1. Abstract

Kasai Oriental is one province of the Democratic Republic of Congo (DRC), in the country's heart. Annual flooding occurs in this province, causing significant damage because of heavy rainfall, especially in parts of Mbuji-Mayi city. Floods are one of the most catastrophic natural catastrophes, and modelling them is extremely difficult. The progress of flood prediction models has contributed to risk reduction, policy recommendations, a decrease in human life loss, and a reduction in flood-related property damage. In this paper, we worked on understanding the principal causes of flooding in this province, and proposing a concrete solution as part of a comprehensive approach to the management of the Kasai Oriental watersheds. Utilized the Neural Networks to predict the occurrence of flooding because of rising waters, as well as the use of a Climate Data Operator (CDO), to study actual temperature and precipitation. It proposed 4 needed Layers outputs (Temperature, Humidity, Pressure, Precipitation) after training the model, and an important variables, Temperature which is the subject of the study because of the lack of the flood data itself. After analysis, we discover that the temperature and precipitation vary according to the seasons. The results show which seasons may or may not have the possibility of flooding with predictive accuracy.

Keywords: *flood prediction; artificial neural network; big data; artificial intelligence ; time series prediction, Climate Data Operator, Watershed Management*

List of Abbreviations

AI	Artificial Intelligent
AIMS	African Institute for Mathematical Sciences
ANN	Artificial Neural Network
API	Application Programming Interface
BP	Back Propagation
CDO	Climate Data Operator
DJF	December,January and February
DMI	Dipole Mode Index
DRC	Democratic Republic of Congo
ENSO	Ei Nio Southern Oscillation
FF-NN	Feed Foward Neural Network
IDP	Internally Displaced Persons
JJA	June, July and August
KNN	K-Nearest Neighbour
MAM	March, April and May
NBI	Nature Based Interventions
NBS	Nature Based Solutions
ONI	Oceanix Nio Index
RNN	Recurrent Neural Network
RoC	Republic of Congo
SON	September, October and November
SWMM	Stomwater Management Model
UN	United Nation

List of Figures

2. Introduction

2.1 Background

Watersheds are natural phenomena of the universe, which can be influenced by human activities such as agriculture, environmental changes, and so on. When watersheds are not adequately modified, natural resources can easily decay to where they are no longer necessary for human improvement. We know that plants, air, water, and soil are natural resources that humans and animals need to live. A watershed is described hydrologically by the water that separate the distinct basins. These various lines are invariably natural boundaries drawn by the relief and are parallel to the crest lines. In certain parts of DRC, the state of the watershed, the amount and form of precipitation, the environment, and the flow of rivers may all contribute in the basis of flooding. Flooding happens when the amount of water in a stream exceeds the bed's potential. When the usual water limit reaches the river land level, flooding will occur along lakes and sea coasts. Floods are amongst the most prevalent natural disasters in the DRC, especially in the Kasai Oriental area, resulting in substantial losses of human lives, vegetation, the economy, and infrastructure Anusha and Bharathi (2020).

Rainstorms in DRC and the Republic of Congo (RoC) have caused rivers to overflow, causing floods in 16 of the DRC's 26 provinces and 8 of the RoC's 12 agencies since October 2019, with the nations bordering the rivers Ubangi and Congo and their tributaries being severely impacted Lateef et al. (2010). The rains were more intense in November and December 2019, causing further floods, catastrophes, and the eviction of hundreds of thousands of people. The most hit provinces in DRC were Nord-Ubangi, Sud-Ubangi, Mongala, Equateur, Kasai Oriental, and Tshopo, while the most hit provinces in RoC were Likouala, Cuvette, Plateaux, and Brazzaville and the United Nations (UN) claimed in December that almost 170,000 people in the departments of Likouala, Cuvette, and Plateaux alone are impacted in the Republic of Congo (Zarocostas, 2011).

2.2 Problem Statement

According to the actual flood situation in the DRC, there has been a significant loss of human life, property destruction, and environmental deterioration in the localized areas affected by these unforeseen incidents (languages French, 1960). Given that most households in the area depend on agricultural practices, trades, and livestock for their living standards, it is important to understand the impact of these floods on their socio-economic development.

Residents of Kasai Oriental province are the most affected by floods, which destabilize and kill, causes ravines, and destroy parcels. Floods also cause soil erosion which degrades the suitability of the soil and its fertility. We offer a method to aid them since we see the pain that the population is going through. (Fatras et al., 2021), Oppong and Woodruff (2007).

Figure 2.1 shows, for example, a depiction of the Flood in Kasai Oriental.

Figure 2.1: Flood in Kasai Oriental
for the Coordination of Humanitarian Affairs
(UN OCHA)

2.3 Aim of the Study and Research Objectives

The aim of this study is to predict flooding based on historical data to see how often flooding may or may not occur from one area to another over time.

Given the disasters brought about the floods (presence of ravines which destroy roads, etc.), the objective is to secure the population against the danger of erosion, with our model using Neural artificial intelligence and Climate Operator Data to study the temperature and rainfall.

To this, the secondary objectives are associated :

- Study the climatology of the province (precipitation and temperature).
- Ensure a perfect return : that is to say, put in place methods or ideas to return to the state before disasters repair the damage that has occurred.
- Maintain the safety of the population : Target inhabited places at lower risk during the period of flood crisis.

Some motivations that led us to choose this subject are :

- the harmful effects of flooding in this region,
- suffering caused by the floods after having ravaged the fields and loss of property of the inhabitants of this region of Kasai Oriental.

3. Literature Review

3.1 Introduction

This chapter gives a short overview of some studies relative to Watershed Management. The discussion is based on some methods and materials used in different past study done by other researchers in flood prediction.

3.2 Watershed Management

After going through (Asdak et al., 2018), The author is based in Jakarta, and the goal is to first evaluate the fundamental factors that cause flooding in Jakarta's coastal environment, and then provide mid to long remedies from the standpoint of an integrated watershed management strategy. In a cross-administrative border scenario, the proposed approach would also require structural and financial arrangements. However, the author of this paper did not take the time to quantify the probability of floods in Jakarta.

When integrated flood risk management is necessary, the identification of risk analysis approaches that involve systematic actions as part of continuing preparations, response, and recovery is necessary. The rainfall that caused the floods in Jeddah in 2009 and 2011 is investigated in Qari et al. (2014), and their return time is predicted. It is shown that, while designing structural defenses is feasible in certain circumstances, it is not possible in others, and non-structural defenses should be built instead. Examples of both structural and non-structural defense interventions in action in various countries are shown in his article.

In Southern Africa the author Tumbare (2000), examines potential flood factors including El Nino/La Nina impacts, Global Warming, human interaction with the atmosphere, reservoir and dam activity. He goes on to discuss how best to prevent flooding, including the use of climate change projections, mathematical models to predict disasters, and the flood disposal and warning mechanisms that must be in place, with a focus on Southern Africa.

While Löschner et al. (2017) examines the impact of climate change and land growth on potential flood risk in a number of Austrian flood-prone communities in this report. He predicted four separate inundation scenarios for current and projected 100 and 300 year floods (including climate change allocations) as part of an anticipatory micro-scale risk assessment, developed predictions of future population development in floodplains, and analyzed improvements in flood damage potentials and flood risk until 2030. His results suggest that potential flood risk would be greatly increased by both climate change and settlement growth. The case study report then emphasizes the need for a more thorough examination of the local causal factors of flood risk in order to improve the efficacy of integrated flood risk management .

3.3 Artificial Neural Network

An artificial neural network (ANN) is a component of a computer system designed to replicate or disassemble the way the human brain processes and analyzes data Jain et al. (1996). It is typically the backbone of artificial intelligence (AI) and addresses challenges in society that would be impractical or difficult to solve using human or mathematical criteria.

Centered on the neural networks approach, in (Feng and Lu, 2010) the authors suggested a new flood prediction algorithm with similar applications for this study. This approach has been used to provide improved production and productivity outcomes. The implementation of this method is supposed to boost sensitivity and improve flood prediction efficiency. As a result, the use of flood forecasting and emergency preparedness is needed to minimize financial losses and human risk. The author presents a flood forecasting model that uses real-time data from the basin (information on dam operation, hydrometric data and rainfall data) to forecast river progression. The system is predicated on artificial neural networks, that have already been successfully used to forecast floods in an uncontrolled basin and water-level progression in the Arno basin below low operating conditions in previous studies. Using a two-year data set and a preferential treatment of data input, accurate forecasts are provided, allowing a balance to be reached between the sensitivity and specificity of precipitation data and the model sophistication (Campolo et al., 2003).

4. Aspects of flooding Systems

4.1 Presentation of the study site

4.1.1 Location. Kasai Oriental is a diamond-bearing province of the DRC, located in the center of the country with Mbujimayi as the Chief town. This province is made up of city of Mbujimayi and 5 territories : Katanda, Lupatapata, Tshilenge, Miabi, Kabeya-Kamwanga, having an area of 9545 km^2, a population estimated at 5.475.398 inhabitants (2015) as well as a density of 574 inhabitants / km^2 (Tshonda).

The figure 4.1 shows the Congo River basin with the Kasai Oriental watershed located between $15°\ 30'$ and $25°$ East and $1°$ and $12°$ South, in general the area of the Kasai Oriental basin is estimated at 730,000 km^2 and covers almost 20% of the total area of the basin. The Kasai River which is the main tributary of the watershed has its source in Angola where the basin shares the ridge line with the Zambezi River basin (Wéry et al., 1976).

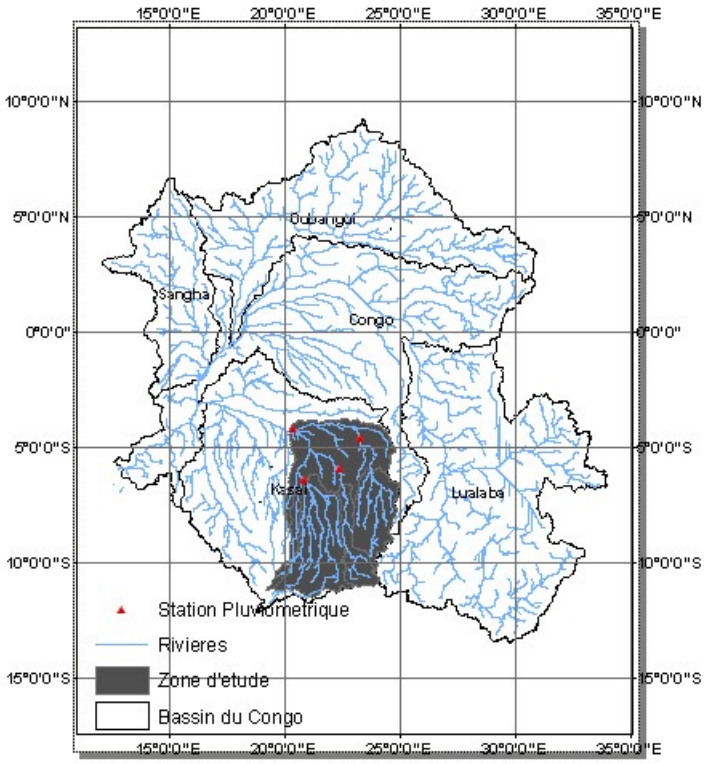


Figure 4.1: Map of the Congo Basin showing the Kasai sub-basin

4.1.2 Vegetation and Soil of Kasai Oriental Basin. We distribute two types of soil in the Kasai Oriental basin : Ferralitic soils (Soils with high fertility) which occupy the northern part of the basin (this southern part of the central basin which is the domain of the equatorial forest), and Tropical ferruginous soils which is found on the Kwango-Kwilu and the Kasai Oriental plateau and abounds in a dense dry forest, clear forest and savannah (de Dorlodot, 1920).

4.1.3 The Geology of Kasai Oriental Basin. Kwango is a part from which Kasai Basin originates. Note that the Mesozoic shows a geological appearance that connects them to the formation of Lualaba unit and this formation rests on the bedrock dating long ago by means of an almost fluvial conglomerate. The cover lands in this vast region are flourishing and the base sometimes only emerges at the bottom of the deep valleys (Tshonda),(Roberts et al., 2015).

From downstream to upstream, the stratigraphic levels in the Kasai Oriental basin are :

- Archean rocks: Which currently abound in three units which are :

  Haute-Luanyi gneiss and granulite complex which is made up of gneissic granulite rocks.

  The gabbro-noritic and charnockitic one of kasai-Lomami

  That of migmatites and migmatitic granites of Dibaya which is currently a large assemblage of calc-alkaline ganites and granitic to tonalitic migmatites with occasional septa of amphybolites and pyroxenolites.

- Upper Proterozoic rocks which are represented by "Bushimaie Supergroup", measuring about 1600 m. it is made up of groups such as :
  - Schisto-limestone (more than 1030 m of power)
  - Schisto-sandstone ($\pm$ 450 m of power)

4.2 Management of the risk of flooding in the susceptible environment.

The concept of disaster management is that of reducing human and also economic vulnerabilities while securing what is safe the place. This notion is divided into three dimensions: prevention and mitigation, crisis management and operations in states of emergency. One of the primary occupations is to know in the valid period the exact extent of the flood as well as the identification of vulnerable populations. The last dimension of the study is based on the analysis of potential floods to assist in the control of physical phenomena and also to limit future risks. This dimension also consists of drawing conclusions about the consequences of the flood on all levels as well as analysing the experience of people living in areas affected by them. Taylor et al. (2011). "This phase makes it possible to build or improve policies for risk prevention and protection of vulnerable elements in the face of future flooding" Mind'je et al. (2019),Zaharia et al. (2017).

4.3 Types of watersheds

Several surveys of hydrological chronicles conducted on various watersheds around the globe indicate, that these hydrological bodies behave very differently during heavy rainfall, whether or not correlated with floods. As a result, the key parameters that describe these events can be identified. The primary stress (except in coastal basins) is undoubtedly rain, with its severity,

spatial distribution, and length. However, based on the characteristics of the basins : type of soil, occupation, and drainage network, the same rainfall distribution may cause different forms of flooding (Brooks et al., 2003). After our research, we found that there is an urban basin, a rapid basin, a coastal basin, an estuary, a middle river, a tablecloth basin, a large river.

4.4 Meteorological and hydrological considerations

In flood forecasting and alert, the capacity to anticipate catastrophic disasters, both temporally and briefly, spatially and quantitatively, is undeniably advantageous. Flooding is caused mostly by weather conditions, which may take the form of rain, snow, or meltwater. Understanding precipitation processes, their seasonality, and the extremes of their activity are all part of climatology. Whitfield (2012).In an arid region where flash floods are common, analysis and forecasting techniques must be dependent on the event's real-time detection. Understanding which types of weather conditions are likely to cause floods can help determine which observing and forecasting systems are essential in such an area as the Sahara Desert. Bartholmes and Todini (2005),Jasper et al. (2002)

5. Data and Prediction Method

This chapter explains the data's source, the data collection processes, and some technical data processing approaches. We use ANN to manage data and predict flood in Kasai Oriental using this model.

5.1 Data

The data used are gathered using different types of sensors to sample temperature, humidity, pressure, precipitation etc. These data are historical from weather forecasts for 1,414 sectors and villages in this area, which covers a range of 10 years of available data. We are able to obtain them by using a Local Weather Application Programming Interface (API) that transmits real-time metrology data, average monthly temperature data, and weather warnings. The local weather API allows both individuals and companies to access weather forecasts for any area helped by using a python code, Du et al. (2018) .

The data we obtained covers the period from December 31, 2008 to December 31, 2020. It is climate data with many factors that may be used to forecast floods in this region of DRC.

5.2 Methodology

5.2.1 Research Methodology. An in-depth examination of the difference of high frequencies for all stations with at least 12 years of data on this form of order. Rainfall stations (sensors) were used to gather historical data from multiple areas in the province of Kasai Oriental and got using World Weather Online historical weather data API Python wrapper. The space was exploited for the flood map in this area using this data.

5.2.2 Artificial Intelligence Neural Network. An ANN is a component of a computer system that is programmed to approximate how the human brain analyzes and processes data Yegnanarayana (2009). This is referred to as the fundamental basis of artificial intelligence (AI) in the sense of resolving issues that are impossible or unsolvable by human standards. The capabilities of the ideal self-learning ANNs, which allow them to have better outcomes when data is available.

We are using python programming as a coding language in this analysis, which is commonly used in a variety of fields and has a high degree of general use. Several methods and instruments for data processing can be used in flood prediction, including Feed-Forward Back Propagation, which we used in our research. An example of ANN operation is depicted in the figure 5.1. Where we have inputs layers $x_1, x_2, x_i ... x_n$, that are interpreted by the model, and the latter regenerates the states of the interpreted outputs layers $y_1, y_2, y_k ... y_l$.

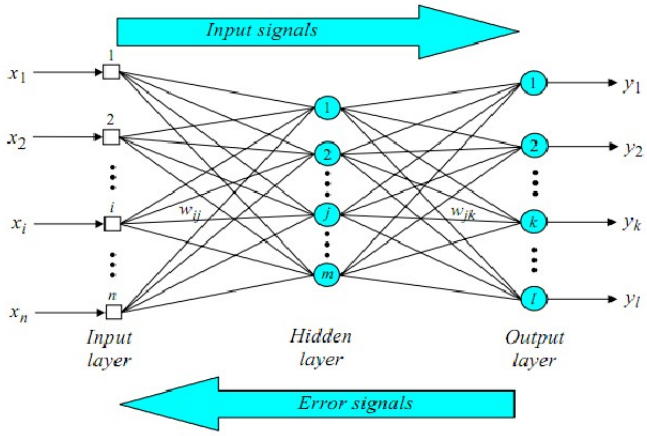


Figure 5.1: Architecture of ANN-with-back propagation (BP)-learning-algorithm

5.2.3 Feed Forward Neural Network. The FFNN approach built by David Rumelhart is the model we've selected for this work because of the amplitude at which it develops in different fields, such as flood prediction, failure prediction, and building. This standard approach focusing specifically related to the algorithm called the back-propagation standard is commonly used for hydrological applications because they are accurate, easy and fast. The information passe into the model, where it is multiplied by a weight that reinforces the corresponding signals, thus engaging an activation function, as seen in the figure 5.2. The output layer can perform calculations on the data by the time it reaches a certain dimension in order to enable an outputs. It means that, this model starts with the sum of the data inputs and then develops them using a nonlinear activation function to produce the final output (Bebis and Georgiopoulos, 1994).

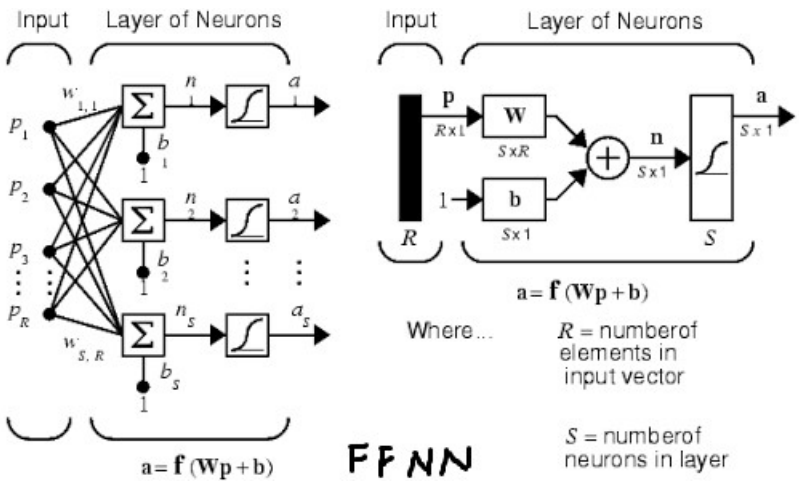


Figure 5.2: Backpropagation: Feed Forward Neural Network

From figure 5.2, let's assume that,

$$\textbf{the input layers} \quad p = x \qquad \textbf{and} \quad \textbf{the output layers} \qquad a = y$$

Where w is the weight, and b is the bias.

The shape of the data is determined by the weight of each individual neuron, and this weight can be changed by algorithms back-propagation in order to achieve desired outcomes.

This method shows neural network construction, deployment, data scaling, performance, review, and, eventually, neural network validation and acceptance. Before the anticipated outcome is attained, data scaling, architecture selection, and network learning all happen at the same time. With x input nodes, y output nodes, and L hidden nodes, a three-layer neural network is used for back propagation and classification. A bias vector b, a weight matrix w, and a transfer function f are common components of the neuron sheet. To discover the highest number of layers and hidden nodes, trial and error is employed. If the data weights are sufficient, a hidden layer network may theoretically snap to any continuous function. Without a doubt, the hidden layer is the finest option. The expected and desired outputs, on the other hand, have the same number of neurons as the output layer (Chibueze and Nonyelum, 2009).

The input neurons accept and interpret input signals before activating an output signal that is sent to other similarly active neurons in the network, which has a single activation function and a single function that may be non-linear or linear and still continues. The transmitted signal passes through the neuron and is altered by weights, which adjust the functions such that the output signal reaches the next neuron. The output and propagation of information from the input layer to the output layer via the hidden layer was changed as a result of this transformation of the weights of neurons in the network (Ogwueleka and Ogwueleka, 2009). In the hidden layer, the input node is given by :

$$x_j = \sum_{i=1}^{N} w_j x_j + b_j, \qquad j = 1, ...N \tag{5.2.1}$$

Thus, the node's output name can be accessed :

$$y_j = f(x_j),$$

Then we have,

$$f(x_j) = [1 + e^{-x_j}]^{-1} \qquad j = 1, 2..., N \tag{5.2.2}$$

So, to include the k^{th} node's input in the output layer, we'll use the following mathematical formula :

$$x = \sum_{j=1}^{N} w_k y_k + b_k, \qquad k = 1, 2, ...N \tag{5.2.3}$$

And the output of k^{th} will be :

$$y_k = f(x_k) = [1 + e^{-x_k}]. \tag{5.2.4}$$

Throughout the process, learning and errors generally leak over from exit nodes to interior nodes. Backpropagation is then used to determine the gradient of the network error, at the expense of the modifiable network weights. The categorization is selected since, unlike learning, it is a fairly quick procedure. The objective function to be lowered is the root mean square error across all

forms of learning. As a result, the root mean square error is calculated as follows:

$$E = \frac{1}{M}\sum_{p=1}^{M}(t_k - y_k)^2, \tag{5.2.5}$$

where M is the number of samples, and t_k is the estimated production.

The activation function is a mathematical function corresponding to a signal output from one of the neurons in ANN. The one that is most often implemented in nodes is a logistic sigmoid function, which generates a binary output between 0 and 1, and admits linear relationship in the network. This maximizes the chances of getting a nonlinear relationship between the input and output values, and the logistic equation is rewritten as . (Leshno et al., 1993) :

$$f(x) = \frac{1}{1 + e^{-x}} \tag{5.2.6}$$

5.2.4 Analogies of FF-NN with polynomial curve fitting. An analog neural network is a convolutional neural network that uses electrically changeable hardware to simulate the function of a neural synapse at the right moment. This network is also used to highlight the reliance on physical hardware in order to replicate neurons, as opposed to software-based alternatives. It's worth noting that human brain creates analogies, which is a unique feature that doesn't always occur in other neural network

We may use the example of a youngster who does not know how to use a phone when he is born to demonstrate this point. The first time he sees an image of the phone, it sticks in his mind, and so on. Even if we bring him a different brand of phone and ask him what he sees, he will reply phone even though he has never used the new model. To put it in other words, the youngster took the time to study based on the phones he or she had seen.

Since each problem may be described by a mathematical curve that interprets the evolution of the latter, we will use this concept in our work because our outcomes are represented by graphs. As a result, fitting a curve is merely a process of constructing the latter, or even a mathematical function that adapts to a set of points subjected to integrity restrictions (O'Neill et al., 1969).

5.2.5 Water level forecast. The historical water level, the historical precipitation, and the precipitation of the current days are all used to anticipate the water level. The output, in turn, is the current day's water level. The example below shows the water level from 4 days of historical data:

$$N(t) = f[P(t-4), P(t-3), P(t-2), P(t-1), P(t), N(t-4), N(t-3), N(t-2), N(t-1)], \tag{5.2.7}$$

Where P is precipitation, N is the expected water level and t is the time / days

6. Data Analysis

6.1 Introduction

In this chapter, we show the result of our research, which were collected utilizing a variety of tools, including the CDO, which is a set of instructions that permits climate studies to be conducted. We used temperature and precipitation data for the whole continent of Africa, that can be found in the two files below : chirps v2.0.1985-2004.days p25.nc and chirts Africa masked.nc, which are Daily Datasets for Climate Extreme Processing in CDO. Then, using Python, we analyzed the data we received from Kasai,and used the FF-NN technique to predict flooding.

6.2 Results and Discussion

Figure 6.1, depicts seasonal temperature in DRC during period from December, January, and February (DJF), March, April, and May (MAM), June, July, and August (JJA), and September, October, and November (SON). To do so, we used data from Africa, which could only offer us the departure status of a whole country after programming, not a province. As a result, we selected the portion of Kasai Oriental that is the subject of our study using this map of the DRC that displays temperature fluctuation in all regions. We can see that the average temperature is $25°C$ for all seasons except the JJA season, which is the dry season. During the dry season, the average temperature ranges from $25°C$ in the north to $22.5°C$ in the south, with lows of $18°C$.

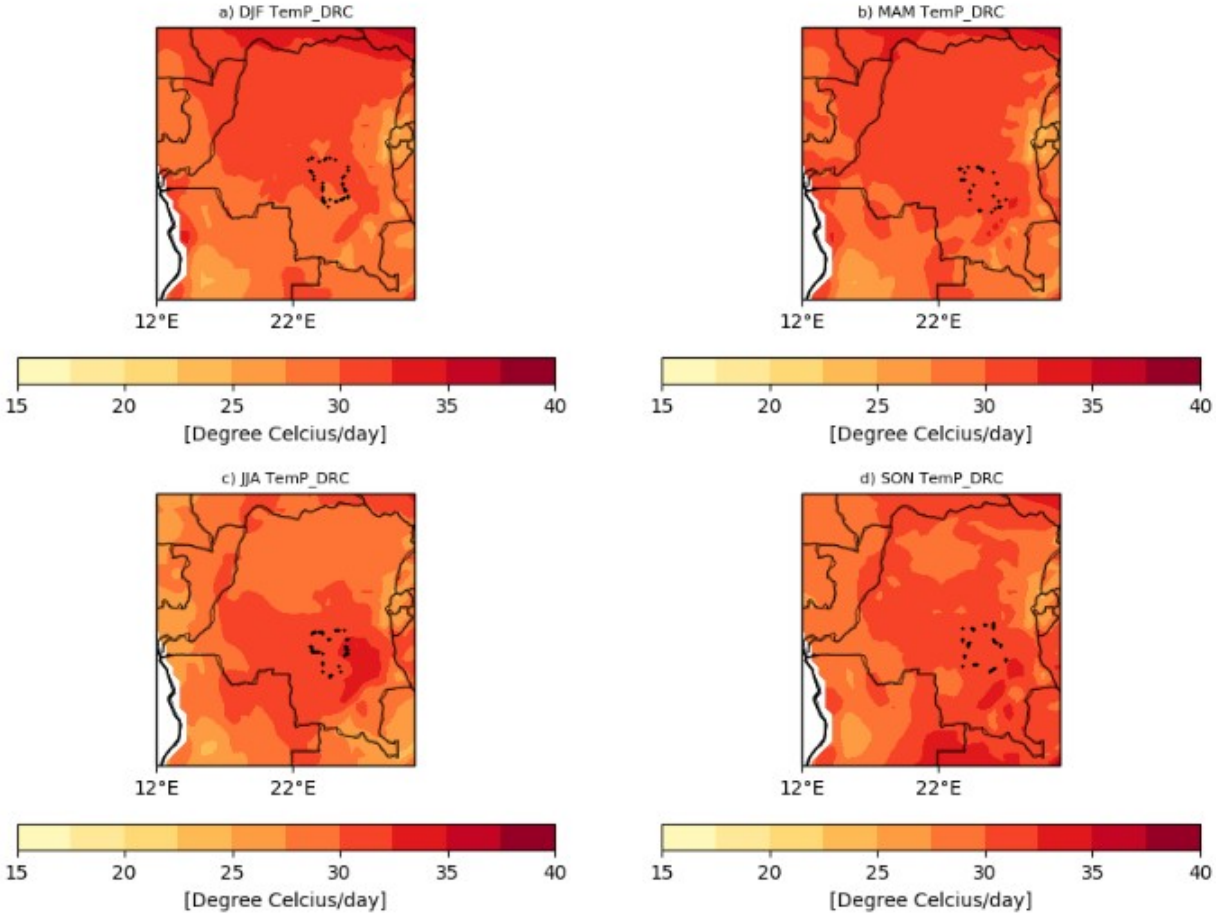


Figure 6.1: Temperature for all Seasons in Congo(Kasai-Oriental)

Figure 6.2, depicts the precipitation in DRC, as previously said, we are concentrated on the Kasai Oriental region. During the DJF season, we note a high amount of precipitation up to $9mm$ in certain Kasai Oriental towns, but the amount drops substantially to $6mm$ in MAM as we approach the dry season, which starts in May and lasts until August. So, as seen in the map below, JJA is a season in which there is no precipitation in this province ($0mm$), and then the amount of rainfall increases to $7mm$ as we go closer to SON.

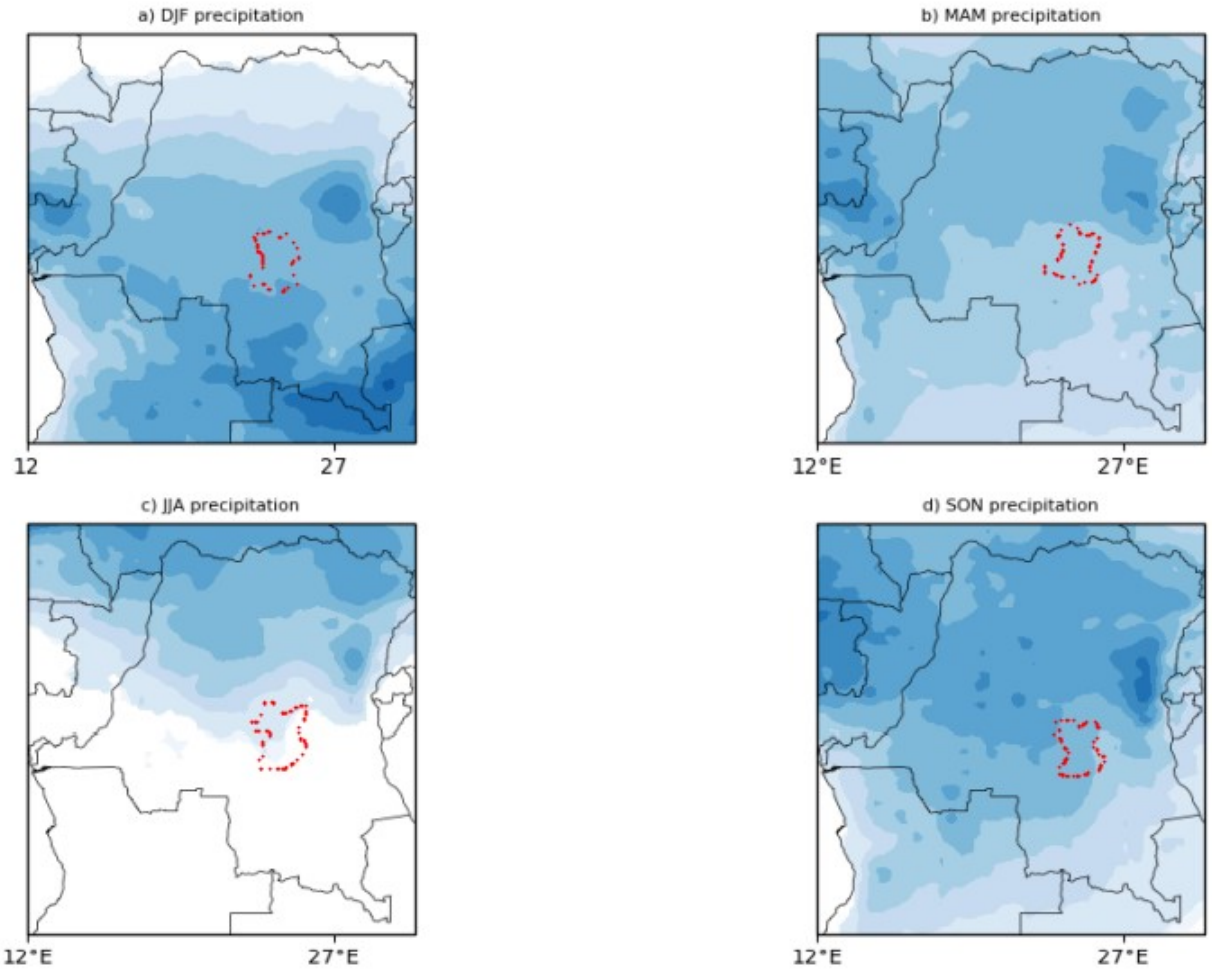
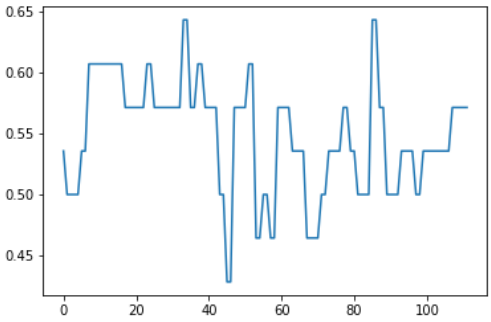


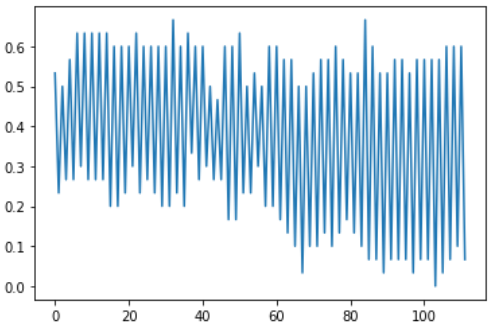
Figure 6.2: Precipitation for all Seasons in Congo(Kasai-Oriental)

Note that the model that we use generates a random batch of 256 sequences based on our data, with each sequence including 112 observations and every observation involving 15 input-signals and 4 output-signals.

