

Detection of Rainfall Probability Distribution from Atmospheric Parameters Using the Neural Network Modelling Algorithm

Okoro, N. O.^{1,*} and Ogbonna R. C.²

¹Department of Industrial Physics, David Umahi Federal University of Health Sciences, Uburu, Ebonyi State, Nigeria.

²Department of Computer Science and Mathematics, Evangel University Akaeze, Ebonyi State Nigeria.

*Corresponding author email: bathonjoku@gmail.com

ORCID ID: <https://orcid.org/0000-0001-9222-7525>

Abstract

The Earth's surface receives different amounts of rainfall during one season in different places and that too during another season throughout the year. It is a priority in Nigeria to detect and understand spatiotemporal rainfall distribution probabilities to anticipate suitable strategies for environmental hazard monitoring and control, agricultural productivity and water management control. In Nigeria, obtaining data and information regarding rainfall distribution probability has been strongly dependent on the use of rainfall forecasting instruments by Nigeria Meteorological Agency (NiMET). This has made information dissemination about rainfall so difficult to reach out to many people especially the ruler dwellers. Therefore, the objective of this study is to use neural network modelling algorithms and atmospheric parameters (temperature, pressure, relative humidity, wind speed and total cloud cover) to detect rainfall probability distributions (that is, when to and when not to expect rainfall) at any given location. A programme was designed using a neural network modelling algorithm to permit the input of atmospheric parameter(s) for the detection of rainfall probability distributions within the area of study for different periods of time. The results obtained showed that rainfall distribution tendencies on different days in the year of study corresponded with the observed rainfall data distribution for the same year of study within the same location. This clearly indicated that the rainfall probability distribution can possibly be detected from the temperature, pressure, relative humidity, wind speed and total cloud cover using a neural network modelling algorithm. This may have implications for agricultural productivity and sustainability when monitoring rainfall tendencies during the actual planting period. It may also be useful in monitoring rainfall activities in any given location to reduce hazards.

Keywords: Atmospheric parameters, modelling algorithm, neural network, rainfall distribution probability.

1. Introduction

Over the years, the rainfall distribution in different parts of Nigeria has generally been perturbed due to the convective movement of air masses in the Earth's atmosphere and has consistently caused climatic change. This phenomenon has become a major interest of atmospheric study due to its inherent contributions both positively and negatively to life since the existence of the universe. Rainfall occurs when condensed water vapour masses in the atmosphere fall back to the earth in the form of raindrops resulting from complex interactions that occur in the Earth-atmosphere environment (Okoro *et al.*, 2022). It is consequently being driven by the condensation of air masses through upwards movement in the atmosphere. To a greater extent, rainfall in Nigeria has been highly important for agricultural growth, food security, water management, economic development and sustainability.

However, no rainfall or extreme rainfall for a long period of time at a particular location may consequently cause many hazards of various types and variable magnitudes (Okoro *et al.*, 2022). Dyer and Tyson (1977) reported that rainfall distribution has periodic tendencies in various regions of Africa. Obviously, rainfall is the most critical and key parameter of interest in both the atmospheric and hydrological cycles.

In Nigeria, a lack of accurate information about rainfall distribution probability has only been managed by the Nigeria Meteorological Agency (NiMET). However, little attention has been given to how the interaction of some atmospheric parameters in the atmosphere can be used to predict when there will be rainfall and when there will be no rainfall using a neural network modelling algorithm rather than focusing on Nimes alone. This has made it very challenging for many people, especially rural farmers, who may not have the necessary tools to access NiMET predictions to find it difficult to obtain precise information about rainfall probability distributions. This, consequently, has caused many problems for farmers, who are in most cases being deceived to commence planting their crops on their farms due to early rainfall, which may eventually cease to continue for a long time. This may eventually cause the crops they plant to die since the rainfall may not continue in order to sustain the growth of crops. In a related development, there may be excessive rainfall most of the time that may eventually lead to environmental hazards such as flooding, and erosion due to a lack of credible information about the probability distribution of rainfall in a given location. Therefore, detecting when there will be rainfall or when there will be no rainfall has become so important that it will help farmers to appropriately know the right time to plant their crops to increase productivity. Additionally, this study will provide accurate information for policy makers in water management and hazard control. Therefore, the objective of this study is to determine the possibility of detecting from atmospheric parameters when there will be rainfall and when there will be no rainfall in any given location using a neural network model algorithm.

Neural network algorithms are a set of codes in neural networks that can be combined to compare relationships in a dataset by possibly mimicking the way the human brain works. Therefore, neural network algorithms resemble human neurons in the brain. In most cases, modelling, forecasting, etc., may be carried out with the help of neural network algorithm combinations since they are capable of manipulating and utilizing a historical dataset for forecasting purposes. Because the combined neural network algorithm mimics the way the human brain works, it has a better approach and accuracy level than the contemporary statistical and mathematical models.

2. Related Works

2.1. Theoretical Review

According to Dyer and Tyson (1977), rainfall is central to the management of major economic activities such as hydroelectric power generation. Zhang *et al.* (2013) reported that rainfall is regarded as the fundamental element for replenishing water bodies and managing water resources worldwide. In a separate work, (Manzanas, *et al.*, 2014; Rezaeianzadeh *et al.*, 2014; Atiah *et al.*, 2019) stated that rainfall globally is an indisputable factor in agricultural planning and management in any given region, but excessive rainfall can lead to a waterlogged environment that can subsequently result in flooding, while drought may result from insufficient rainfall distribution. Ankita and Geeta (2015) reported that rainfall plays a very important role in agriculture and human life, and thus, it is essential for irrigation.

Rainfall is an important factor for agricultural activities and management in any given region because moderate rainfall leads to an increase in agricultural produce, while excessive and inadequate rainfall can trigger floods and induce drought. According to Neitsch *et al.* (2011), floods cause economic losses and surface runoff, which are important driving forces of the transport of sediment, nitrate, phosphorus and other chemical compounds in watersheds. Drought can have widespread effects on many water-related sectors, such as agriculture, irrigation, and energy, in regions with low or variable rainfall (Olusola and Samuel, 2022).

Generally, rainfall exhibits spatial and temporal distribution patterns in Nigeria and invariably has undoubtedly affected agricultural production and water supply management and has equally resulted in environmental hazards and degradation on many occasions. Ankita and Geeta (2015) stated that changes in many atmospheric parameters, such as temperature, humidity, wind speed, and pressure, can possibly affect rainfall distribution over a given area and within a specific interval of time. According to Okoro and Onugwu (2023), rainfall distribution can be classified into four major patterns, namely, no rainfall distribution pattern, low rainfall distribution pattern, moderate rainfall distribution pattern and heavy rainfall distribution pattern, depending on the magnitude of rain droplets on the surface of the Earth during a particular period. Many researchers have used various mathematical and statistical models to predict rainfall in Nigeria. Ighile *et al.* (2022) predicted flood areas due to excessive rainfall in Nigeria using logistic regression (LR) and artificial neural networks (ANNs). They observed that the ANN model produced a higher performance and prediction rate than did the LR model (76.4% and 62.5%) respectively. Lawal *et al.* (2021) used three machine learning algorithms, namely, decision tree, logistic regression and support machine vector methods, with thirty-three historical rainfall datasets to predict flood cases in Kebbi State, Nigeria.

3. Methodology

3.1 Study Area

Two (2) data coordinate locations in Nigeria that covered the spatial location of two states were studied for one year. These regions included Cross River (Calabar), with a latitude of 4.98° N and a longitude of 8.35° E and Minna (Kaduna), with a latitude of 10.60° E and a longitude of 7.45° E to 45° E, respectively.

3.2 Data Collection, Computation and Analysis

The daily rainfall and atmospheric parameter datasets used in this work consist of a satellite (Reanalysis-Interim) dataset obtained from the European Centre for Medium-Range Weather Forecasting (ECMWF) and are jointly managed by the National Center for Atmospheric Research data support section (Okoro *et al.*, 2022). The spatial resolution of the ECWFM datasets is 0.125° latitude by 0.125° longitude (spectral truncation T159), and they extend from 1979 through the present. The Panoply and Ferret software packages, sourced from the Goddard Institute for Space Studies (GISS) and the National Oceanic and Atmospheric Administration (NOAA), respectively, were used equally for data conversion (Okoro *et al.*, 2022). After data collection, data downscaling, conversion and analysis were performed with Ferret-PMEL software, the Excel spreadsheet package, MATLAB R2007b and the Minitab 19 statistical software package. The Panoply and Ferret software packages, which were sourced from the *Goddard Institute for Space Studies* (GISS) and the *National Oceanic and Atmospheric Administration* (NOAA), respectively, were used equally for data conversion and analysis.

3.3 Neural Network Architecture

The neural network (NN) model, which is a model that has roots in artificial intelligence, uses a series of algorithms to recognize underlying relationships in a set of data through a process that mimics the same way the human brain works. According to Ankita and Geeta (2015), a neural network model can be used for prediction because it has the capability of examining and determining the historical data used for prediction. It was adopted for algorithm design and rainfall probability detection and analysis. This model was preferred because it has the ability to vary input such that the network can possibly generate the best result or output without redesigning the output processes all over and, as such, has a better accuracy level than the mathematical and statistical models. Neural network modelling algorithms were created, trained, validated and tested for the detection (prediction) of daily rainfall probability distributions using longitude, latitude and altitude coordinate values of atmospheric parameters as input variables. The model was adopted in this work because of its inherent advantages over other models in its mode of processing data, accuracy level and ability to compare the relationships between inputs without necessary redesigning the output over and over.

The neural network model has three different (architectures) layers: an input layer, a hidden layer and an output layer.

Input layer: This layer contains the initial dataset used for neural network modelling. It is usually represented as a vector X .

Hidden layer: This layer in most cases is often referred to as the activation node or the weight, and it is usually represented as W or ϑ . It forms the intermediate layer between the input and output layers and serves as an anchor point for all computations in neural network modelling.

Output layer: This layer produces an output (result) for a given input(s) in neural network modelling. It is generally represented as y .

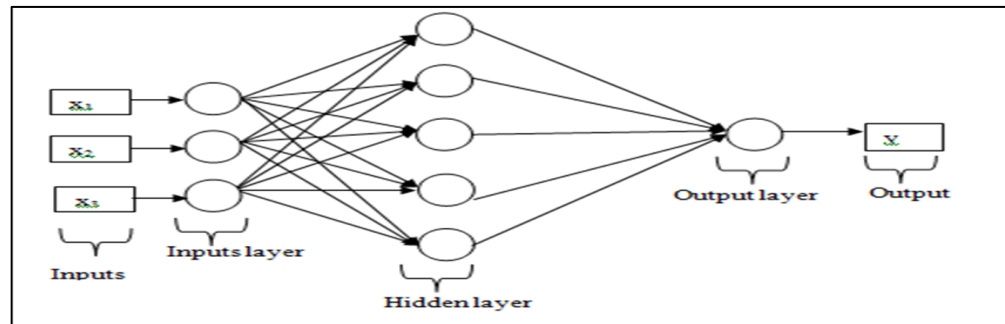


Figure 1: Schematic diagram of the neural network architecture.

Figure 1 shows a schematic diagram of the neural network architecture adopted during modelling in this work. It has 3 input layers, 5 hidden layers and 1 output layer connected with nodes.

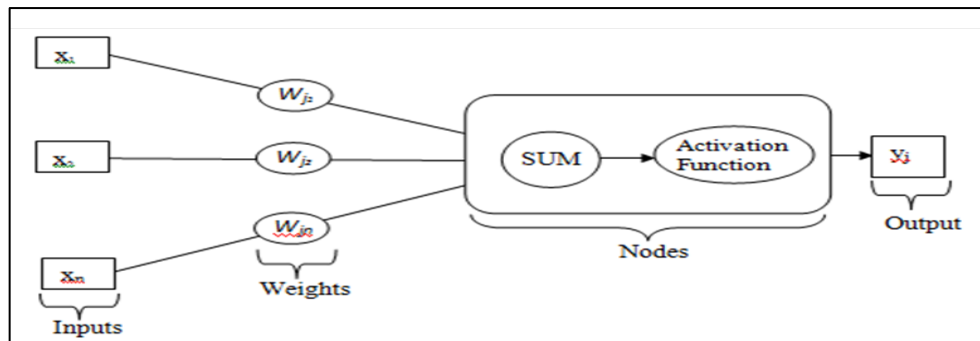


Figure 2: Neural network activation function.

As shown in Figure 2 above, every node is connected to another node from the next layer, and each connection has a particular weight. The weight can be visualized as the impact each layer exerts on the node from the next layer. By implication, each node in the input layer only acts as a buffer for delivering the input data x_i ($i = 1, 2, \dots, n$) to the nodes within the hidden layer. Therefore, the equation of the activation function of the i^{th} neuron in the hidden layer is given as:

$$h_i = f(\varphi_i) = f\left(\sum_{i=1}^n w_{ji}x_i\right) \tag{1}$$

where h_i represents the i^{th} hidden layer neuron, $f(\varphi_i)$ represents the activation function, w_{ji} represents the weight in the ji^{th} entry and x_i represents the j^{th} input value. Therefore, every processing node j in the hidden layer adds up

its input dataset x_i after weighting with the strengths of the particular connections w_{ji} from the input layer and computes the resultant output y_j as a function of the sum of the hidden layer neurons. This implies that:

$$y_i = f\left(\sum_{i=1}^n w_{ji} x_i h_i\right) \tag{2}$$

where y_i is the j^{th} output value from the output layer.

3.4 Data arrangement

The datasets used in this study were first arranged using the MATLAB in the format required for NN training. In this case, the datasets for NN training were arranged in an array of matrix formats that mapped the inputs to the output as follows (Figure 3):

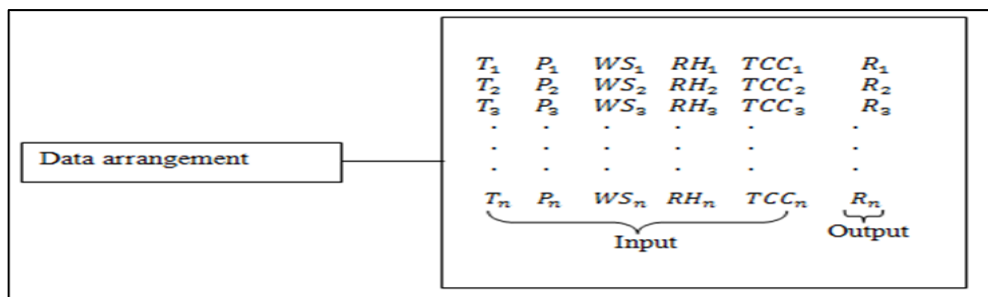


Figure 3: Neural network activation function.

Here n is an integer. For T , n ranges from $-2.82\text{ }^\circ\text{C}$ – $43.37\text{ }^\circ\text{C}$, for P , n ranges from 96272 Pa – 97428 Pa ; for WS , n ranges from 0.1172 ms^{-1} – 3.1558 ms^{-1} ; for RH , n ranges from 44% – 107.7% ; for TCC , n ranges from 0 – 1 ; and for rainfall, n ranges from 0.1 mm – 120 mm .

3.5 Designing Algorithms

The designed NN algorithms implemented in MATLAB were used to generate vectors of binaries for detected rainfall (predictions) as well as the observed (initial) rainfall and rainfall probability distribution from the observed atmospheric parameter datasets for a test year in the study location.

3.6 Rainfall detection and analysis

For detection, the algorithm was designed to load the network that was chosen as the suitable network during training, validation and testing of the data with the NN model. It was equally designed to read the inputs, which are the atmospheric parameters (T , P , WS , RH , and TCC), simulate the inputs for the possible prediction of rainfall probability distributions across each location and to create vector binaries, choose an assumed boundary that was used for inferences and compute the percentage of agreements or disagreements that were used for possible analysis. To detect the rainfall probability distribution for all the locations from the chosen network using the designed algorithm, the network was first loaded. Second, inputs (i.e., T , P , WS , RH and TCC) that were used for the prediction of rainfall were provided. Third, the network was used to simulate the output (predicted rainfall). Fourth, assumed boundaries for rainfall were created to enable us to detect when there will be rainfall and when there will be no rainfall. Based on the assumed boundary, vectors of binaries (0s or 1s) were generated, and from the assumed boundary, any value of predicted rainfall ranging from 0.1 - 0.4 mm was considered “0” and represented no rainfall, and any value of predicted rainfall ranging from 0.5 mm and above was considered “1” and represented rainfall, as displayed below. Afterward, the vector of binaries created for predicted rainfall as well as the actual observations and the percentage of agreements and disagreements were computed, compared and analysed. The detected value is a function of the sum of the hidden layer neurons and is given as:

$$\text{Detected (Predicted) rainfall} = f\left(\sum_{i=1}^{12} (T, AP, WS, RH, TCC) \times H_i\right), \quad (3)$$

where T, AP, WS, RH and TCC are the inputs and H_i represents the number of hidden layer neurons.

