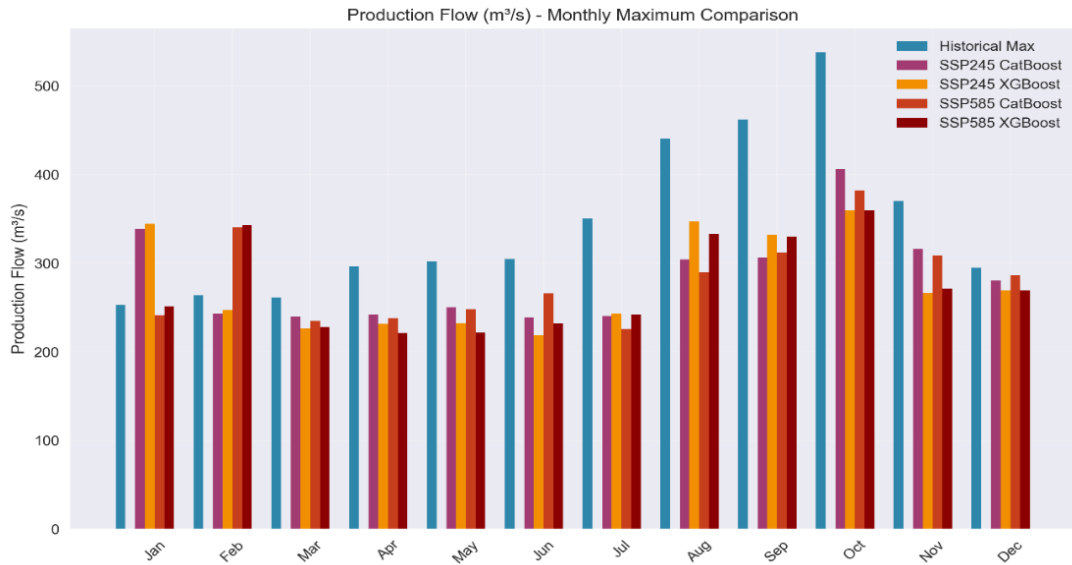


Figure 3.20: Monthly changes in the maximum production flow

The figure presents a bar chart illustrating the maximum monthly production flow. It exhibits distinct seasonal trends, with peak values generally occurring in the rainy months. In comparison to historical data, the scenarios for both SSP2-4.5 and SSP5-8.5 display minor differences, while the estimations from the XGBoost and CatBoost models vary, emphasizing the significance of algorithm selection for future predictions.

Figure 3.21 indicates the highest monthly variations in effective production.

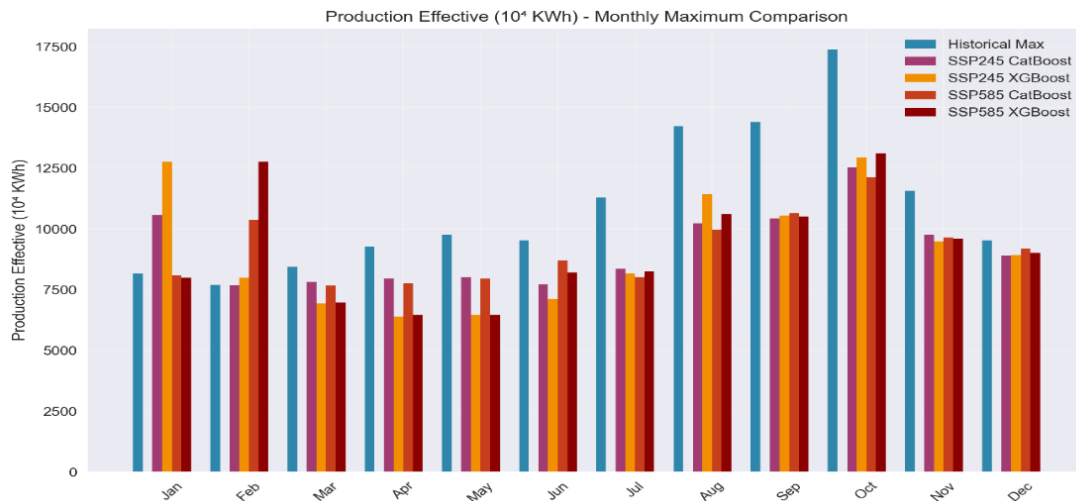


Figure 3.21: Monthly changes in the maximum effective production

The second subplot illustrates the production effectiveness (10⁴ kWh) and reveals seasonal patterns. In both SSP scenarios, future values often exceed the historical maximum, particularly during the wet season. This suggests that electricity production may increase

under future climate conditions. However, the results heavily depend on the model employed, as CatBoost and XGBoost demonstrate different levels of increase.

As highlighted in Figure 3.22 below, the production flow shows monthly fluctuations in all months.

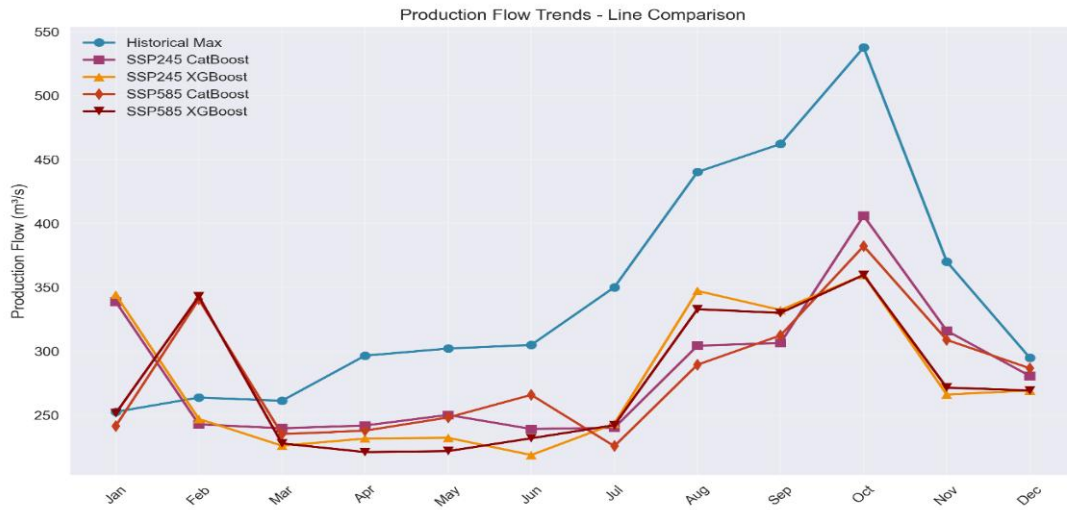


Figure 3.22: Monthly maximum production flow trends changes

Figure 3.23 displays the changes in monthly maximum production flow trends.

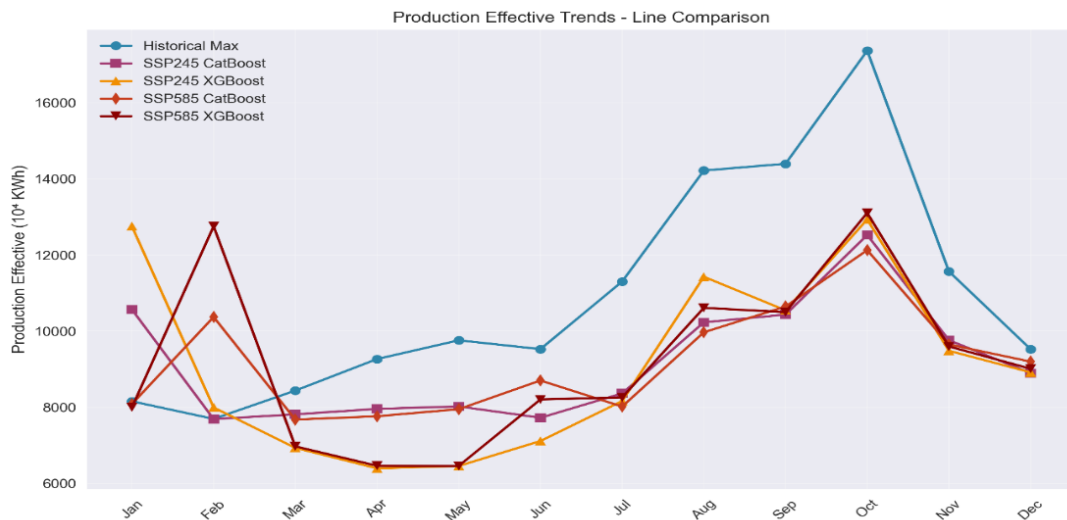


Figure 3.23: Monthly maximum production flow trends changes

Figures 3.22 and 3.23 provide a clearer picture of monthly trends. The line for historical data shows consistent peaks during the wet months (typically June to September), which correspond to periods of higher rainfall and runoff. The SSP scenarios, while maintaining the general trend, show either intensified or moderated production flow and production

effective, implying future shifts in the hydropower regime. This reinforces the seasonal vulnerability and opportunity zones for hydropower planning.

Figures 3.20 to 3.23 illustrate that climate change could have a significant impact on hydropower generation at the Kaléta Dam. Under the SSP5-8.5 scenario, the predicted maximum values are decreasing from the historical values. This means that there could be increased hydropower production in the future, but only if water management systems are improved and adapted. Meanwhile, the maximum values should decrease under SSP2-4.5 and SSP5-8.5 2026-2034 compared to the maximum values of the historical period 2016-2024.

However, these increases are not certain. The results depend on the machine learning model used. For example, CatBoost and XGBoost give different predictions. This difference makes it hard to say precisely what will happen. It indicates that future outcomes remain uncertain, despite the generally positive trend. While monthly runoff is projected to rise in June and August (2021-2060) and increase in all months 2061-2100 by Chen et al. (2022).

The seasonal pattern of production remains largely unchanged, with higher values during the wet season. However, the size of these peaks in the future could be larger, which may pose a problem for the current infrastructure. If not managed well, this could cause stress on the system. Also, in the dry season, the models show a decrease in production on a monthly observation, which reaches its peak in October with almost 540 m³/s to decrease 385 m³/s, 380 m³/s, for model CatBoost in the scenarios SSP2-4.5 and SSP5-8.5, and nearly 370 m³/s for all the 2 scenarios SSP2-4.5 and SSP5-8.5 for XGBoost model. This could make the energy supply less stable in those months. Chen et al. (2022) outperformed an increase of monthly runoff in 4 months (April, May, June, July, and October) and a decrease in 6 months (January, February, August, September, and December).

Finally, the differences between CatBoost and XGBoost models demonstrate that relying solely on one model is risky. It is better to use several models and compare their results. Additionally, verifying the models against real-world data from the ground can help improve accuracy. These findings suggest that future hydropower plans should use different models and scenarios to prepare for uncertain climate conditions.

3-9 Anomalies in production flow and production effective

Figure 3.24 below reveals the monthly variations in production flow.

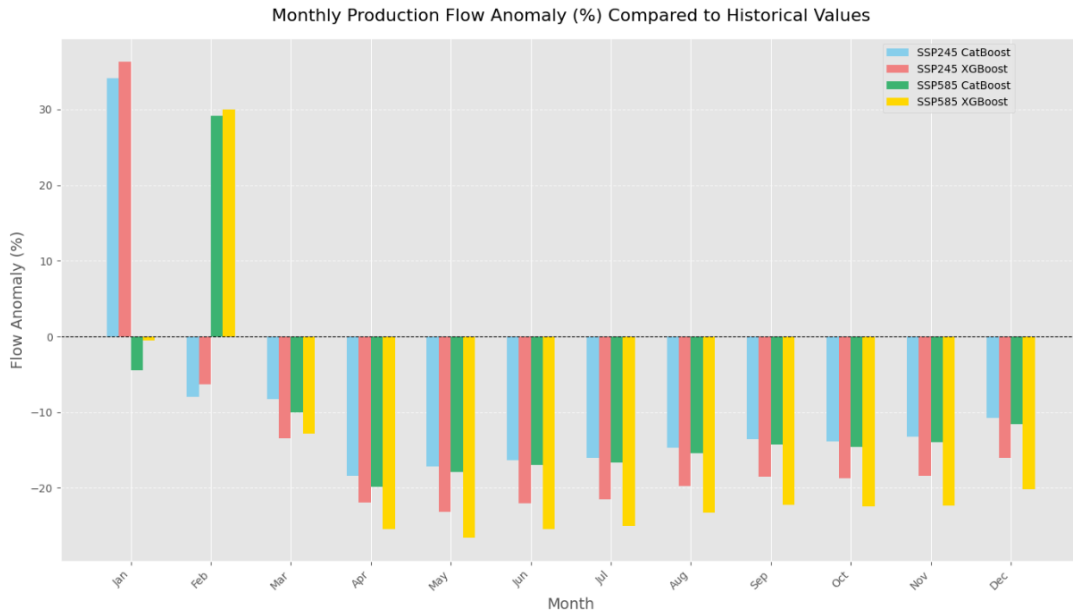


Figure 3.24: Monthly anomaly production flow

Figure 3.25 provides a clear overview of the monthly extremes and anomalies in effective production, as shown below.

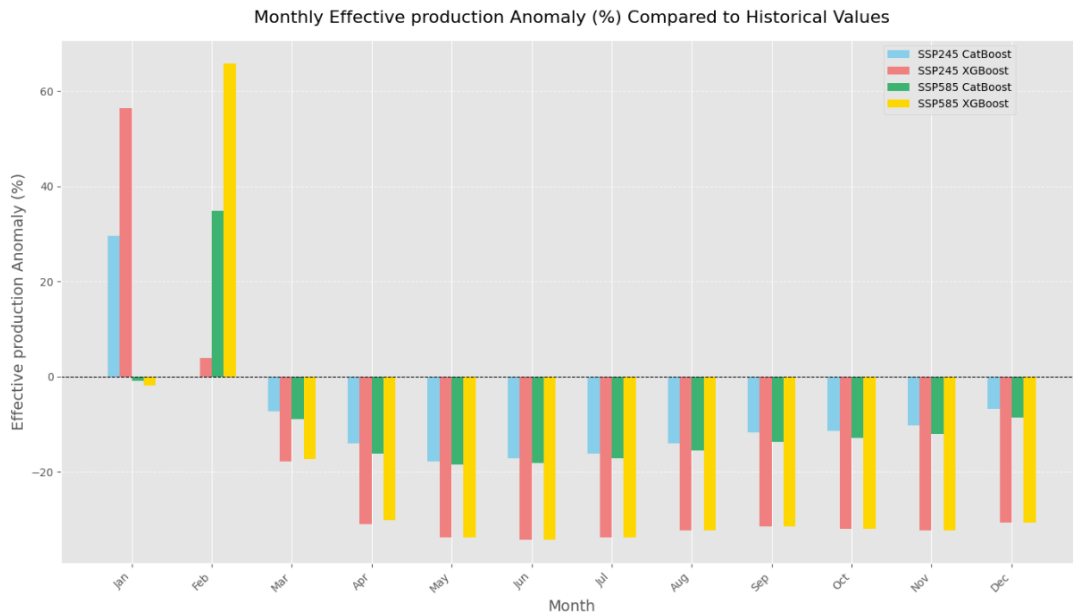


Figure 3.25: Monthly production effective anomaly

The results are presented using two key visualizations: one chart showing monthly streamflow anomalies and another depicting monthly energy effectiveness anomalies, each

comparing projections under SSP2-4.5 and SSP5-8.5 scenarios to historical data. Each chart includes four color-coded bars per month corresponding to predictions from CatBoost and XGBoost under each SSP. This format enables a clear and direct comparison across models and scenarios.

In both figures, months are displayed on the x-axis and anomaly percentages are displayed on the y-axis. A dashed horizontal line at 0% represents the historical baseline. Bars extending above the line indicate months where future values exceed historical averages (positive anomaly), while those below suggest declines (negative anomaly). The use of distinct colors and grouped bars facilitates model-to-model and scenario-to-scenario comparison for each month.

Figures 25 and 26. Monthly streamflow anomalies under SSP2-4.5 and SSP5-8.5 scenarios compared to historical flow (2016-2024). Positive values indicate projected increases in flow; negative values represent reductions.

The anomaly plots provide important insights into how climate change may affect hydropower performance at the Kaléta hydropower plant. For production flow, the SSP2-4.5 projections generally indicate minor positive values for January and March, while negative deviations from historical production flow for other months in the years in the prediction period 2026-2034, depending on the month and model, XGBoost predicts the most negative anomalies in the two scenarios than the CatBoost model. In contrast, the SSP5-8.5 scenario exhibits more consistent decreases in production flow, particularly in the later months, indicating potential future water shortages that could limit hydropower generation capacity (Ekwezu et al. 2018).

Regarding effective production, we observe that the CatBoost model under the SSP2-4.5 project shows moderate production increases, effective in the early months (e.g., January-March), which could be linked to higher water availability or improved operational efficiency. However, under the SSP5-8.5 scenario, XGBoost shows notable decreases across many months, indicating that higher emissions pathways might reduce electricity production reliability. These findings suggest greater vulnerability under SSP5-8.5 compared to SSP2-4.5.

Overall, the comparative analysis between historical and projected values helps identify critical periods where climate change could stress hydropower systems. The use of both

CatBoost and XGBoost also highlights the model uncertainty, reinforcing the importance of using multiple approaches in climate impact assessments. The seasonal breakdown enables policymakers and engineers to develop more effective adaptive strategies tailored to monthly operational needs.

Conclusion

this chapter presents the results and discussion of the research from the beginning to the end of the results and their discussion, such as climate variability, variability of hydroelectricity from 2016 to 2024, the choice of the best models of CMIP6 in the Konkouré study area, and the hydropower prediction under two shared socio-economic pathways, SSP2-4.5 and SSP5-8.5 from 2026 to 2034 using two machine learning models CatBoost and XGBoost.

CONCLUSION AND PERSPECTIVE

CONCLUSION

In this study, we examined how climate change may impact electricity production at the Kaléta hydropower Dam. We used two advanced machine learning models, CatBoost and XGBoost. These models were trained using historical data from 2016 to 2024 and were then used to predict production flow and electricity production for the years 2026 to 2034. We used two climate scenarios: SSP2-4.5 and SSP5-8.5. This method supports long-term planning for hydropower by utilizing modern data analysis tools.

To check how well the models worked, we used essential performance measures. The CatBoost model yielded a strong R^2 value of 0.7193, an RMSE of 51.01, and a Pearson correlation of 0.8564, indicating good accuracy. XGBoost also gave good results, with an R^2 of 0.6586 and a correlation of 0.8341. These results indicate that both models can accurately predict production flow, with CatBoost performing slightly better.

For electricity production, measured in units of 10^4 kWh, CatBoost again gave better results than XGBoost. It had a higher R^2 value (0.6776 vs. 0.6515) and a lower RMSE (1736.28 vs. 1805.13). Other error values, such as MAE and MAPE, also indicated that both models followed the general trend in energy production well, even though some small errors occurred during months with unusual water flows. These results demonstrate that machine learning is a valuable tool for predicting the impact of climate change on the production flow of hydropower.

Our analysis of past data showed that the average production flow from 2016 to 2024 was about 233 m^3/s . For the years 2026 to 2034, the models predicted an increase in the number of cases. For example, CatBoost predicted a 10.28% increase under SSP2-4.5 and a 9.92% increase under SSP5-8.5. XGBoost showed smaller increases (6.29% and 5.54%). This suggests that future rainfall may increase the amount of water entering the Kaléta reservoir.

In terms of energy production, CatBoost predicted an average of 8404.22×10^4 KWh under SSP2-4.5 and 8413.91×10^4 KWh under SSP5-8.5. These values are about 13% higher than the historical average. XGBoost predicted smaller increases of about 8.32% and 7.91%, but

still showed an upward trend. This means Kaléta Dam could have better water availability and higher power output in the near future.

These results fit with basic hydrological knowledge. When more water flows through a river, more energy can be produced because more water passes through the turbines. However, our study focused solely on the production flow and its relative change in electricity production. Other important factors, such as the efficiency of the turbines, the amount of sediment that accumulates in the reservoir, and the dam's management, were not considered and could also impact the actual amount of electricity generated.

Even with this simple approach, the clear rise in expected electricity production indicates that Kaléta Dam could continue to function effectively under climate change, at least in the short term. This is important because some studies in West Africa predict lower hydropower output due to reduced rainfall. The Konkouré River, where Kaléta is located, might be less affected by this, or current models might not fully capture local rainfall patterns.

Still, some risks remain. Strong storms and extreme rainfall, which the models may not predict well, can cause problems such as flooding or damage to turbines. Although the predictions for 2026-2034 are positive, the long-term future is less certain. Some climate models suggest that West Africa may experience more pronounced year-to-year rainfall fluctuations, along with shifts in rainy seasons and changes in storm behavior. These changes could affect both the water supply and the electricity demand. Using many climate models (CMIP6) in machine learning can help improve the reliability of long-term predictions.