From figure 6.3a, we present the plot of one of the 15 input signals and we could also visualize in figure 6.3b one of the output signals that the model should learn to predict given all 15 input signals.



(a) Figure 5.3, One of Inputs Signals



(b) Figure 5.3, one of the output signals

We can easily conduct several training epochs since the neural network trains rapidly. However, there is a danger of overfitting the model to the training data, resulting in poor generalization to unknown data. As a result, we'll track the quality of the model on the test-set after each epoch but only store the model's weights if the test-set health improves.

From figure 6.4, we can see the summary of our model. The (None, None, 4) output shape indicates that the model generates a batch with such an arbitrary number of iterations, each of which contains an arbitrary number of observations, and each observation includes 4 signals. This correlates to the 4 target signals we're attempting to forecast.

```
model.summary()

Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
gru_2 (GRU)                  (None, None, 512)         812544
_________________________________________________________________
dense_1 (Dense)              (None, None, 4)           2052
_________________________________________________________________
module_wrapper_1 (ModuleWrap (None, None, 4)           0
=================================================================
Total params: 814,596
Trainable params: 814,596
Non-trainable params: 0
_________________________________________________________________
```

Figure 6.4: Model: sequential1

We can see from figure 6.5 and 6.6 that precipitation is monthly and varies throughout the years. The lowest frequency was seen in 2008, once that occurred every 10th month of the years 2009 and 2010, then every 7th month of the years 2009 and 2010, then every 6th month of the years 2009 and 2010. However, the highest frequency was recorded in December 2009 and 2013, followed by 2012 and Avril 2009.

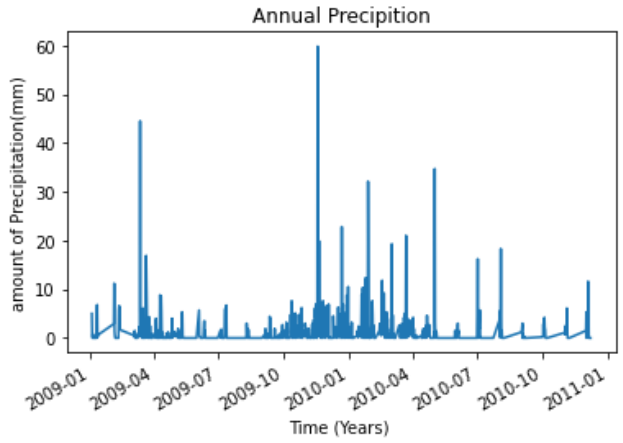


Figure 6.5: Precipitation Over some years

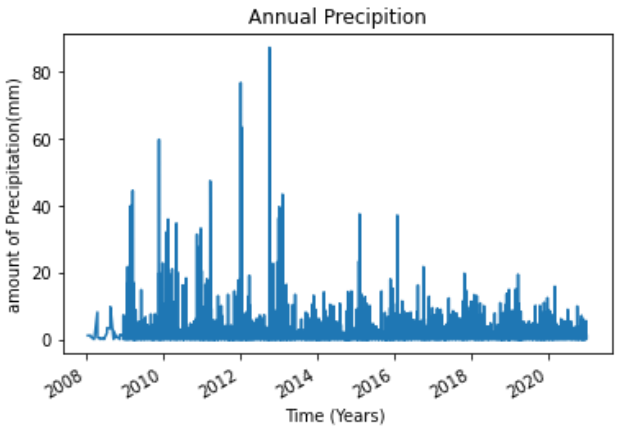


Figure 6.6: Precipitation Over some years

We can now visualize a sample of expected output-signals after training the model. Only the output-signals are shown in these figures 6.11, not the 15 input-signals that were used to forecast the output-signals. In these plots, the period between the input and output signals is kept constant. The model always forecasts output-signals in the future, for example 2 days in the future, but only throughout the time period chosen. As a result, the x-axis of the graphic simply indicates how many time-steps of input-signals have been detected by the prediction model thus far. Because this model creates a single time-step of output data for each time-step of input data, it knows very little about the history of the input signals after only a few time-steps and cannot produce an accurate forecast.

We can forecast 4 variables : temperature, pressure, precipitation, and humidity, which are the study's essential variables. The time series forecasts the future using previous data, and we should remember that the blue color in the graphs reflects reality (true) and the orange color indicates forecast. Consider that, figure 6.7 shows the temperature which is reasonably well predicted in the test set, while the peaks are occasionally wrong. It changes with time, and the actual highest temperature of the year at position 55 is at least $28°C$ on 12/21/2009 in Kasai Oriental, while the model forecasts $26°C$ in the future on the same date. The present temperature is $20°C$ in 2009, and the model forecasts a future temperature of $24°C$, and so on.

With the exception of a lag and a more noisy signal than the genuine time-series, the atmospheric pressure shown in figure 6.8 is pretty accurately predicted. We can observe that the air pressure was low in 2009 at 1008 in altitude, and that the future pressure projection is 1010 in altitude , and so on. To begin, it should be remembered that air pressure must be measured at a specified altitude in order to be reasonable. In reality, as altitude rises, the pressure drops (by half at 5 500 m), since the weight of the air above us becomes less essential. This fluctuation, however, is not linear. There are calculations that take into consideration variations in altitudes and temperature to determine it. At low altitudes, the decline is higher than at high altitudes.

We have precipitation in figure 6.9, which is the main variable because we do not have flood data for this region of the Kasai, so we take a second option of predicting flooding by the amount of the rainfall, and we project into the future using these precipitation data to see the possibility of

flooding. We can see from the graph that the forecast is ineffective, it does not follow virality, and the model did not do a good job of predicting precipitation. However, according to the data provided by this same model, there was a flood in 2017 and the program forecasts that there will be no flood in the future.

Finally, the figure 6.10 presents the humidity, which is effectively predicted despite a few notable minor errors from the first years to the nineteenth year. In recent years, we've seen that the current humidity is so high that the model only predicts a medium humidity.

A 3 % rise in relative humidity is possible for every degree decrease in temperature. The humidity can be decreased to a more respectable level by reducing the temperature. Such as illustration, if the relative humidity of air warmed to 23°C is 18%, by reducing the temperature by 4°C, it may be raised to 30%.

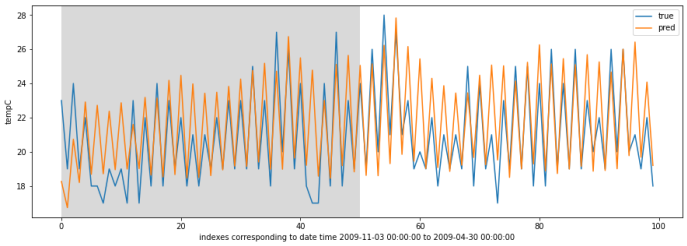


Figure 6.7: Plot of Temperature

Figure 6.8: Plot of Pressure

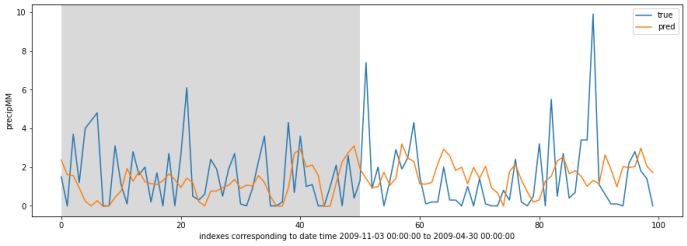


Figure 6.9: Plot of Precipitation

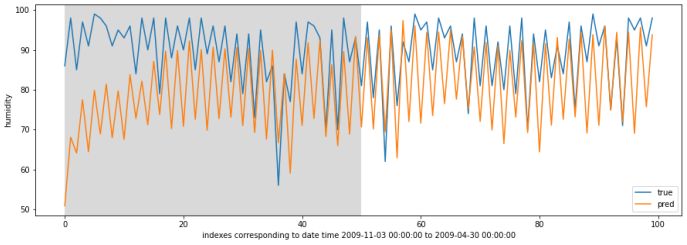


Figure 6.10: Plot of Humidity

Figure 6.11: Output Signals

Since we discovered that the aforementioned model does not accurately predict precipitation very vell, we must all come up with a persuasive conclusion for this project. To solve this issue, we chose to test another model called Flood Prediction Model, which can forecast changeable precipitation with constant precision by concentrating on K-Nearest Neighbor (KNN), which is among the simplest Machine Learning algorithms focused on the Supervised Learning methodology.

To find the nearest neighbors, KNN use a similarity metric. The Euclidean distance among both our unknown variable and some other variables in the dataset is frequently used as this similarity measure. The following is the general formula for Euclidean distance:

$$d(p,q) = d(q,p) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + ... + (p_n - q_n)^2}, \tag{6.2.1}$$

where the attribute values for one observation range from q_1 to q_n, while the attribute values for the second observation range from p_1 to p_n.

