IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 25, Issue 6, Ser. 3 (Nov. – Dec. 2023), PP 08-18
www.iosrjournals.org

Forecasting Seasonal Rainfall using a Feed Forward Neural Network with Back-Propagation: A Case of Zambia

Lillian Mzyece[1]
University of Zambia
School of Natural Sciences
Department Computer Science
Lusaka. Zambia
Email: mamzyece@gmail.com

Mayumbo Nyirenda PhD[2]
University of Zambia
School of Natural Sciences
Department Computer Science
Lusaka. Zambia
Email: mnyirenda@unza.zm

Prof Jackson Phiri PhD[3]
University of Zambia
School of Natural Sciences
Department Computer Science
Lusaka. Zambia
Email: jackson.phiri@unza.zm

Abstract
Weather forecasting is a scientific estimation of future weather conditions that applies science and technology to predict the state of the atmosphere for a given time and location. Rainfall forecasting is one of the most difficult and important weather forecasting tasks for national meteorology services worldwide. The current statistical method used for rainfall forecasting in Zambia relies on a relationship between sea surface temperatures and station rainfall observations and ignores other factors that affect rainfall. Such statistical models have limitations for long range forecasts as correlations change over time and lose their significance. Furthermore, the station point rainfall observations used in the current method also have high spatial variability that increases the uncertainty of the forecasts. Some steps in the current method are subjective and based on visual inspection. This research proposes the use of artificial neural networks to forecast seasonal rainfall in Zambia and also incorporates other factors that influence rainfall into the model, which will improve the forecast accuracy. Three artificial neural network techniques were analyzed in order to select a suitable one for generating seasonal rainfall forecasts in Zambia. Of the three, the Feed-Forward Back-Propagation (FFBP) neural network was selected due to its principles of simplicity and effectiveness. Monthly data from 1961 to 2010 for six weather parameters was used to train the FFBP model. It was tested to forecast seasonal rainfall from 2011 to 2016, one season at a time.

Experimental results showed that the ANN technique can perform better with acceptable precision than the current fore-casting method with a higher confidence level of 0.997 compared to 0.309 for the currently used method.

Keywords: *Artificial Intelligence; