Journal of Computer Science and Information Technology
https://jcsitech-upiyptk.org/ojs
2021 Volume 7 Issue 4 Hal: 70-81 e-ISSN: 2502-1486

Towards Food Security: the Prediction of Climatic Factors in Nigeria using Random Forest Approach

Charity Ojochogwu Egbunu[1], Matthew Tunde Ogedengbe[2], Terungwa Simon Yange[3]✉, Terlumun Gbaden[4], Malik Adeiza Rufai[5], Habibat Iye Muhammed[6]
[1,2,3,4] Department of Computer Science, Joseph Sarwuan Tarka University, Makurdi, Nigeria
[5,6] Department of Computer Science, Federal University Lokoja, Lokoja, Nigeria
lordesty2k7@ymail.com

Abstract

With the explosive growth in the world's population which has little or no corresponding rise in the food production, food insecurity has become eminent, and hence, the need to seek for opportunities to increase food production in order to cater for this population is paramount. The second goal of the Sustainable Development Goals (SDGs) (*i.e.,* ending hunger, achieving food security and improved nutrition, and promoting sustainable agriculture) set by the United Nations (UN) for the year 2030 clearly acknowledged this fact. Improving food production cannot be achieved using the obsolete conventional methods of agriculture by our farmers; hence, this study focuses on developing a model for predicting climatic conditions with a view to reducing their negative impact, and boosting the yield of crop. Temperature, wind, humidity and rainfall were considered as the effect of these factors is more devastating in Nigeria as compared to sun light which is always in abundance. We implemented random forest algorithm using Python programming language to predict the aforementioned climate parameters. The data used was gotten from the Nigerian Meteorological (NiMet) Agency, Lokoja, Kogi State between 1988 and 2018. The result shows that random forest algorithm is effective in climate prediction as the accuracy from the model based on the climatic factors considered was 94.64%. With this, farmers would be able to plan ahead to prevent the impact of the fluctuations in these climatic factors; thus, the yield of crops would be increased. This would dwarf the negative impact of food insecurity to the populace.

Keywords: Food Security, Random Forest, Prediction, Climatic Condition, Sustainable Development Goals.

1. Introduction

The International programme 'Millennium Development Goals' signed in the year 2000 which was focused on developing countries birthed the Sustainable Development Goals (SDG) from its new set of goals targeted at every nation including the developed countries to pursue a better world. The SDG second goal to end hunger and achieve food security by 2030 especially for poor and vulnerable people and ensuring food sufficiency all year round requires promoting sustainable agriculture that would influence agricultural practices and improve farmers' productivity. Investment from international cooperation to boost agricultural capacity in developing countries like Nigeria had been impactful as food production and consumption permeates every society and is fundamental to any economy [1]. Small scale farmers in Nigeria produces about 90% of agricultural products and low yield in production of some crops like maize is observed compared to the output in other countries with similar climate patterns [2].

The Nigerian Government has not relented in its effort to cut cost in food importation in the bid to ameliorate the effect of food insecurity and promote agriculture. In December 2018, the Governor of Central Bank of Nigeria (CBN), Mr. Godwin Emefiele announced that Nigeria saved $21 billion from monthly food import bill from January 2015 to October 2018[3]. Much effort has been put in place after then in the bid to promote food security and secure a good economy. Thus, harnessing more opportunities to curb food insecurity is a welcome idea. Again, small scale farming increases access to food and improves income.

Although, there are several challenges militating the sustainable food production, but increased demand for food which would span from a rapid growing population would sure have a negative effect on food availability. Projected that by 2050 Nigeria would become the third most populated country in the world [4]. The growing population in Nigeria is mounting intense pressure on the economy which is gradually drifting towards food insecurity. This is because the available resources are not sufficient and some environmental factors that should propel agricultural growth are helplessly been watched to take on their natural negative effect. Changes in climatic conditions pose high risks to agriculture and this is presently causing serious problems in the universe [5].

Climate is an important and independent factor in agriculture which has significant effect on plants' outputs in terms of quality and quantity. Crop growth,

Recived: 30-08- 2021 | Accepted: 11-09-2021 | Publication: 31-10-2021 | doi: 10.35134/jcsitech.v7i4.15

development and yield under standard conditions are usually determined by the climatic factors during the growing season. Deviations from the normal climate could decrease the efficiency of applied inputs, thus, food production could be greatly impaired [6].

The climatic factors include: temperature, humidity, wind, rainfall and light. Among these factors, the first four have more devastating effect on crops in Nigeria and attention has always been given to them. Light has always not been an issue as sunlight is in abundance in this part of the world [7]. For instance, most rural farmers depend so much on rainfall. Its absence could result to drought which could drastically reduce crop yields as usually suffered by farmers in Northern Nigeria. When rainfall is in excess, leaching, erosion and flooding are eminent and this also reduce the output of crops as experienced by farmers in the Southern Nigeria. It is therefore important to identify obstacles confronting the growth and productivity of agriculture to devise technological approaches to address them.

The accuracy of climate forecasting is of great value for different people depending on their needs and interest. Sailors, travelers, air navigators and farmers are not left out. The rapid development in science and technology has exposed different methods in which weather data is been collected with different machine learning techniques employed on the data by researchers. In spite of the data collected, the challenge of accuracy in climate prediction has remained with us. The unpredictability in changes in climate has affected productivity in agriculture such that timely decisions that could enhance production has been jeopardized with the meandering changes experienced among climatic factors.

Factors that affect crop yield can be classified as genetic and external. About 50% influence comes from the climatic factors which includes atmospheric gases, temperature, humidity, wind velocity and Light. In an earlier study carried out by some of these factors referred to as abiotic factors were considered; factors like soil type and nutrient level, temperature, humidity, and wind were found to have influence on crop yield [7]. There exists a strong relationship between climate change and agriculture and to ensure food security, collective effort towards combating its impact on agro-systems is highly required [8]. Increase in temperature, changes in the patterns of rainfall affects crop production and even leads to decline in crop production [9]. The ability of farmers to make the best decisions in terms of sales and storage lies to a great deal on their ability to predict crop yield which to a large extent depends on the climatic factors. Hence, accurate climate prediction is needed to achieve agricultural sustainability from precise planning [10].

Hence, this paper aims at developing a model that would maximize the use of meteorological data to uncover associations among climate data from history and accurately forecast climate. Climatic factors considered includes temperature, humidity, rainfall, wind. These are external environmental factors that affect the growth and development of plants that if reliable climate forecast is known ahead of time, substitute arrangements can be put in place where there would likely be deficiency for effective growth and subsequently high crop yields. The demand for accurate climate prediction is crucial for farmers and we hope to create a model that would outweighs the challenge of inaccuracy been faced by the traditional climate forecast. The remaining part of this work is organized as follows: Section 2 discussed the climatic factors and and related works on the impact of climate on Agriculture and technology applied for climate forecast. Section 3 described the materials and methods used while the results are presented in Section 4; and lastly, the conclusion is presented in Section 5.

2. Review of Related Works

The optimal use of data in finding of patterns and the use of the patterns to predict future events has put data analytics to lime light in terms of its efficiency and speed. Understanding how a machine learns from these previous data and predict the future is key and very interesting. Through better scientific understanding, more reliable climate change projections can be achieved which again can help to tackle climate change prediction problems [11]. Some countries see climate variability as a business opportunity while its life threatening to some others and even pose difficult challenges in regards to economic development as it plays a huge role on agriculture. Africa possess a high susceptibility to climate change and climate variability, a condition that stern from the interaction of multiple sensors coupled with low adaptability capacity. Relatively, accessibility to natural and socio-economic resources mutually determines the degree to which an individual, community, or region is vulnerable to climate change irrespective of climate change induced stressors. In this section we reviewed some existing works that had been done on climate prediction.

2.1 Climatic Factors

Climatic factors play a crucial role in the development of a crop In this section, four (4) climatic factors are considered: rainfall, temperature, humidity and wind. Light and other factors are not given more attention in this research due to their insignificant contribution in plant development in Nigeria. For instance, sunlight is always in abundance in Nigeria, hence the little attention given to it.

Rainfall plays an important role in the growth of crops. Its amount and regularity vary with location and affects the dominance of certain types of vegetation as well as crop growth, development and yield. Hence, its availability or scarcity, can mean a successful harvest, or diminution in yield depending on plant species.

Temperature is another important climatic factor that influences all plant growth processes such as photosynthesis, respiration, transpiration, breaking of seed dormancy seed germination, protein synthesis and translocation. At high temperatures the translocation of photosynthesis is faster so that plants tend to mature earlier. In general, plants survive within a temperature range is 0 to 50^{o}C, also, enzyme activities and the rate of most chemical reactions generally increase with the rise in temperature. Up to a certain point, there is doubling of enzymatic reaction of 10^{o}C temperature increase. But at excessively high temperatures, denaturation of enzymes and other proteins occur. Excessively low temperatures can also cause limiting effects on plant growth and development. For instance, water absorption is inhibited when the soil temperature is low because water is more viscous at low temperature and less permeable. At low temperature- the freezing point of water, there is change in the form of water from liquid to solid. The expansion of water as it solidifies in living cells causes the rupture of the cell walls.

Humidity is a climatic factor that affects the opening and closing of the stomata which regulates loss of water from the plant through transpiration as well as photosynthesis. A substantial understanding of this climatic factor is likewise important in plant propagation. Newly collected plant cutting and bare-root seedlings are protected against desiccation by enclosing them in a sealed plastic bag. The propagation chamber and plastic tent are also commonly used in propagating stem and leaf cuttings to ensure a condition with high relative humidity.

Wind is a climatic factor that serves as vector of pollens from one flower to another thus aiding in the process of pollination. It is therefore essential in the development of fruit and seed from wind – pollinated flowers as in many grasses. Moderate winds favour gas exchanges, but strong winds can cause excessive water loss through transpiration as well as lodging or toppling of plant. When transpiration rate is high, excesses of water absorption and partial or complete closure of the stomata may ensure which will restrict the diffusion of carbon dioxide into the leaves. As a result, there will be a decrease in the rate of photosynthesis growth and yield.

2.2 Impact of Climate Change

It is expected that the world population would reach 9.7 billion by 2025 and this would imply more pressure on agricultural land as the food demand which is already affected by climate change would increase. Climate change refers to some anomalies of the system caused by human activities which are ultimately leading to global warming. Nigeria is recognized as being vulnerable to climate changes and if left unchecked would cause adverse effects on livelihood in Nigeria such as livestock production, forestry, crop production, and post-harvest activities, because the rainfall regimes and patterns would be altered, floods which devastate farmlands would occur, increase in temperature and humidity which increases pest and disease would occur and other natural disasters like floods, ocean and storm surges, which would not only cause damage on Nigerians' livelihood but also cause harm to life and property. It is possible to promote and actualize the strategies for limiting and adapting to the impact of climate change in Nigeria and globally provide cost effective and sustainable collaboration between governments. Their study to assess the impacts of climate change on groundnut crop opined that crop models should capture all extreme cases identified with certain crops and the best recommendations should be suggested in the model [12]. Increase in temperature, changes in the patterns of rainfall affects crop production and even leads to decline in crop production. The climate inputs like rainfall varies between years and also locations. Water supply and temperature over long-term also varies. The effect of these variations on crop production would certainly affect home and global food security.

2.3 Review of Previous works

Different data mining approaches have been used to predict weather based on knowledge derived from studies in the history of climate change. Artificial Neural Network and Decision Tree Algorithms on meteorological data over a span of nine (9) years from Ibadan Nigeria to investigate the efficiency of data mining techniques in forecasting of rainfall, wind speed, evaporation and maximum temperature [13]. They developed a data model to train the classifier algorithms and when the performance results from the model was compared with standard performance metrics, classification rules were generated from the algorithm that gave the best result. Weather prediction programme was achieved with the neural predictive network model and the result showed that with adequate data, data mining techniques are capable of predicting weather and climate change studies. It would be interesting to know the variation in the prediction for lesser or more than 9 years of meteorological data as used here. Exploration of other data mining techniques to determine higher level of accuracy in prediction for different areas would be valuable to stakeholders who rely on such information for effective decision making. To acquire reliable information on climate prediction in order to avoid meteorological disaster applied deep learning networks in their research on the prediction of precipitation using climate big data and the experimental results showed the feasibility of their model in weather forecasting [14]. Perturbations in weather systems as a result of different atmospheric conditions in complex weather prediction models prompted the use of machine learning by Jakaria *et al.* (2018) where the Random Forest Regression was used and their study showed that leveraging weather station data to the area where forecasting is being performed is more profitable. Artificial Neural Network in predicting temperature using Python API to obtain data from multiple online meteorological databases [15].

Rainfall prediction is of great importance to prevent flooding and manage water resources, saving lives and property and securing economic activities. ANN to accurately forecast rainfall [16]. The study shows an artificial neural network model to predict rainfall in late spring and early summer for the Geum River Basin, South Korea. Stated that air data from meteorological agency were used to train and predict values up to 72 hours with low error rate [17]. This data was then used to train decision trees to evaluate input feature importance over different time prediction horizons. The number of features used to train the long – short term memory model was reduced from 25 features to 5 features, resulting in improved accuracy as measured by Mean Absolute Error (MAE). Parameter sensitivity analysis identified look-back nodes associated with the Recurrent Neural network proved to be a significant source of error if not aligned with the prediction horizon. In all, MAEs of less than 2 were calculated for predictions up to 72 hours. K–means and naïve Bayes algorithm for forecasting weather with parameters such as temperature, humidity and wind. The study also showed that the use of data mining techniques for weather prediction yields good result and could be considered as an alternative to traditional meteorological approach. It concluded that after comparison, the decision tree and k – means clustering are best suitable data mining technique for this application [18].

Stated that weather predictions are important since the help in the formulation of the first level of preparation against natural disaster which shows the difference between life and death, and also helps in decreasing the loss of resources [19]. The research made use of C4.5 random forest algorithms after the comparison with data mining techniques that are used to boost the model performance to develop a model that can predict weather. Neural Network to forecast and generate climate dataset [20]. They found that to reproduce climate of general circulation models with a seasonal cycle using neural network is challenging in contrast to the promising results of Neural Network on a model without seasonal circle. Forecast rainfall used numerical weather prediction model to generate ensemble rainfall forecast making use of post processing of raw numerical weather prediction [21]. The Bayesian joint probability approach was used to forecast individual locations and the ensemble forecast produced using their approach was more skilful in forecasting than the raw numerical weather predictions. The latitudinal and longitudinal variations categorizes rainfall in Nigeria carried out an empirical study comparing metrological data from the different locations to established that a linear relationship exists between sea surface temperature and rainfall amount [22]. The research emphasized the need to have a quantitative means of probing anticipated rainfall for the purpose of planning and policy formulation. The artificial neural network to train and predict rainfall data collected from the province of Khorasan in Northern Iran and the prediction fell within acceptable precision [23]. Predicting the weather is essential to help prepare for the best and worst climate. Accurate weather prediction has been one of the most challenging problems around the world [24].

The research compared different data mining techniques and found decision tree and k-mean clustering to have higher prediction accuracy than ANN, KNN and multilayer regression and recommended them as the alternatives to traditional meteorological approach. It is important to note that the pre-processing steps carried out on different data mining techniques can have impact in analysis hence some researchers even go as far as linking different techniques for better accuracy. Also, predictions are sometimes unique with the data that applies to specific region hence, it is advised for climate predictions not to be generalized with specific methods to create room for optimal results as it fits specific geographical regions. Uncertainty cannot be ruled out regarding regional and local scale climatic changes that would arise from signals of natural variability in improving climatic predictions on all time scales. Therefore advised that models can be coupled for days to decadal prediction or using numerical data weather prediction models for seasonal to decadal prediction. Several systems of climate condition prediction have been built to increase the rate at which humans can be aware of their atmospheric conditions [25].

Considering the changes in the earth's climate, it is important for climate scientists to find and develop models to help not just to predict the future climatic condition but a model that checks the effect associated with the events from the climate change like soil, water condition, extreme humidity *etc*. In this research, Random Forest Regression was used because it deals with mean prediction. Also, the learning level of random forest algorithm is detailed enough for this prediction because it deals with multitude of decision trees.

3. Research Method

This section provides detailed description of the study area, and materials and the methods used for this research. The paper focused on predicting suitable climate condition for farmers using Random Forest (RF) algorithm with dataset collected from Lokoja Local Government Area, Kogi State. The RF model considered five climatic condition parameters namely: humidity, maximum temperature, minimum temperature, rainfall and wind speed for future climate prediction.

3.1 Study Area

The area considered for this research is Kogi State, which falls within the Southern Guinea Savanna Belt in Nigeria. Kogi State is located in the Middle Belt (North-Central) Zone, and was created in 1991 from portions of Eastern Kwara and western Benue State of Nigeria. It is popularly called the Confluence State because of the confluence of Rivers Niger and Benue at its Capital City, Lokoja. The climate condition in Kogi State is the local steppe climate. The average temperature in Kogi State is 26.8^{0}c daily and 747mm of rain falls annually. Kogi State is a transition from the rainforest to a savanna ecological landscape. Credible enough, cultivation of root crops and grains flourishes as with other parts within the Southern Guinea Savanna Belt of Nigeria. Most imperatively, Kogi State exhibits increasing annual rainfall and mean temperature trends in tandem with a decreasing annual rain-days trend amid seasonal fluctuations [26]. Like other states in the Middle Belt region of Nigeria, Agriculture is the hallmark of the economy amidst other resources.

Major crops farmed include: yam, cassava, rice, corn, beans, sorghum and cotton. Riverine fishing is also important as quarrying, mining and other activities. With four (4) major ethnic groups: Igala, Yoruba, Ebira and Ogori. The Igala people are the main ethnic group to the eastern part, while the Ogori, Ebira and Yoruba ethnic groups are to the west of the river. Some of the daily activities have hugely affected the environmental resources in the State. These activities range from environmental pollution from heavy duty vehicles, oil spillage, deforestation, erosion, quarry to overflown drainages, that is erosion, quarry to overflown that is un-sanitized and speculation of flooding and other future environmental hazards if not duly and mandatorily addressed.

Investigation shows that local farmers in Kogi State show little or no attention to climatic changes and this has grave impact on food security [27]. Hence, it is imperative to seek notable innovations which search for the best scientific and technological practices to address climate changes, and also, improve agricultural practice. Again, it is important to identify prevailing practices among these local farmers for mainstreaming and putting forth planned climatic change intervention programmes.

3.2 Data Collection

The dataset for this research was sourced from Nigerian Meteorological Office, Lokoja, Kogi State. This dataset was average monthly climate records from 1988 to 2018. The following are the most important factors considered for climate prediction using Random Forest: rainfall, maximum temperature, minimum temperature, humidity and wind speed [28].

a. Rainfall (mm): The total precipitation within 1988 to 2018 in Kogi State.
b. Maximum Temperature (degree Celsius): The highest temperature recorded per day from 1988 to 2018 in Kogi State.
c. Minimum Temperature (degree Celsius): The lowest temperature recorded per day from 1988 to 2018 in Kogi State.
d. Humidity: The presence of water vapor in the air which sometimes slow down evaporation. It has been established that, the higher the temperature, the lower the relative humidity and hence the faster the drying rate of any material.
e. Wind speed: The rate at which air is moving in Kogi State.

3.3 Data Processing

In this research, bootstrap from bagging technique was adopted in building the random forest and splitting the dataset for training across the 42 decision trees. When building the random forest, we drew the bootstrap sample set by sampling with replacement for each decision tree and 1/3 of the original instances are left out. The idea was to repeatedly sampled the data with replacement from the original training set in order to produce multiple separate training sets. These were then used to allow "meta-learner" or "ensemble" methods to *reduce the variance of their predictions*, thus greatly improving their predictive performance. This is known as Out-of-bag data (OOB). Each of the 42 decision trees has its own OOB data set which was used for error estimation of individual tree in the forest, called as OOB error estimation. However, we trained 42 classifiers on different samples of training data. Bagging Regressor class was adopted to generate an ensemble of regressors from the 42 decision trees known as random forest regressor. A random forest regressor is a meta estimator that fits a number of classifying decision trees on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting.

3.4 Random Forest

Random forest is a popular and powerful supervised machine learning algorithm capable of performing both classification and regression tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. The more trees in a forest the more robust the prediction. Random decision forests correct the decision trees habit of over fitting to their training set [29]. In this study, the data sets considered are rainfall, maximum temperature, minimum temperature, humidity and wind speed to construct the random forest, a collection of decision trees by considering two-third of the records in the datasets. These decision trees are applied on the remaining records for accurate classification. The resultant training sets can be applied on the test data for correct prediction of crop yield based on the input attributes [30].

3.5 Proposed Model

A RF is a classifier which comprises of multiple decision tree classifiers and can be expressed as

$$\{h(x,\Theta k), k=1...)\} \tag{1}$$

where Θk are the independently and identically distributed random trees and each tree casts a unit vote for the final prediction of input x. RF uses the Gini index for determining the final prediction in each tree. However, the proposed model is an ensemble model which consists of forty-two (42) decision trees which were adopted for the prediction of climatic condition in Lokoja, Kogi State. The decision trees make use of humidity, rainfall, maximum temperature, minimum temperature and wind speed. The dataset was divided into 75% for training and 25% for testing. Fig. 1 depicts the diagrammatic representation of the proposed model. The model consists of a database, the decision trees and the aggregate prediction.

a. Database: This is the source of data and it comprised of the climatic data obtained from NiMet Agency of Kogi State from the year 1988 to 2018, with the following features: rainfall, humidity, maximum temperature, minimum temperature and wind speed.
b. Decision Trees: The proposed model comprises of 42 decision trees which make their individual prediction.
c. Aggregate Prediction: At this phase the final prediction of each decision is aggregated and voted by weighted values to generate the final predicted value.

3.6 Gini Index

Random Forest uses Gini index derived from the Classification and Regression Tree (CART) learning system to develop sets of decision trees. The Gini index is a measure of how each variable contributes to the homogeneity of the nodes and leaves in the resulting random forest. Each time a particular variable is used to split a node, the Gini coefficient for the child nodes is calculated and compared to that of the original node.

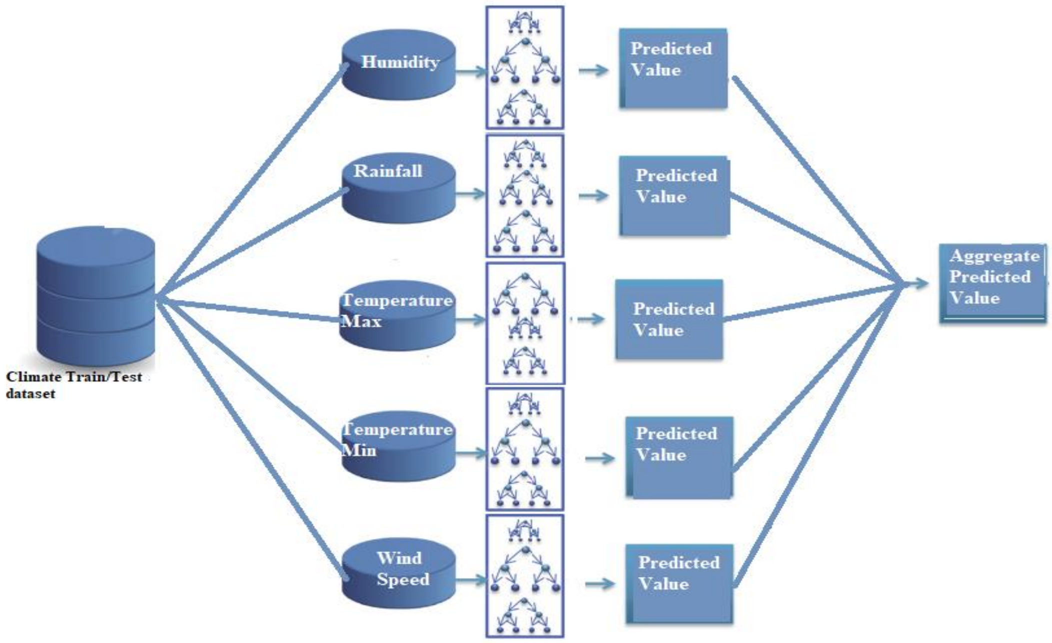

Fig. 1: The Random Forest Model for the Proposed Climatic Prediction

The Gini index for a dataset Q which contains n classes, is denoted by Gini(Q) and defined as:

$$Gini(Q) = 1 - \sum_{i=1}^{n} (f_j)^2 \qquad (2)$$

where f_j is the relative frequency of class j in Q

If the dataset Q is split into two subsets of Q_1 and Q_2 with sizes N_1 and N_2 respectively, the Gini (Q) of the split data which contains samples from n classes, would be expressed as follows:

$$Gini_{split}(Q) = \frac{N_1}{N} gini(Q_1) + \frac{N_2}{N} gini(Q_2) \qquad (3)$$

However, the feature value that provides the smallest Split Gini(Q) is selected to split the node.

3.7 RF Algorithm

Fig. 2 depicts the step-by-step illustration of the proposed RF algorithm. The following steps are listed below;

Start the proposed RF for climate prediction for the data.

Input: Number of cases (*N*), number of variables in the model (*M*)

Output: Aggregate prediction (V)

Step 1: Let m be the number of input variables used to determine the decision node of the tree such that m < M.

Step 2: Split the dataset into 75% for training and 25% for testing and choose N times replacement for all N available in training cases.

Step 3: Use the remaining cases to estimate the error of the tree by predicting each classes.

Step 4: For each node randomly choose m variables on which to base the decision at the node.

Step 5: Calculate the best split (Gini index) at each split point based on these m variables in the training set.

Step 6: Compute the prediction error

Step 7: Compute the aggregate prediction value from all the trees.

End

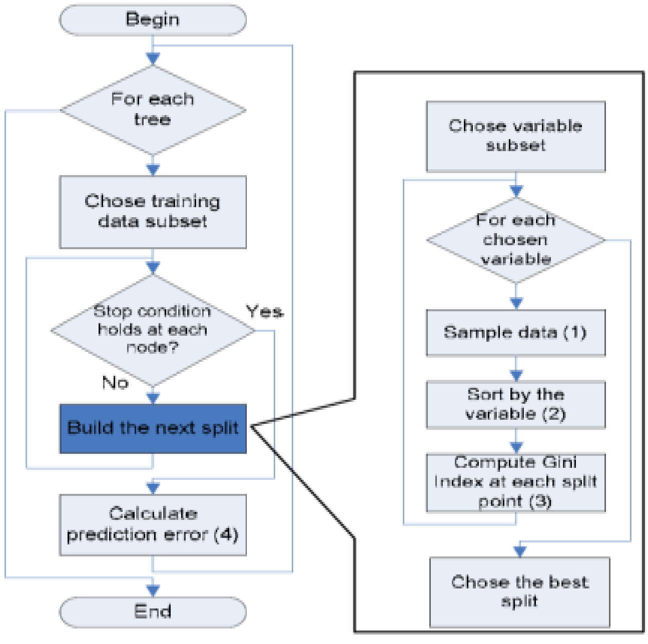

Fig. 2: Proposed RF algorithm

3.8 Metrics Measure for Random Forest Performance

Random Forest Regressor uses some standard splitting criterion to measure the quality of a split and its prediction accuracy [31]. The supported criteria are, the mean squared error (MSE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

Mean Square Error (MSE): This measures the squared average distance between the real data and the predicted data. Here, larger errors are well noted (better than MAE). We used a simple squared error *E* as our *cost function* given as:

$$E = \sum (y_i - \hat{y}_i)^2 \qquad (4)$$

The MSE is given as;

$$MSE = \sum \frac{(y_i - \hat{y}_i)^2}{n} \qquad (5)$$

Where $i = 1,\ldots n$ is each point in the dataset, *y* is the real value, $\hat{y}$ is the predicted value, and *n* is the total number of observations in the dataset.

Mean Absolute Error (MAE): This measures the absolute average distance between the real data and the predicted data, but it fails to punish large errors in prediction. We took the absolute difference between y and ŷ for each of the *n* available observations: $|y_i - \hat{y}_i|$. The total Mean Absolute Error is given as:

$$MAE = \sum \frac{|y_i - \hat{y}_i|}{n} \qquad 6$$

Root Mean Squared Error (RMSE): This is actually the square root of MSE and is given as follows:

$$RMSE = \sqrt{\sum \frac{(y_i - \hat{y}_i)^2}{n}} \qquad (7)$$

Hence, these metrics measures the average model prediction error ranging between 0 to infinity with negatively oriented scores which means the lower the evaluation value, the better is the model. However, the difference between MSE and RMSE is that RMSE has the same units as the target variable while MSE has squared units. Also, MSE is the Variance of the error value, while RMSE is Standard Deviation of errors

d. Result and Discussion

The system was implemented using python programming language. The system was used to predict five (5) different climatic parameters, which include rainfall, humidity, wind speed, maximum temperature, and minimum temperature.

4.1 Experiments

To evaluate the system, the entire dataset was split into 75/25 train-test split, that is, 75% of the data was used as the training set, and 25% of the data as the testing set. The data was then fitted into the system. First, it was the training data and then it was followed by the testing data to ascertain the accuracy of the model. Fig. 3 shows the predicted humidity for the period of 20years (*i.e.*, 1988-2018), while Fig. 4 presented the prediction of rainfall for the same period. The prediction of maximum temperature for the period under consideration (*i.e.*, 1988-2018) while Fig. 6 presented the result of the predicted minimum temperature for the same period. Lastly, Fig. 7 shows the result of the predicted wind for a period of 20years (*i.e.*, 1988-2018).

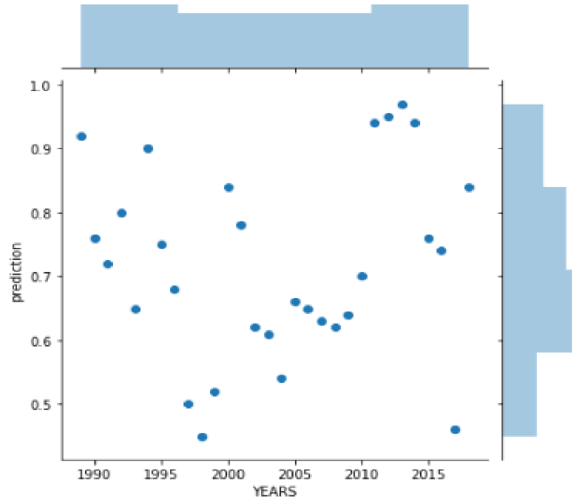

Fig. 3. Graphical Representation of Humidity

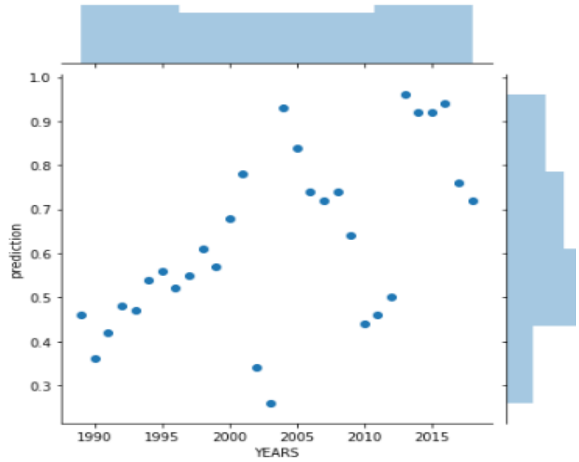

Fig. 4. Graphical Representation of Rainfall

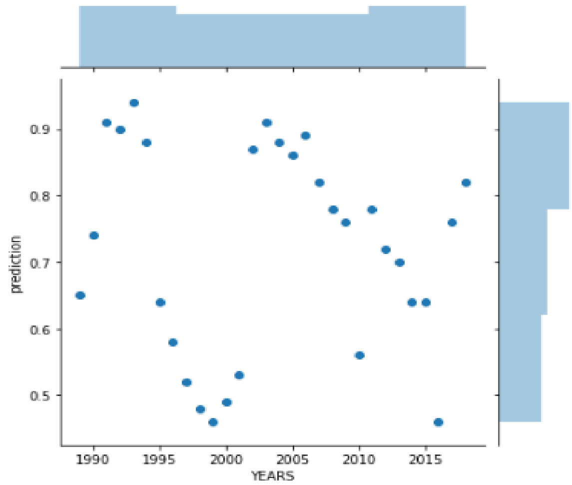

Fig. 5. Graphical Representation of Temperature

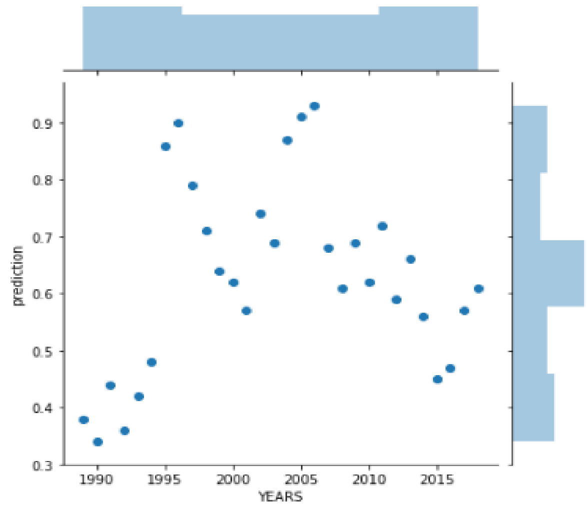

Fig. 6. Graphical Representation of Temperature

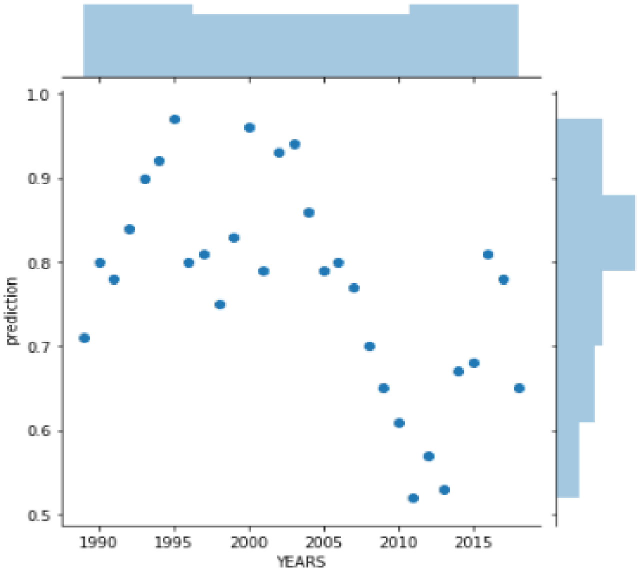

Fig. 7: Graphical Representation of Wind speed

4.2 Evaluation

During the development phase, testing was carried out on the five (5) different climatic parameters to determine their accuracies. This was achieved using the Mean Absolute Error, Mean Squared Error and Root Mean Squared Error of all the parameters in the random forest algorithm as described in Section 3.8. The summary of the evaluation results is presented in Table 1 with the average accuracy of 94.64%.

Table 1: Result Evaluation

Parameters	MAE	MSE	RMSE	Accuracy (%)
Rainfall	0.11	0.004	0.063	93.70
Humidity	0.11	0.005	0.066	93.40
Wind Speed	0.07	0.001	0.032	96.80
Max Temperature	0.15	0.008	0.070	93.00
Min Temperature	0.09	0.002	0.037	96.30

4.3 Discussion

In this research, a system for predicting climatic factors was developed. This was with a view to reduce the drop in yield that farmers experience due to the variation in climatic factors. Maximization of output is crucial, most especially now that the world's population is growing astronomically without a corresponding food production which has posed serious danger to food security. As outlined in MDG/SDG, fighting this menace is of great importance to humanity.
This paper used random forest technique in predicting rainfall, humidity, wind speed, maximum temperature and minimum temperature using the dataset collected from NiMet between 1988-2018 in Kogi State. The dataset was split into 75/25 train-test split set and the accuracy was computed using mean absolute error, mean squared error and root mean squared error for each of the five (5) climatic factors considered. The result of the evaluation is presented in Table 1.
As discussed in the review of related works, food insecurity is a threat to the existence of mankind and researchers are working tirelessly to arrest the danger it poses before it gets to a level where the world would not contain it. In view of this, many researches are ongoing in all disciplines on how to reduce the scourge of this menace. In agricultural informatics, data mining and machine learning techniques such as the Support Vector Machine [32], Artificial Neural Network [16],[32], Deep Learning [14],[33], [34], Decision Tree [35];[32], K-Means [18[, naïve Bayes [18], and many more have been used.

One of the major undoing of the existing works is that most of them did not consider all the climatic factors. For instance [15], and considered the prediction of temperature [34],[14]. considered wind. None of these works took a wholistic look at these five most prevalent climatic factors in Nigeria. But according to Howden and White (2007), climatic factors always operate and interact with each other under natural conditions for the benefit/detriment of crops; therefore, in order to mitigate the effect of climatic factors, it must be treated wholistically which has made this research to stand tall above the previous researches.

To reduce the drop in yield of crops, this research has provided a tool to predict climatic conditions to enable farmers to forecast what would happen in the next farming season(s). With this, if the climatic factors for the next farming seasons are not favourable for the crop the farmer is intending to plant, an alternative option would be explored. This would avert the future lost by the farmer; hence, food production would not be affected. Thus, food security would be guaranteed

4. Conclusion

The availability of food is fundamental in a healthy environment and to actualise that, environmental factors must be checked and closely monitored. Ending waste is crucial in the move towards efficient food production. In the bid to ensure food security and better the lives of Nigerian citizens, we realize the need for an accurate foreknowledge of weather climate to ascertain its effect on agriculture (crop production, livestock farming, fishing *etc.*) and to act promptly to avoid the negative effect of climate changes on food production.

In this paper, we used random forest algorithm for predicting climate parameters such as maximum temperature, minimum temperature, wind speed, humidity, and rainfall yearly. The data used was gotten from meteorological agency in Kogi State between 1988 and 2018. Random forest algorithm proved to be quite effective in its accuracy in climate prediction. However, just been aware of the forecast without any action is as good as being ignorant of it. We therefore recommend this system for timely weather forecasting and that the information from weather forecast be disseminated to remote villages via MDG/SDG group members through the NYSC target strategy. Farmers should be told what they should prepare to do like in the case of draught where channels for irrigation can be put in place.

Similarly, for predicted flooding times, the government should discourage farmers from investing in such communities and lend temporal lands to willing farmers at no cost for that farming year. To eradicate hunger, sound policies and sustained political commitment must thrive. Future research should consider an expert system for certain crop climatic requirements for maximum yield. It should also consider building a model that would recommend the crop that would be most favourable for a particular climate forecast year. Also, researchers should consider conducting studies that could achieve higher accuracy prediction considering more climate factors on larger datasets.

References

1. Gill, J.D.B., Reidsma, P., Giller, K. et al. (2019). Sustainable development goal 2: Improved targets and indicators for agriculture and food security. *Ambio 48, 685–698* https://doi.org/10.1007/s13280-018-1101-4
2. GYGA. 2018. Global yield gap atlas. Wageningen University and Research, University of Nebraska—Water for Food. http://www.yieldgap.org/web/guest/home. Accessed July 2020.
3. Pmnewsnigeria (2018). Nigeria has achieved food imports reduction, saved $21bn in 34 months —Emefiele. https://www.pmnewsnigeria.com/2018/12/03/nigeria-has-achieved-food-import-reduction-saved-21bn-in-34-months-emefiele/. Accessed in July, 2020.
4. Van Ittersum, M.K., L.G. Van Bussel, J. Wolf, P. Grassini, J. Van Wart, N. Guilpart, L. Claessens, H. de Groot (2016). Can sub-Saharan Africa feed itself? Proceedings of the National Academy of Sciences 113: 14964–14969.
5. Manyong, V.M. 2005. Agriculture in Nigeria: Identifying opportunities for increased commercialization and investment, IITA.
6. Neenu, S., Biswas, A. K. and Rao, A. S. (2013). Impact of Climatic Factors on Crop Production - A Review. *Agricultural Reviews*, 34 (2): 97-106.
7. Yange, S. T., Egbunu, C. O., Onyekwere, O. and Foga, K. A. (2020). Prediction of Agro Products Sales using Regression Algorithm. American Journal of Data Mining and Knowledge Discovery (AJDMKD), 5(1): 11-19
8. Arora, N.K. (2009). Impact of climate change on agriculture production and its sustainable solutions. Environmental Sustainability 2, 95–96. https://doi.org/10.1007/s42398-019-00078-w
9. Mall R. K., Gupta K. I., Sonkar G. (2017). Effect of Climate Change on Agricultural Crops. *Current Developments in Biotechnology and Bioengineering*. *Pp 23-46.* Elsevier.
10. Priyanka, T., Soni, P., Malathy, C. (2018). Agricultural Crop Yield Prediction Using Artificial Intelligence and Satellite Imagery. *Eurasian Journal of Analytical Chemistry, 13(SP), emEJAC181194.*
11. Dessai S, Hulme M, Lempert RJ, Pielke R. (2009). Climate prediction: a limit to adaptation?
12. Kumar U., Singh P. and Boote K. J. (2012) Effect of Climate Change Factors on Processes of Crop Growth and Development and Yield of Groundnut (Arachis hypogaea L.), 116: 41-69.
13. Olaiya, F. and Adeyemo, A. B. (2012). Application of Data Mining Techniques in Weather Prediction and Climate Change Studies. I.J. Information Engineering and Electronic Business MECS, 1, 51-59. DOI: 10.5815/ijieeb.2012.01.07
14. Du J, Liu Y, Liu Z. Study of Precipitation Forecast Based on Deep Belief Networks. *Algorithms*. 2018; 11(9):132. https://doi.org/10.3390/a11090132
15. Abrahamsen E., Brastein O M., Lie B. (2018). Machine Learning in Python for Weather Forecast based on Freely Available Weather Data. Conference: Exergy Analysis for Combined Heat and Power (CHP) Plants. DOI: 10.3384/ecp18153169
16. Lee, J., Kim, G.C., (2018). Application of artificial neural network to rainfall forecasting in geum river basin, korea. *Water*, *10*, 1448.
17. Brian S. F., Taylor, G., Gharabaghi, B. and The, J. (2018). Forecasting Air Quality Time Series Using Deep Learning. Journal of the Air and Waste Management Association, 1-21.
18. Biradar, P., Ansari, S., Paradkar, Y. and Lohiya, S. (2017). Weather Prediction Using Data Mining. International Journal of Engineering Development and Research, 5(2): 211-214.
19. Karthick, S., Arun, C., Malathi, D., (2018). Forecasting of monthly temp-variation using radom forest. *International Journal of Pure and Applied Mathematics* Vol 118. 255 -262
20. Scher, S. and Messori, G. (2019). Weather and climate forecasting with neural networks: using general circulation models (GCMs) with different complexity as a study ground, Geosci. Model Dev., 12, 2797–2809, https://doi.org/10.5194/gmd-12-2797-2019.
21. Robertson, D. E., Shrestha, D. L., and Wang, Q. J. (2013). Post-processing rainfall forecasts from numerical weather prediction models for short-term streamflow forecasting, Hydrol. Earth Syst. Sci., 17, 3587–3603, https://doi.org/10.5194/hess-17-3587-2013.
22. Omogbai B. E. (2010). Empirical Prediction of Seasonal Rainfall in Nigeria. *J Hum Ecol, 32(1): 23-27.* http://environmentportal.in/files/Rainfall%20in%20Nigeria.pdf
23. Fallah-Ghalhary G. A., Mousavi-Baygi M and Habibi-Nokhandan M. (2009). Seasonal Rainfall Forecasting Using Artificial Neural Network. *Journal of Applied Sciences, 9: 1098-1105.* 10.3923/jas.2009.1098.1105
24. Choi C., Kim J., Kim D., Bae Choi C., Kim J., Kim J., Kim D., Bae Y. and Kim H. S. (2018). Development of Heavy Rain Damage Prediction Model Using Machine Learning Based on Big Data. https://doi.org/10.1155/2018/5024930
25. Hurrell, J. & Co-Authors (2010). "Decadal Climate Prediction: Opportunities and Challenges" in *Proceedings of OceanObs'09:*