The only data we need from now is precipitation data for all years from 2009 to 2020, which will allow us to build a forecast. The structure of the data used is depicted in the diagram 6.12 below.

```
[ ] df = read_csv('/Kasai_OR_precipMM.csv')
```

```
df.head(12)
```

	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual Rainfall
0	2009	147.5	155.8	141.5	61.6	21.3	36.2	34.4	18.2	40.2	87.6	201.2	135.1	1080.6
1	2010	150.5	125.0	128.6	47.6	66.9	11.9	42.2	46.9	23.0	50.1	156.7	218.5	1067.9
2	2011	89.2	83.2	195.8	40.1	25.9	41.2	38.0	18.7	32.5	44.9	157.5	125.5	892.5
3	2012	218.8	104.6	129.0	19.2	10.9	16.3	14.0	21.7	17.2	131.6	106.1	140.8	930.2
4	2013	97.1	91.3	107.9	24.2	28.5	18.5	15.5	21.3	34.5	47.2	103.4	76.0	665.4
5	2014	30.1	50.9	81.9	33.2	27.4	34.9	20.9	22.7	10.9	53.8	80.9	10.3	457.9
6	2015	82.7	137.1	78.4	58.0	18.1	16.3	17.2	10.0	33.9	34.6	77.7	100.1	664.1
7	2016	156.7	90.0	86.6	9.8	24.8	31.2	19.1	33.0	23.7	46.0	84.9	89.7	695.5
8	2017	95.3	76.0	109.4	33.6	17.2	32.4	30.8	26.6	24.6	82.6	124.9	131.7	785.1
9	2018	157.1	95.4	100.0	52.1	21.4	13.5	15.8	26.9	14.8	61.8	105.9	166.5	831.2
10	2019	125.8	148.6	100.6	78.8	38.2	16.6	31.6	31.4	32.5	86.9	125.9	140.9	957.8
11	2020	78.1	65.5	107.9	37.0	24.8	22.1	18.6	16.2	23.1	34.7	81.1	67.2	576.3

Figure 6.12: Structure of Precipitation Data

We can see the behavior or fluctuation of precipitation during the last 12 months of each year from 2009 to 2020 in figure 6.13. The graph shows that the amount of precipitation was quite high during year 0 (considered the first year) until it reached 1099 mm, and then gradually decreased until year 5 (when it reached less than 500 mm). After this year, we saw a significant change until the ninth year, which accumulated 1000 mm, and finally, there is rigorous decreasing up to the last year, which accumulated 600 mm.

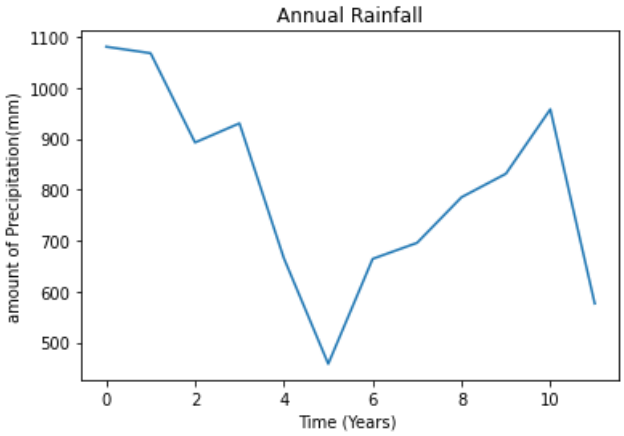


Figure 6.13: Annual Railfall Plot

Figure 6.14 indicates the amount of precipitation for 4 seasons of all years : DJF, JJA, and SON. We may examine the amount of precipitation for each month of each year, as well as each season of each year. Because it is rainy season in Kasai Oriental, the months of December, January, and March have a higher intensity of precipitation than the others in years 1, 3, and 2. Since May, June, July, and August are the 4 months of the rainy season in Kasai Oriental, their amount is too low every year.

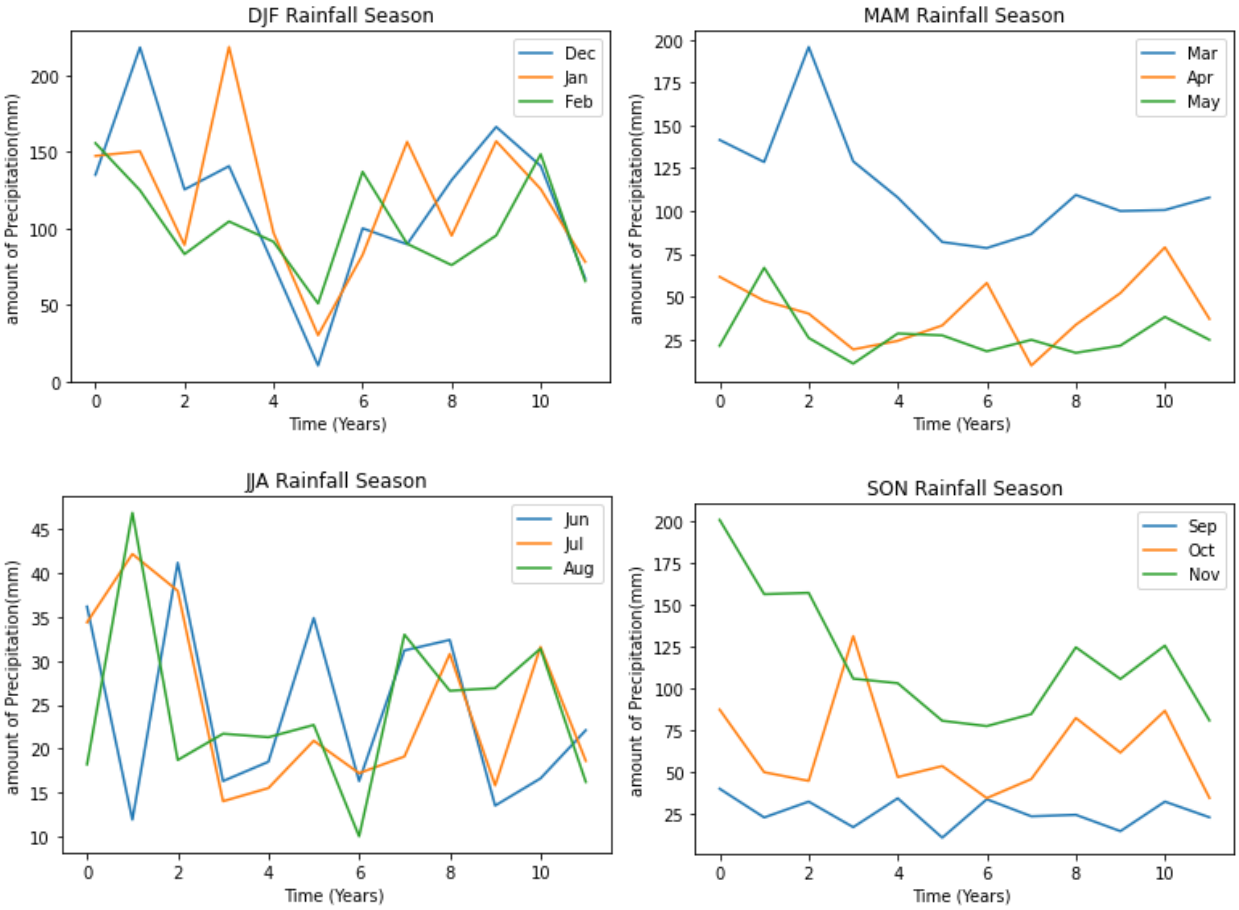


Figure 6.14: Seasonal precipitation

We may take at random the water level threshold of Lubilanji river in Kasai Oriental and use to declare a flood year. The flood threshold has been set at $700mm$, which means that each time the amount is more than or equal to $700mm$ throughout the year, there is a chance of a flood (so 1). If the amount of the rain is less than 700 mm, it means there will be no flood this year (therefore 0), figure 6.15 depicts how the model is able to forecast. We can see that the first four years were flooded, the next four were not, and the next three years were still flooded, despite the fact that the last year was not.