Verification of the hypothesis of this research:

The initial hypotheses suggested that increasing temperature and evaporation due to climate change would reduce the potential electricity production at the Kaléta Hydropower Dam. However, the analysis of observed and projected data does not fully support these assumptions. Despite climate change impacts such as higher temperatures and evaporation, the Kaléta Dam may experience improved water availability and greater electricity generation. Therefore, the original hypothesis that climate change would decrease hydropower production was not confirmed by the current data and model projections.

PERSPECTIVE

To ensure a sustainable electricity supply for the growing population in Conakry and other regions served by the Kaléta Hydropower Plant, future research should assess the alignment between projected electricity demand and the plant's generation potential under future climate scenarios. Such an evaluation will support stakeholders in strategic planning and promote the long-term sustainability of energy production.

In addition, future studies should consider other critical factors that influence the performance of the Kaléta Hydropower Plant and other hydropower facilities in the Konkouré River Basin. One such factor is sedimentation in the reservoir, which can reduce storage capacity, limit water availability, and negatively impact energy production.

We also need to consider the impact of climate change on hydropower dams before they are constructed. This is a crucial point for stakeholders when planning hydropower projects and managing the dam's reservoir.

REFERENCES

- Agnoun, Yves, Samadori S. H. Biaou, M. Sié, R. S. Vodouhè, and A. Ahanchédé. 2012. “The African Rice *Oryza Glaberrima* Steud: Knowledge Distribution and Prospects.” *International Journal of Biology* 4(3). doi:10.5539/ijb.v4n3p158.
- Ansoumane Sakouvogui, Elhadj Ousmane Camara, Nènè Aïssata Balde, and Mamby Keita. 2023. “Sizing and Simulation of a Hybrid Hydroelectricity and Photovoltaic System with Storage for Supplying the Tamagaly District in Mamou, Guinea.” *Journal of Energy and Power Engineering* 17(3). doi:10.17265/1934-8975/2023.03.001.
- Arora, Vivek K., Aranildo Lima, and Rajesh Shrestha. 2025a. “The Effect of Climate Change on the Simulated Streamflow of Six Canadian Rivers Based on the CanRCM4 Regional Climate Model.” *Hydrology and Earth System Sciences* 29(1): 291–312. doi:10.5194/hess-29-291-2025.
- Arvizu, Dan, Thomas Bruckner, John Christensen, Jean-Michel Devernay, Andre Faaij, Manfred Fischedick, Barry Goldstein, et al. *Final Draft Special Report Renewable Energy Sources (SRREN) Summary for Policy Makers*.
- Bah, A., Zhang, H., Luo, Z. et al. 2024. “A Study of Land Use Changes and Its Impacts on Flood Inundation in the Konkouré River Basin, Republic of Guinea.” *Environ Monit Assess* 196.
- Barasa, Maulidi, Dmitrii Bogdanov, Ayobami Solomon Oyewo, and Christian Breyer. 2018. “A Cost Optimal Resolution for Sub-Saharan Africa Powered by 100% Renewables in 2030.” *Renewable and Sustainable Energy Reviews* 92: 440–57. doi:10.1016/J.RSER.2018.04.110.
- Bishop CM. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag, New York Inc; 2006.
- Chen, Changzheng, Rong Gan, Dongmei Feng, Feng Yang, and Qiting Zuo. 2022. “Quantifying the Contribution of SWAT Modeling and CMIP6 Inputting to Streamflow Prediction Uncertainty under Climate Change.” *Journal of Cleaner Production* 364: 132675. doi:10.1016/J.JCLEPRO.2022.132675.
- Chen, Tianqi, and Carlos Guestrin. 2016. “XGBoost: A Scalable Tree Boosting System.” In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, 785–94.

doi:10.1145/2939672.2939785.

country priority plan and diagnostic of the electricity sector Guinea 3. 2021.

Dagne, Yasmin. 2020. “*We’re Leaving Everything behind*”: *The Impact of Guinea’s Souapiti Dam on Displaced Communities*. Human Rights Watch.

discover pia capabilities portfolio insights news careers contacts kaléta and souapiti hydropower schemes, a success story in public private partnerships.

Ekwezu, C. S., H. C. Nnamchi, and P. O. Phil-Eze. 2018. “Future Changes in Seasonal-Mean Precipitation over West Africa under the Representative Concentration Pathways.” *Nigerian Journal of Technology* 37(1): 233. doi:10.4314/njt.v37i1.31.

F. Borowski, Piotr. 2020. “Nexus between Water, Energy, Food and Climate Change as Challenges Facing the Modern Global, European and Polish Economy.” *AIMS Geosciences* 6(4): 397–421. doi:10.3934/geosci.2020022.

Fan, Jing Li, Jia Wei Hu, Xian Zhang, Ling Si Kong, Fengyu Li, and Zhifu Mi. 2020. “Impacts of Climate Change on Hydropower Generation in China.” *Mathematics and Computers in Simulation* 167: 4–18. doi:10.1016/J.MATCOM.2018.01.002.

Friedman, Jerome. 2001. “Greedy Function Approximation: A Gradient Boosting Machine.” *Annals of Statistics* 29(5): 1189–1232.

Fukś, Maksymilian, Mariola Kędra, and Łukasz Wiejaczka. 2024. “Assessing the Impact of Climate Change and Reservoir Operation on the Thermal and Ice Regime of Mountain Rivers Using the XGBoost Model and Wavelet Analysis.” *Stochastic Environmental Research and Risk Assessment*. doi:10.1007/s00477-024-02803-2.

Goodarzi, Mohammad Reza, Hamed Vagheei, and Rabi H. Mohtar. 2020. “The Impact of Climate Change on Water and Energy Security.” *Water Science and Technology: Water Supply* 20(7): 2530–46. doi:10.2166/ws.2020.150.

Goodfellow, Ian. 2016. *Deep Feedforward Networks*. www.deeplearningbook.org.

Guo, Li Na, Chen She, De Bin Kong, Shuai Ling Yan, Yi Peng Xu, Majid Khayatnezhad, and Fatemeh Gholinia. 2021a. “Prediction of the Effects of Climate Change on Hydroelectric Generation, Electricity Demand, and Emissions of Greenhouse Gases under Climatic Scenarios and Optimized ANN Model.” *Energy Reports* 7: 5431–45. doi:10.1016/j.egyr.2021.08.134.

Harina Azman, Azreen, Nurul Nadrah, Aqilah Tukimat, and M A Malek. 2021. 16 *Journal of Engineering Science and Technology climate change impacts on hydropower generation at kenyir lake*.

- He, Sheng, Geng Niu, Xuefeng Sang, Xiaozhong Sun, Junxian Yin, and Heting Chen. 2023. "Machine Learning Framework with Feature Importance Interpretation for Discharge Estimation: A Case Study in Huitanggou Sluice Hydrological Station, China." *Water (Switzerland)* 15(10). doi:10.3390/w15101923.
- "IPCC_WGI-AR6-Press-Release_en."
- Ishfaqe, Muhammad, Saad Salman, Khan Zaib Jadoon, Abid Ali Khan Danish, Kifayat Ullah Bangash, and Dai Qianwei. 2022. "Understanding the Effect of Hydro-Climatological Parameters on Dam Seepage Using Shapley Additive Explanation (SHAP): A Case Study of Earth-Fill Tarbela Dam, Pakistan." *Water (Switzerland)* 14(17). doi:10.3390/w14172598.
- Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. *LightGBM: A Highly Efficient Gradient Boosting Decision Tree*. <https://github.com/Microsoft/LightGBM>.
- Khaniya, Bhabishya, Chamaka Karunanayake, Miyuru B. Gunathilake, and Upaka Rathnayake. 2020. "Projection of Future Hydropower Generation in Samanalawewa Power Plant, Sri Lanka." *Mathematical Problems in Engineering* 2020. doi:10.1155/2020/8862067.
- Kim, Su Jin, Masoud Asadzadeh, and Tricia A. Stadnyk. 2022. "Climate Change Impact on Water Supply and Hydropower Generation Potential in Northern Manitoba." *Journal of Hydrology: Regional Studies* 41: 101077. doi:10.1016/J.EJRH.2022.101077.
- Kim, Yeon Hee, Seung Ki Min, Xuebin Zhang, Jana Sillmann, and Marit Sandstad. 2020. "Evaluation of the CMIP6 Multi-Model Ensemble for Climate Extreme Indices." *Weather and Climate Extremes* 29: 100269. doi:10.1016/J.WACE.2020.100269.
- Kling, Harald, Philipp Stanzel, and Martin Fuchs. 2016. "Regional Assessment of the Hydropower Potential of Rivers in West Africa." *Energy Procedia* 97: 286–93. doi:10.1016/J.EGYPRO.2016.10.002.
- Kling, Harald, Philipp Stanzel, Martin Fuchs, and Hannes Bauer. 2017. 6 Conference Proceedings Paper *water storage and hydropower development for africa International Conference and Exhibition Theoretical Hydropower Potential for Different Plant Sizes for All Rivers in West Africa*. www.ecowrex.org.
- Kouadio, Koffi Claude Alain, Ernest Amoussou, Talnan Jean Honoré Coulibaly, Arona Diedhiou, Houebagnon Saint Jean Patrick Coulibaly, Regis Didi, and Issiaka Savané.

2020. “Analysis of Hydrological Dynamics and Hydropower Generation in a West African Anthropized Watershed in a Context of Climate Change.” *Modeling Earth Systems and Environment* 6(4): 2197–2214. doi:10.1007/s40808-020-00836-4.
- Kumar, Vijendra, Naresh Kedam, Kul Vaibhav Sharma, Darshan J. Mehta, and Tommaso Caloiro. 2023. “Advanced Machine Learning Techniques to Improve Hydrological Prediction: A Comparative Analysis of Streamflow Prediction Models.” *Water (Switzerland)* 15(14). doi:10.3390/w15142572.
- Lee, Jyun Long, and Wen Cheng Huang. 2014. “Impact of Climate Change on the Irrigation Water Requirement in Northern Taiwan.” *Water (Switzerland)* 6(11): 3339–61. doi:10.3390/w6113339.
- Maciejewski, Dawid, Krzysztof Mudryk, and Maciej Sporysz. 2024. “Forecasting Electricity Production in a Small Hydropower Plant (SHP) Using Artificial Intelligence (AI).” *Energies* 17(24). doi:10.3390/en17246401.
- Mailhot, Edouard, Biljana Music, Daniel F. Nadeau, Anne Frigon, and Richard Turcotte. 2019. “Assessment of the Laurentian Great Lakes’ Hydrological Conditions in a Changing Climate.” *Climatic Change* 157(2): 243–59. doi:10.1007/s10584-019-02530-6.
- Mohammed, Safwan, Sana Arshad, Firas Alsilibe, Muhammad Farhan Ul Moazzam, Bashar Bashir, Foyez Ahmed Prodhan, Abdullah Alsalman, et al. 2024. “Utilizing Machine Learning and CMIP6 Projections for Short-Term Agricultural Drought Monitoring in Central Europe (1900–2100).” *Journal of Hydrology* 633: 130968. doi:10.1016/J.JHYDROL.2024.130968.
- Nowak, Bogumił, Anna Andrzejak, Grzegorz Filipiak, Mariusz Ptak, and Mariusz Sojka. 2022. “Assessment of the Impact of Flow Changes and Water Management Rules in the Dam Reservoir on Energy Generation at the Jeziorsko Hydropower Plant.” *Energies* 15(20). doi:10.3390/en15207695.
- Obahoundje, Salomon, and Arona Diedhiou. 2022. “Potential Impacts of Climate, Land Use and Land Cover Changes on Hydropower Generation in West Africa: A Review.” *Environmental Research Letters* 17(4). doi:10.1088/1748-9326/ac5b3b.
- Obahoundje, Salomon, Arona Diedhiou, Laurent Dubus, Eric Adéchina Alamou, Ernest Amoussou, Komlavi Akpoti, and Eric Antwi Ofori. 2022. “Modeling Climate Change Impact on Inflow and Hydropower Generation of Nangbeto Dam in West Africa Using Multi-Model CORDEX Ensemble and Ensemble Machine Learning.” *Applied Energy*