Sustained Ocean Observations and Information for Society (Vol. 2),

26. Venice, Italy, 21-25 September 2009, Hall, J., Harrison, D.E. & Stammer, D., Eds., ESA Publication WPP-306, doi:10.5270/OceanObs09.cwp.45
27. Atedhor, G. O. (2015) Perceptions of Local Farmers' Vulnerability to Climate Change in Kogi State, Benin. International Journal of Agricultural Economics and Extension Services, 4 (1): 1-15
28. Ajadi, D. A., and Sanusi, Y. K. (2013) Effect of Relative Humidity on Oven Temperature of Locally Design Solar Carbinet Dryer. *Global Journals Inc. (USA)* ISSN: 2249-4626. 13**(1)**.
29. Narasimhamurthy, V., and Kumar, P. (2017). Rice Crop Yield Forecasting Using Random Forest Algorithm. *International Journal for Research in Applied Science & Engineering Technology* (IJRASET) *ISSN: 2321-9653;* 5**(X)**. *Available at www.ijraset.com*
30. Gandhi, N., Armstrong, L.J. and Petkar, O. (2016). Rice Crop Yield Prediction in India using Artificial Neural Network. *International Conference on 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), Chennai, India.*
31. Dobbali, M.K. (2019). Basic Metrics to Understand Regression Models in Plain English. Available at https://towardsdatascience.com/metrics-to-understand-regression-models-in-plain-english-part-1-c902b2f4156f. Accessed on 14/07/20
32. Yange, S. T., Egbunu, C. O., Rufai, M. A., Onyekwere, O., Abdulrahman, A. A. and Abdulkadri, A. (2020). Using Prescriptive Analytics for the Determination of Optimal Crop Yield. International Journal of Data Science and Analysis (IJDSA), 6(3): 72-82.
33. Khan, A., Zamee, A., Jamal, T., (2018). Deep belief network based features generation and regression for wind power. Cornell University arXiv:1807.11682 **[cs.LG]**
34. Agarwal, A., Shrimali, M., Jain, A., Sirohi, A., (2019). Forecasting using machine learning. *International Journal of Recent Technology and Engineering 2*(6c).
35. Naing W. Y. N and Htike Z. Z. (2015). Forecasting of Monthly Temperature Variations using Random Forests. ARPN Journal of Engineering and Applied Sciences. 10 (21). Pp 10109 - 10112
36. In: Adger WN, Lorenzoni I, O'Brien K (eds) Adapting to climate change: thresholds, values,
37. governance. Cambridge University Press, Cambridge (in press).
38. Howden, S. M. and White, D. H. (2007). Climate and its Effects on Crop Productivity and *Management - Soils, Plant Growth and Crop Production, 1: 1-9.*
39. Idowu, A.A., Ayoola, S.O., Opele, A.I., (2011). *I*mpact of climate changein nigeria. *Iranica Journal of Energy and Environment*. 2(2) Pp 145 - 152
40. Jakaria A. H. M., Hossain, M. M., and Rahman M. A. (2018). Smart Weather Forecasting Using Machine Learning: A Case Study in Tennessee. In Proceedings of ACM Mid-Southeast conference (Mid-Southeast'18). ACM, New York, NY, USA, 4 pages.
41. Odjugo A. O. (2009). Quantifying the Cost of Climate Change Impact in Nigeria: Emphasis on Wind and Rainstorms. Journal of human ecology. 28(2):93-101 DOI: 10.1080/09709274.2009.11906223