	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual Rainfall	Flood
0	2009	147.5	155.8	141.5	61.6	21.3	36.2	34.4	18.2	40.2	87.6	201.2	135.1	1080.6	1
1	2010	150.5	125.0	128.6	47.6	66.9	11.9	42.2	46.9	23.0	50.1	156.7	218.5	1067.9	1
2	2011	89.2	83.2	195.8	40.1	25.9	41.2	38.0	18.7	32.5	44.9	157.5	125.5	892.5	1
3	2012	218.8	104.6	129.0	19.2	10.9	16.3	14.0	21.7	17.2	131.6	106.1	140.8	930.2	1
4	2013	97.1	91.3	107.9	24.2	28.5	18.5	15.5	21.3	34.5	47.2	103.4	76.0	665.4	0
5	2014	30.1	50.9	81.9	33.2	27.4	34.9	20.9	22.7	10.9	53.8	80.9	10.3	457.9	0
6	2015	82.7	137.1	78.4	58.0	18.1	16.3	17.2	10.0	33.9	34.6	77.7	100.1	664.1	0
7	2016	156.7	90.0	86.6	9.8	24.8	31.2	19.1	33.0	23.7	46.0	84.9	89.7	695.5	0
8	2017	95.3	76.0	109.4	33.6	17.2	32.4	30.8	26.6	24.6	82.6	124.9	131.7	785.1	1
9	2018	157.1	95.4	100.0	52.1	21.4	13.5	15.8	26.9	14.8	61.8	105.9	166.5	831.2	1
10	2019	125.8	148.6	100.6	78.8	38.2	16.6	31.6	31.4	32.5	86.9	125.9	140.9	957.8	1
11	2020	78.1	65.5	107.9	37.0	24.8	22.1	18.6	16.2	23.1	34.7	81.1	67.2	576.3	0

Figure 6.15: the prediction of years of flooding

We take three variables x-test, y-test, and y-predict after training the model with the KNN Classifier method. The x-test variable provides data for three years chosen at random from a list of 12, while the y-test variable provides the present condition of those 3 years, indicating if flooding occurred during those years.

The data for the 3 years: 9, 1, and 3 are shown in figure 6.16.

```
x_test
```

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
9	157.1	95.4	100.0	52.1	21.4	13.5	15.8	26.9	14.8	61.8	105.9	166.5
1	150.5	125.0	128.6	47.6	66.9	11.9	42.2	46.9	23.0	50.1	156.7	218.5
3	218.8	104.6	129.0	19.2	10.9	16.3	14.0	21.7	17.2	131.6	106.1	140.8

Figure 6.16: the Variable Xtest

The figure 6.17 shows the current flood values for the years 9, 1, and 3, and informs us that flooding occurred during each of these three years as shown in 6.15. After training, figure 6.18 the model predicts that floods will occur in years 9 and 1, but that flooding will not occur in year 3. The accuracy of the model is 0.66666 so which is 66 percent as shown in figure 6.19.

```
actual values of floods:
9    1
1    1
3    1
Name: Flood, dtype: int64
```

Figure 6.17: Actual Values of floods

```
predicted chances of flood
[1 1 0]
```

Figure 6.18: Predicted Chances of flood

```
knn_accuracy.mean()
```

```
0.6666666666666666
```

Figure 6.19: Accuracy

Figure 6.20: Results

Both of these strategies are helpful in various ways, yet they both have flaws. The FF-NN is ineffective in predicting precipitation, hence the Flood Prediction Model based on the KNN is more appropriate and easier to use. In terms of comparison, we can state that the second approach has less mistakes and is faster, but the first does not accurately anticipate precipitation,so the error is larger.

7. Conclusion and Recommendation

FF-NN can accurately forecast temperature, humidity, and pressure, but not precipitation. The study's findings revealed that while prediction of precipitation is ineffective, prediction of other variables using the FF-NN is realistic, suitable, and of sufficient precision. The use of these methodologies would increase data confidence, allowing for better watershed management. We used data from the Kasai Oriental province in north-east Africa to develop a computer model that forecast rainfall, air pressure and humidity. It functioned well enough to forecast temperature in situations when daily oscillations were adequately anticipated, but peaks were not always precisely anticipated. Although the modest inaccuracies and the expected signal were stronger and had a shorter lag, atmospheric pressure was also reasonably anticipate. A neural network can readily find patterns in data that are just transitory correlations that do not transfer well to hidden data. Humidity is also accurately predicted, but precipitation is not. This is why we employed the second KNN approach to accurately forecast the latter with 66 percent accuracy. This strategy may be applied with other time series, but one should be cautious to discern between causation and correlation in the data. When a causal link is likely, we should choose the input and output data. Should have a lot of data for training, and you strive to avoid overfitting the model to the training data. Finally, we achieved the goals we set for ourselves in this project, and the results are convincing and true in light of the situation in Kasai Oriental.

As for recommendations, we can complete this work by collecting flood data and using an extensive amount of data.

References

Ibidun O Adelekan. Flood risk management in the coastal city of lagos, nigeria. *Journal of Flood Risk Management*, 9(3):255–264, 2016.

Berhanu Fanta Alemaw, Thebeyame Ron Chaoka, and Nata Tadesse Tafesse. Modelling of nature-based solutions (nbs) for urban water management—investment and outscaling implications at basin and regional levels. *Journal of Water Resource and Protection*, 12(10):853–883, 2020.

N Anusha and B Bharathi. Flood detection and flood mapping using multi-temporal synthetic aperture radar and optical data. *The Egyptian Journal of Remote Sensing and Space Science*, 23(2):207–219, 2020.

Chay Asdak. Hydrology and watershed management. *Yogyakarta (ID): Gadjah Mada University Pr*, 1995.

Chay Asdak, Sudradjat Supian, et al. Watershed management strategies for flood mitigation: A case study of jakarta's flooding. *Weather and climate extremes*, 21:117–122, 2018.

Jens Bartholmes and Evwzio Todini. Coupling meteorological and hydrological models for flood forecasting. *Hydrology and Earth System Sciences Discussions*, 9(4):333–346, 2005.

George Bebis and Michael Georgiopoulos. Feed-forward neural networks. *IEEE Potentials*, 13(4): 27–31, 1994.

Kenneth N Brooks, Peter F Ffolliott, Hans M Gregersen, Leonard F DeBano, et al. *Hydrology and the Management of Watersheds.* Number Ed. 3. Iowa State University Press, 2003.

Marina Campolo, Alfredo Soldati, and Paolo Andreussi. Artificial neural network approach to flood forecasting in the river arno. *Hydrological Sciences Journal*, 48(3):381–398, 2003.

Ogwueleka Toochukwu Chibueze and Ogwueleka Francisca Nonyelum. Application of artificial neural networks in estimating wastewater flows. *IUP Journal of Science & Technology*, 5(3), 2009.

Léopold de Dorlodot. Contribution à la géologie du bassin du kasaï d'après des échantillons récoltés par g. gustin, reçus au musée le 15 juillet 1910.—rg: nos 1370 à 1399. *Annales de la Société géologique de Belgique*, 1920.

APJ De Roo and RJE Offermans. Lisem: a physically-based hydrological and soil erosion model for basin-scale water and sediment management. *International Association of Hydrological Sciences, Publication*, (231):399–407, 1995.

Giuliano Di Baldassarre, Alberto Montanari, Harry Lins, Demetris Koutsoyiannis, Luigia Brandimarte, and Günter Blöschl. Flood fatalities in africa: from diagnosis to mitigation. *Geophysical Research Letters*, 37(22), 2010.

Hu Du, Phillip Jones, Eva Lucas Segarra, and Carlos Fernández Bandera. Development of a rest api for obtaining site-specific historical and near-future weather data in epw format. 2018.

Sulafa Hag Elsafi. Artificial neural networks (anns) for flood forecasting at dongola station in the river nile, sudan. *Alexandria Engineering Journal*, 53(3):655–662, 2014.

Christophe Fatras, Marie Parrens, S Peña Luque, and A Al Bitar. Hydrological dynamics of the congo basin from water surfaces based on l-band microwave. *Water Resources Research*, 57 (2):e2020WR027259, 2021.

Li-Hua Feng and Jia Lu. The practical research on flood forecasting based on artificial neural networks. *Expert systems with Applications*, 37(4):2974–2977, 2010.

William W Finney III. *Effects of oceanic and atmospheric phenomena on precipitation and flooding in the Manafwa River Basin*. PhD thesis, Massachusetts Institute of Technology, 2014.

UN Office for the Coordination of Humanitarian Affairs (UN OCHA). Dr congo – thousands hit by floods in haut-uélé and tshopo. http://floodlist.com/africa/dr-congo-floods-hautuele-tshopo-november-2019, 18 NOVEMBER, 2019 par FLOODLIST NEWS IN AFRICA, NEWS).

André Nzamonga Gamo, Amédée Kundana Gbatea, Trésor Mbombo Limbaya, Adèle Chimanuka Mwinja, Gédéon Bongo Ngiala, Colette Masengo Ashande, Ruphin Djoza Djolu, et al. Knowledge on the environmental disaster occurrence in the democratic republic of the congo: The case of flooding and bushfire in businga territory, nord ubangi. *Asian Journal of Geographical Research*, pages 1–15, 2019.

Masahiko Haraguchi and Upmanu Lall. Flood risks and impacts future research questions and implication to private investment decision-making for supply chain networks. *Background paper prepared for the global assessment report on disaster risk reduction*, 2013.

Masahiko Haraguchi and Upmanu Lall. Flood risks and impacts: A case study of thailand's floods in 2011 and research questions for supply chain decision making. *International Journal of Disaster Risk Reduction*, 14:256–272, 2015.

James J Hunt, Kiem-Phong Vo, and Walter F Tichy. Delta algorithms: An empirical analysis. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 7(2):192–214, 1998.

Jean Jacques, Marc-André Lagrange, Thierry Vircoulon, and Enrica Picco. Aide humanitaire et corruption. *Diplomatie*, (107):37–41, 2021.

Anil K Jain, Jianchang Mao, and K Moidin Mohiuddin. Artificial neural networks: A tutorial. *Computer*, 29(3):31–44, 1996.

Karsten Jasper, Joachim Gurtz, and Herbert Lang. Advanced flood forecasting in alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model. *Journal of hydrology*, 267(1-2):40–52, 2002.

Colm J Jordan, Stephen Grebby, Tom Dijkstra, Claire Dashwood, and Francesca Cigna. Risk information services for disaster risk management (drm) in the caribbean: operational documentation. 2015.

A Khan, Khan M Ayaz, A Said, and Z Ali. Khan h. *Ahmad N. and Garstang*, 1950.

Aboulghasem Kazemi Nia Korrani, Kamy Sepehrnoori, Mojdeh Delshad, et al. Significance of geochemistry in alkaline/surfactant/polymer (asp) flooding. In *SPE Improved Oil Recovery Conference*. Society of Petroleum Engineers, 2016.

Official languages French. Democratic republic of the congo. 1960.

Abdul Sahib A Lateef, Max Fernandez-Alonso, Luc Tack, and Damien Delvaux. Geological constraints on urban sustainability, kinshasa city, democratic republic of congo. *Environmental Geosciences*, 17(1):17–35, 2010.

DRC Government Abdul Sahib A Lateef, Unicef, Max Fernandez-Alonso, Luc Tack, and Damien Delvaux. Democratic republic of the congo and republic of congo:flood. *Information Bulletin*, 4(1):1–3, 2020.

Moshe Leshno, Vladimir Ya Lin, Allan Pinkus, and Shimon Schocken. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural networks*, 6(6):861–867, 1993.

Lukas Löschner, Mathew Herrnegger, Benjamin Apperl, Tobias Senoner, Walter Seher, and Hans Peter Nachtnebel. Flood risk, climate change and settlement development: a micro-scale assessment of austrian municipalities. *Regional Environmental Change*, 17(2):311–322, 2017.

Lindsey McEwen, Alison Stokes, Kate Crowley, and Carolyn Roberts. Using role-play for expert science communication with professional stakeholders in flood risk management. *Journal of Geography in Higher Education*, 38(2):277–300, 2014.

Richard Mind'je, Lanhai Li, Amobichukwu Chukwudi Amanambu, Lamek Nahayo, Jean Baptiste Nsengiyumva, Aboubakar Gasirabo, and Mapendo Mindje. Flood susceptibility modeling and hazard perception in rwanda. *International journal of disaster risk reduction*, 38:101211, 2019.

Subhadra Mishra, Debahuti Mishra, and Gour Hari Santra. Applications of machine learning techniques in agricultural crop production: a review paper. *Indian J. Sci. Technol*, 9(38):1–14, 2016.

H Abd-El Monsef. A mitigation strategy for reducing flood risk to highways in arid regions: a case study of the el-quseir–qena highway in egypt. *Journal of Flood Risk Management*, 11: S158–S172, 2018.

Toochukwu Chibueze Ogwueleka and Francisca Nonyelum Ogwueleka. Feed-forward neural networks for precipitation and river level prediction. *Adv. Natl. Appl. Sci*, 3:350–356, 2009.

George A Olah, Stephen J Kuhn, and Sylvia H Flood. Aromatic substitution. viii. 1 mechanism of the nitronium tetrafluoroborate nitration of alkylbenzenes in tetramethylene sulfone solution. remarks on certain aspects of electrophilic aromatic substitution2. *Journal of the American Chemical Society*, 83(22):4571–4580, 1961.

George A Olah, Stephen J Kuhn, Sylvia H Flood, and John C Evans. Aromatic substitution. xiii. 1a comparison of nitric acid and mixed acid nitration of alkylbenzenes and benzene with nitronium salt nitrations. *Journal of the American Chemical Society*, 84(19):3687–3693, 1962.

Jean-Michel Onana, Jean Louis Fobane, Elvire Hortense Biye, Eric Ngansop https://www.overleaf.com/project/608a812f616ab96f2b1c23d2Tchatchouang, and Marguérite Marie Abada Mbolo. Habitats naturels des écosystèmes du cameroun. *International Journal of Biological and Chemical Sciences*, 13(7):3247–3265, 2019.

M O'Neill, IG Sinclair, and Francis J Smith. Polynomial curve fitting when abscissas and ordinates are both subject to error. *The Computer Journal*, 12(1):52–56, 1969.

Joseph R Oppong and Tania Woodruff. *Democratic Republic of the Congo*. Infobase Publishing, 2007.

World Health Organization et al. Ebola virus disease democratic republic of congo: External situation report 52. 2019.

Khalid Oubennaceur, Karem Chokmani, Miroslav Nastev, Marion Tanguy, and Sebastien Raymond. Uncertainty analysis of a two-dimensional hydraulic model. *Water*, 10(3):272, 2018.

Philip Tetteh Padi, Giuliano Di Baldassarre, and Attilio Castellarin. Floodplain management in africa: Large scale analysis of flood data. *Physics and Chemistry of the Earth, Parts A/B/C*, 36(7-8):292–298, 2011.

Huda A Qari, Ibrahim Jomoah, and Stefano Mambretti. Flood management in highly developed areas: problems and proposed solutions. *Journal of American Science*, 10(3):10, 2014.

XS Qin and Yan Lu. Study of climate change impact on flood frequencies: a combined weather generator and hydrological modeling approach. *Journal of Hydrometeorology*, 15(3):1205–1219, 2014.

Eric Roberts, Hielke A Jelsma, and Thomas Hegna. Mesozoic sedimentary cover sequences of the congo basin in the kasai region, democratic republic of congo. In *Geology and Resource Potential of the Congo Basin*, pages 163–191. Springer, 2015.

John W Sammon. A nonlinear mapping for data structure analysis. *IEEE Transactions on computers*, 100(5):401–409, 1969.

Jochen Schanze. Flood risk management–a basic framework. In *Flood risk management: Hazards, vulnerability and mitigation measures*, pages 1–20. Springer, 2006.

Francis J Smith. An algorithm for summing orthogonal polynomial series and their derivatives with applications to curve-fitting and interpolation. *Mathematics of Computation*, 19(89): 33–36, 1965.

Jonathon Taylor, Ka man Lai, Mike Davies, David Clifton, Ian Ridley, and Phillip Biddulph. Flood management: prediction of microbial contamination in large-scale floods in urban environments. *Environment international*, 37(5):1019–1029, 2011.

AH Te Linde, P Bubeck, JEC Dekkers, H de Moel, and JCJH Aerts. Future flood risk estimates along the river rhine. *Natural Hazards and Earth System Sciences*, 11(2):459–473, 2011.

Jean Omasombo Tshonda. Kasaï-oriental.

MICHAEL JAMES Tumbare. Mitigating floods in southern africa. In *First WAFSA/WATERNET Symposium, Maputo*, 2000.

C Van Westen and Y Yifru. Caribbean handbook on risk information management, 2016.

Christelle Vancutsem, J-F Pekel, C Evrard, Francois Malaisse, and Pierre Defourny. Mapping and characterizing the vegetation types of the democratic republic of congo using spot vegetation time series. *International Journal of Applied Earth Observation and Geoinformation*, 11(1): 62–76, 2009.

M Wéry, K Maertens, S Wéry-Paskoff, and A Fain. Contribution à l'étude de l'infestation par onchocerca volvulus dans la région de lusambo (kasai oriental, zaïre). aspects parasitologique, ophtalmologique et immunologique. *Ann. Soc. belge Méd. trop*, 56(2):95–124, 1976.

PH Whitfield. Floods in future climates: a review. *Journal of Flood Risk Management*, 5(4): 336–365, 2012.

Bayya Yegnanarayana. *Artificial neural networks*. PHI Learning Pvt. Ltd., 2009.

Liliana Zaharia, Romulus Costache, Remus Prăvălie, and Gabriela Ioana-Toroimac. Mapping flood and flooding potential indices: a methodological approach to identifying areas susceptible to flood and flooding risk. case study: the prahova catchment (romania). *Frontiers of Earth Science*, 11(2):229–247, 2017.

John Zarocostas. Integrate health services for refugees into national healthcare, says un report. *BMJ: British Medical Journal (Online)*, 343, 2011.