Algorithm for the detection of rainfall probabilities

```
net=load('net number'); %load the network
% inputs=[]; %provide the inputs
inputs=dlmread('1month.txt', '\t', 1, 0); %provide the inputs
targets=inputs(6, :);
inputs=inputs(1:5, :);
predicted_rainfall=sim(net.net, inputs); %simulate
assumed_boundary=0.5;
f=find(predicted_rainfall>=assumed_boundary);
ff=find(predicted_rainfall<assumed_boundary);
predicted_rainfall(f)=1;
predicted_rainfall(ff)=0;
[predicted_rainfall' targets'
```

4. Results and Discussion

Figure 3 represents the rainfall probability distribution for the observed (initial) dataset for Calabar in 2013, while Figure 4 represents the rainfall probability distribution for the detected (predicted) rainfall from atmospheric parameters using a neural network modelling algorithm for Calabar in 2013. Figure 5 represents the rainfall probability distribution for the observed (initial) dataset for Kaduna in 2013, and Figure 6 represents the rainfall probability distribution for the detected (predicted) rainfall from atmospheric parameters using a neural network modelling algorithm for Kaduna in 2013. Figure 3 shows the observed (initial) rainfall probability distribution. The plot was characterized by a series of “no rainfall” or “rainfall”, as represented by small circles and vertical lines on the plot beginning from the first day of the year to the last day of the year in the study area. There was no evidence of rainfall from approximately days 1–9, 13–27, or other days, as clearly indicated in the plot. However, there was early commencement of more frequent rainfall distributions from around day 30, which were consistent until approximately day 350 of 2013. The plot of the observed rainfall probability distribution at Calabar suggested that rainfall commenced earlier and maintained a steady pattern almost throughout the year 2013.

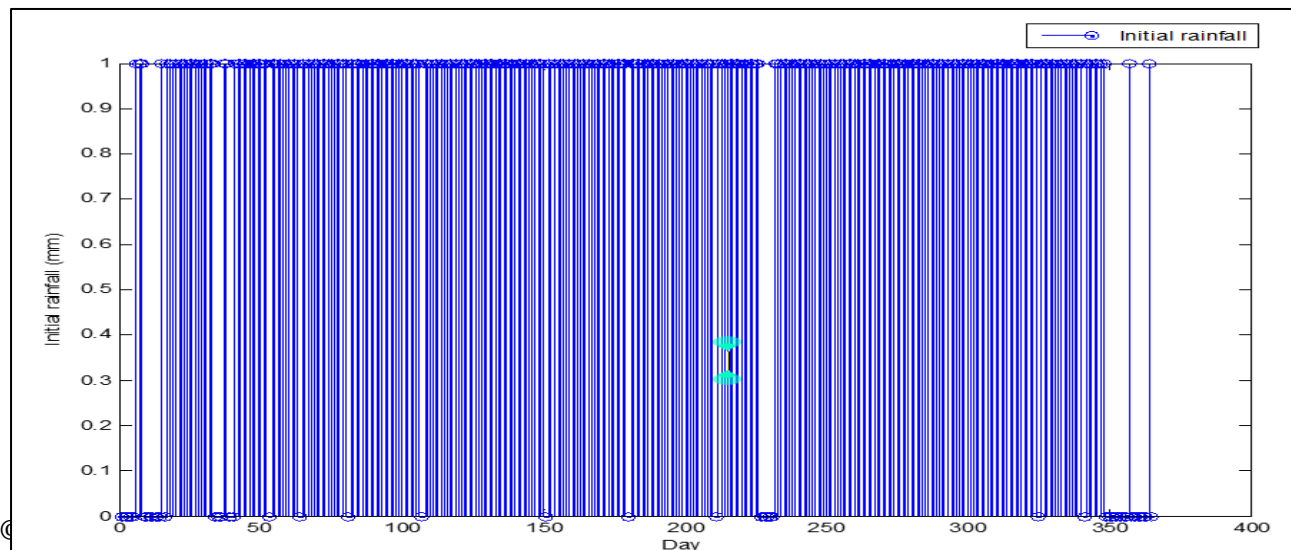


Figure 3: Probability distribution of observed (initial) rainfall with days of the year at Calabar in 2013.