325: 119795. doi:10.1016/J.APENERGY.2022.119795.

- Olabanji, Mary Funke, Thando Ndarana, Nerhene Davis, and Emma Archer. 2020. "Climate Change Impact on Water Availability in the Olifants Catchment (South Africa) with Potential Adaptation Strategies." *Physics and Chemistry of the Earth, Parts A/B/C* 120: 102939. doi:10.1016/J.PCE.2020.102939.
- Oti, Jonathan Opoku, Amos T. Kabo-bah, and Eric Ofosu. 2020. "Hydrologic Response to Climate Change in the Densu River Basin in Ghana." *Heliyon* 6(8). doi:10.1016/j.heliyon.2020.e04722.
- Sakouvogui, Ansoumane, Wogbo Dominique Guilavogui, Mamadou Lamarana Diallo, Thierno Amadou Barry, and Mamby Keita. 2023. "Revue RAMReS-Sciences Appliquées et de l'Ingénieur Influences Des Changements Climatiques Sur Le Potentiel Énergétique Du Barrage Hydroélectrique de Garafiri (République de Guinée). Résumé." *RAMReS-Sci. Appl. & de l'Ing* 5(1): 122–29. <http://publication.lecames.org/>.
- Sattar Hanoon, Marwah, Ali Najah Ahmed, Arif Razzaq, Atheer Y. Oudah, Ahmed Alkhayyat, Yuk Feng Huang, Pavitra kumar, and Ahmed El-Shafie. 2023. "Prediction of Hydropower Generation via Machine Learning Algorithms at Three Gorges Dam, China." *Ain Shams Engineering Journal* 14(4). doi:10.1016/j.asej.2022.101919.
- Şebcioğlu Mutlu, Şeydanur, Abdulhadi Pala, and Aytaç Guven. 2025. "Comparing Traditional Hydrological Forecasting Models with CatBoost Algorithm: Insights from CMIP6 Climate Scenarios." *Journal of Water and Climate Change* 16(3): 1186–1208. doi:10.2166/wcc.2025.775.
- Shahani, Niaz Muhammad, Xigui Zheng, Xiaowei Guo, and Xin Wei. 2022. "Machine Learning-Based Intelligent Prediction of Elastic Modulus of Rocks at Thar Coalfield." *Sustainability (Switzerland)* 14(6). doi:10.3390/su14063689.
- Shrestha, Durga L., and Dimitri P. Solomatine. 2006. "Machine Learning Approaches for Estimation of Prediction Interval for the Model Output." *Neural Networks* 19(2): 225–35. doi:10.1016/J.NEUNET.2006.01.012.
- Da Silva, Marx Vinicius Maciel, Cleiton Da Silva Silveira, José Micael Ferreira Da Costa, Eduardo Sávio Passos Rodrigues Martins, and Francisco Das Chagas Vasconcelos. 2021a. "Projection of Climate Change and Consumptive Demands Projections Impacts on Hydropower Generation in the Sao Francisco River Basin, Brazil." *Water (Switzerland)* 13(3). doi:10.3390/w13030332.

- Singh, Raj, Nawraj Bhattarai, Anita Prajapati, and Shree Raj Shakya. 2022. "Impact of Variation in Climatic Parameters on Hydropower Generation: A Case of Hydropower Project in Nepal." *Heliyon* 8(12). doi:10.1016/j.heliyon.2022.e12240.
- Sun, Lijie, Dongxiao Niu, Keke Wang, and Xiaomin Xu. 2021. "Sustainable Development Pathways of Hydropower in China: Interdisciplinary Qualitative Analysis and Scenario-Based System Dynamics Quantitative Modeling." *Journal of Cleaner Production* 287: 125528. doi:10.1016/J.JCLEPRO.2020.125528.
- Teotónio, Carla, Patrícia Fortes, Peter Roebeling, Miguel Rodriguez, and Margarita Robaina-Alves. 2017. "Assessing the Impacts of Climate Change on Hydropower Generation and the Power Sector in Portugal: A Partial Equilibrium Approach." *Renewable and Sustainable Energy Reviews* 74: 788–99. doi:10.1016/J.RSER.2017.03.002.
- Turner, Sean W.D., Jia Yi Ng, and Stefano Galelli. 2017. "Examining Global Electricity Supply Vulnerability to Climate Change Using a High-Fidelity Hydropower Dam Model." *Science of The Total Environment* 590–591: 663–75. doi:10.1016/J.SCITOTENV.2017.03.022.
- Waheed, Abdul, Muhammad Hidayat Jamal, Muhammad Faisal Javed, and Khairul Idlan Muhammad. 2024. "A CMIP6 Multi-Model Based Analysis of Potential Climate Change Effects on Watershed Runoff Using SWAT Model: A Case Study of Kunhar River Basin, Pakistan." *Heliyon* 10(8). doi:10.1016/j.heliyon.2024.e28951.
- Wang, Chunyang, Chao Li, Yudong Feng, and Shoufeng Wang. 2025. "Predicting Hydropower Generation: A Comparative Analysis of Machine Learning Models and Optimization Algorithms for Enhanced Forecasting Accuracy and Operational Efficiency." *Ain Shams Engineering Journal* 16(3). doi:10.1016/j.asej.2025.103299.
- Wang, Haizhen, Xuerui Wu, and Fatemeh Gholinia. 2021. "Forecasting Hydropower Generation by GFDL-CM3 Climate Model and Hybrid Hydrological-Elman Neural Network Model Based on Improved Sparrow Search Algorithm (ISSA)." *Concurrency and Computation: Practice and Experience* 33(24). doi:10.1002/cpe.6476.
- Wasti, A, P Ray, C. Folch, and M Ubierna. 2022. "Climate Change and the Hydropower Sector: A Global Review." *Wiley Interdisciplinary Reviews: Climate Change* 12(2).
- Wei, Li, Li Jiheng, Guo Junhong, Bao Zhe, Fu Lingbo, and Hou Baodeng. 2020. "The Effect of Precipitation on Hydropower Generation Capacity: A Perspective of Climate

- Change.” *Frontiers in Earth Science* 8. doi:10.3389/feart.2020.00268.
- Yalcin, Emrah. 2023. “Quantifying Climate Change Impacts on Hydropower Production under CMIP6 Multi-Model Ensemble Projections Using SWAT Model.” *Hydrological Sciences Journal* 68(13): 1915–36. doi:10.1080/02626667.2023.2245815.
- Yohannes, Hamere, Teshome Soromessa, Mekuria Argaw, and Ashraf Dewan. 2021. “Impact of Landscape Pattern Changes on Hydrological Ecosystem Services in the Beressa Watershed of the Blue Nile Basin in Ethiopia.” *Science of The Total Environment* 793: 148559. doi:10.1016/J.SCITOTENV.2021.148559.
- Zhang, X. M., K. Grolinger, M. A. M Capretz, and L. Seewald. 2018. “Forecasting Residential Energy Consumption: Single Household Perspective.” In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, , 110–17.
- Zhao, Yinmao, Kui Xu, Ningpeng Dong, and Hao Wang. 2022. “Projection of Climate Change Impacts on Hydropower in the Source Region of the Yangtze River Based on CMIP6.” *Journal of Hydrology* 606: 127453. doi:10.1016/J.JHYDROL.2022.127453.

Scikit-learn Documentation: <https://scikit-learn.org/stable/>
 XGBoost Documentation: <https://xgboost.readthedocs.io/>

(<https://www.analytixlabs.co.in/blog/gradient-boosting-algorithm/>)

(<https://www.analytixlabs.co.in/blog/gradient-boosting-algorithm/>) 04 June 4, 2025, 8 pm. IPCC Climate Change, “Synthesis report. 2014. 4th assessment report of the intergovernmental panel on Climate Change,” in Intergovernmental Panel on Climate Change, R. K. Pachauri and L. A. Meyer, Eds., pp. 1–151, Geneva, Switzerland, 2014. Intergovernmental Panel on Climate Change IPCC, 2014. Summary for policymakers. In: Field, C.B., Barros, V.R., Dokken, D.J., Mach, K.J., Mastrandrea, M.D., Bilir, T.E., Chatterjee, M., Ebi, K.L., Estrada, Y.O., Genova, R.C., Girma, B., Kissel, E.S., Levy, A.N., MacCracken, S., Mastrandrea, P.R., White, L.L. (Eds.), *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, pp. 1–32.

IPCC (Intergovernmental Panel on Climate Change), 2021. Summary for policymakers. In: V. Masson-Delmotte, et al., eds. *Climate change 2021: the physical science basis. Contribution of working group I to the IPCC sixth assessment report*. Cambridge and New York: Cambridge University Press, 3–32.

Mitchell, T. M. (1997). *Machine learning*. In McGraw Hill Series in Computer Science. Retrieved from <http://www.worldcat.org/oclc/61321007> (au="Mitchell, Tom M." - Search Results, mars 30, 2025)

Coyne, G. et Bellier, EDF (1999). Complexe hydroélectrique de Souapiti-Kaléta : Étude de faisabilité. Volumes I, II, III et IV.

hydrography map of the Republic of Guinea Source: www.srp-guinne.org.....13

United Nations Framework Convention on Climate Change, 1992. UNFCCC.