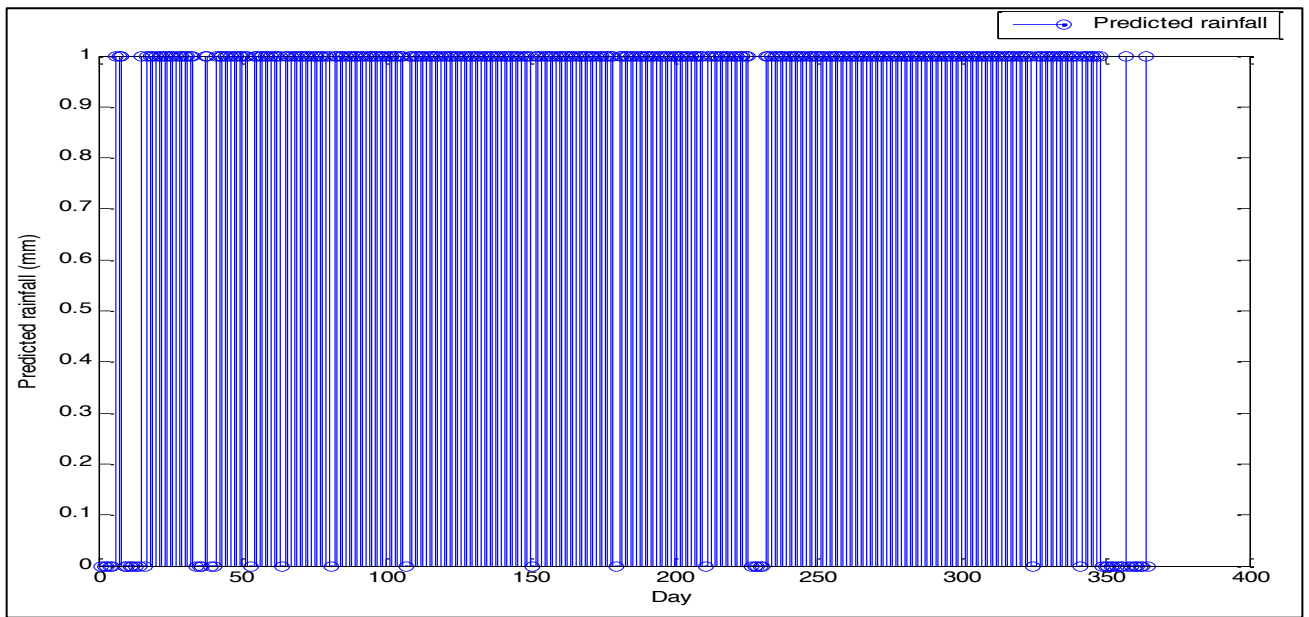


Figure 4: Probability distribution of detected (predicted) rainfall with days of the year at Calabar in 2013.

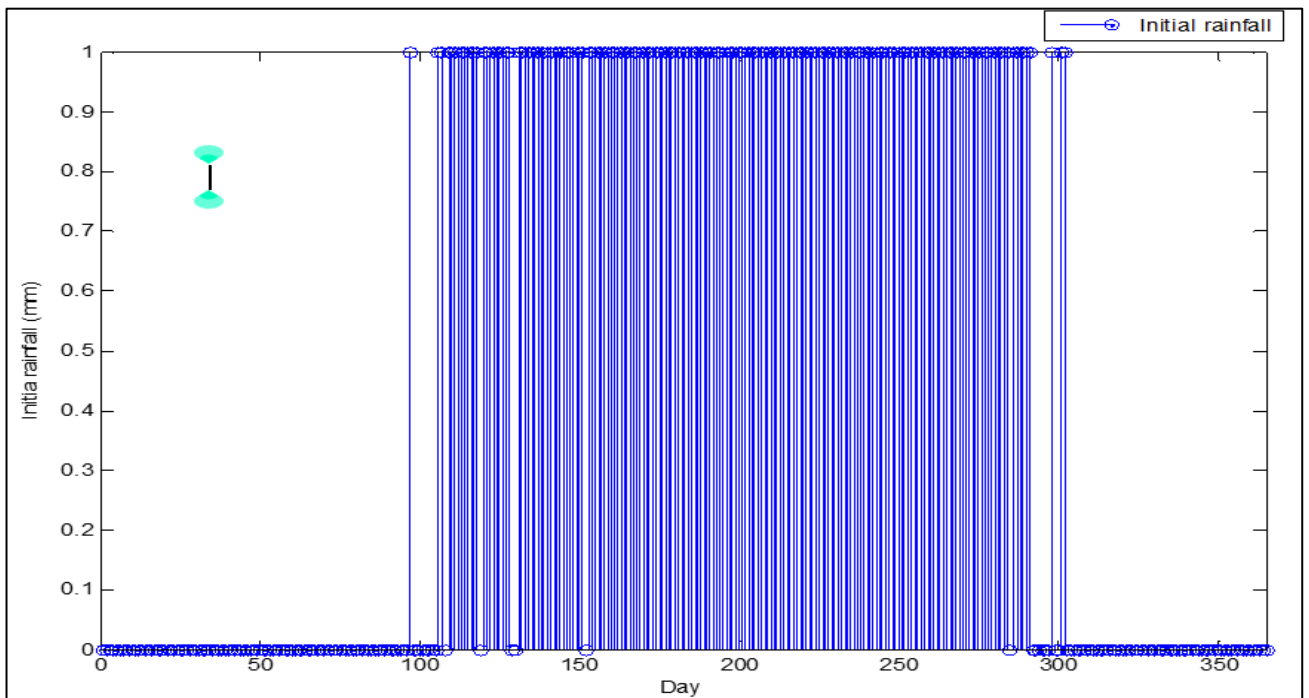


Figure 5: Probability distribution of observed (initial) rainfall with days of the year at Kaduna in 2013.

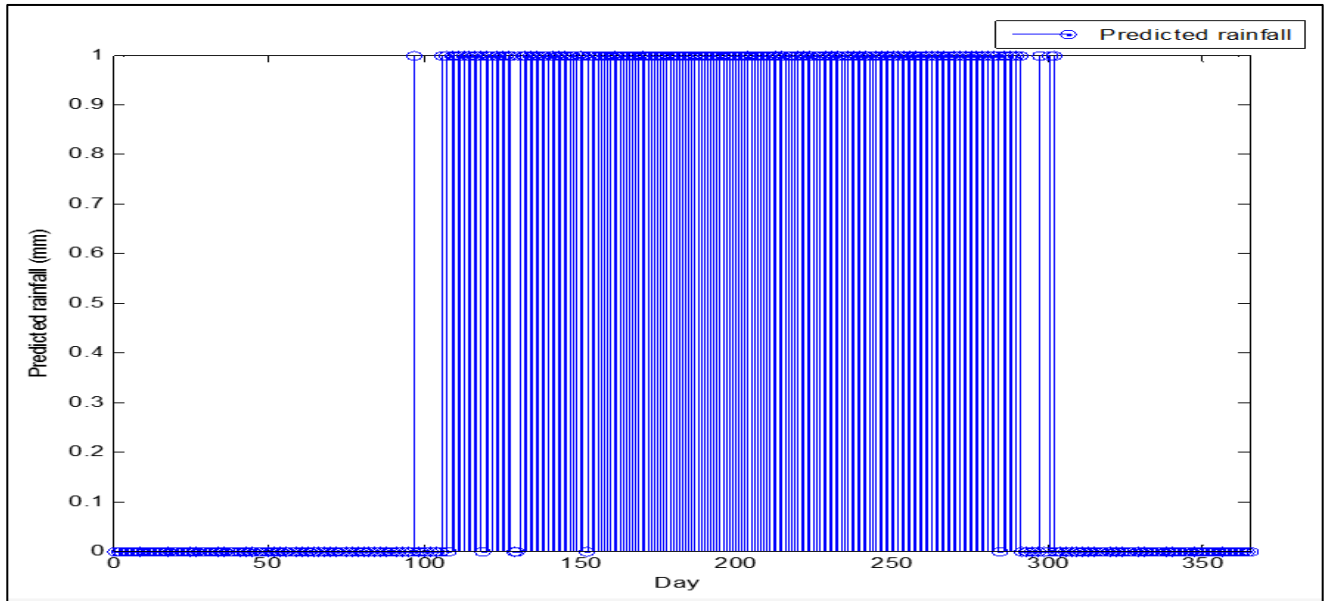


Figure 6: Probability distribution of detected (predicted) rainfall with days of the year at Kaduna in 2013.

Figure 4 shows the detected (predicted) rainfall from atmospheric parameters using the neural network model algorithm, which clearly indicates that there was no rainfall between days 1 and 9. Similarly, between days 13 and 27, there was no rainfall, and the same pattern was observed on some other days within the year of the study. The model detected an early commencement of rainfall in the study area from day 30 to approximately day 350. The observations in this scenario imperatively correspond with the results obtained in Figure 3 and thus suggest that rainfall can actually be detected from atmospheric parameters using a neural network modelling algorithm. In Kaduna (Figure 5), it was detected from atmospheric parameters using a neural network modelling algorithm that from days 1 to 100 (January to April) and from days 300 to 365 (October to December), there were no rainfall probability distributions throughout the days in 2013; thus corresponded with, Figure 6 of the same year showing rainfall probability distribution for the initial dataset at Kaduna. However, extending towards days 151 to 299 (July to September), the rainfall probability distributions detected were regular and more prevalent for the days of the study in the same year, which further confirmed the prediction accuracy of the neural network model implemented algorithm employed in this work.

5. Conclusion

Rainfall is of inherent importance to human existence, and when accurate information about it is obtained, it can apparently result in an increase in agricultural productivity that will in turn lead to sustainable economic development. On the other hand, if accurate information about rainfall in a given location is not obtained, then environmental hazards are most likely inevitable. In Nigeria, to gather accurate information about rainfall probability distributions, mostly by the ruler dwellers, is often limited. This is because information regarding rainfall is solely managed by the NiMET, and therefore limits its timely dissemination; most especially to the ruler dweller

that may not have the necessary tools to access their records, which may either affect them positively or negatively. As a result, there is an emergent need for an alternative method that can also be used to accurately obtain rainfall probability distributions. In this work, a neural network modelling algorithm was used to detect rainfall probability distributions from atmospheric parameters (temperature, pressure, relative humidity, wind speed and total cloud cover). The neural network modelling algorithms used were created, trained, validated and tested for the detection (prediction) of daily rainfall probability distributions using longitude, latitude and altitude coordinate values of atmospheric parameters as input variables. The results showed that the detected (predicted) rainfall probability distribution pattern from atmospheric parameters using the neural network modelling algorithm has approximately the same trend as the trend of the probability distribution pattern of observed (initial) rainfall in all the days in the study areas. This means that if the data of atmospheric parameters considered in this work are readily available, then the rainfall probability distribution (i.e., when there will be rainfall and when there will be no rainfall) for any given coordinate location may be determined using a neural network modelling algorithm. Finally, the interactions among temperature, pressure, relative humidity, wind speed and total cloud cover can be efficiently used to account for the probability distribution of rainfall via a neural network modelling algorithm. This study can be adopted by policy makers to help in monitoring rainfall tendencies to enhance agricultural productivity and reduce environmental hazards in Nigeria and beyond. It can equally be employed for water resources management and control in our society.

Acknowledgement

The authors wish to appreciate those whom their material (s) was used in this work for making them available.

References

- Ankita, S and Geeta, N. (2015). Rainfall Prediction Using Neural Network. *International Journal of Computer Science Trends and Technology (IJCST)*, 3, 3.
- Atiah, W.A., Amekudzi, L.K. Quansah, E., Preko, K. (2019). The spatiotemporal variability of rainfall over the agro-ecological zones of Ghana, *Atmos. Clim. Sci.* 9 (03), 527.
- Dyer, T. G and Tyson, P. H. (1977). Estimating Above and Below-Normal Rainfall Periods over South Africa 1972-2000. *Journal of Applied Meteorology*, 16,145-147.
- Ighile, E.H. Shirakawa, H and Tanikawa, H. (2022). A study on the application of GIS and machine learning to predict flood areas in Nigeria, *Sustainability* 14 (9), 5039 .
- Lawal, Z.K., Yassin, H and Zakari, R.Y. (2021). Flood Prediction using Machine Learning Models: A Case Study of Kebbi State Nigeria, in: 2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), IEEE, pp. 1–6.
- Manzanas, R., Amekudzi, L., Preko, K., Herrera, V and Gutiérrez, J.M. (2014) Precipitation Variability and Trends in Ghana: an Intercomparison of Observational and Reanalysis Products. *Clim. Change.* 124 (4),pp 805–819.
- Neitsch, S. L., Arnold, J. G., Kiniry, J. R and Williams, J.R. (2011). Soil and Water Assessment tool Theoretical Documentation Version 2009, Technical Report, Texas Water Resources Institute.
- Okoro, N. O and Onugwu, A. C. (2023). Study of Rainfall Distribution Pattern Changes and their Relationships with Atmospheric Parameters at Enugu and Kano States using a Regression Model. *Nigerian Journal of Physics (NJP)*, 32:3.
- Okoro, N. O., Agbo, G. A., Ofoma, J. N and Mbamara, C. E. (2022). Testing Neural Network (NN) Algorithms Implemented in MATLAB for the Study of Rainfall Distribution Patterns at Different Locations in Nigeria. *World Journal of Advanced Research and Reviews*, 15(03), 027–032
- Olusola, S. O. and Samuel, T. O. (2022). Machine Learning Models for Prediction of Rainfall over Nigeria. *Scientific African.* 16:e01246
- Rezaeianzadeh, M., Tabari, H., Yazdi, A .A ., Isik, S and. Kalin, L. (2014). Flood Flow Forecasting using ANN, ANFIS and Regression Models, *Neural Comput. Appl.* 25 (1), 25–37.
- Zhang, W., Wang, J., Fan, J., Gao, D and Ju, H. (2013). Effects of Rainfall on Microbial Water Quality on Qingdao no. 1 Bathing beach, China, *Mar. Pollut. Bull.* 66 (1–2),185–190 .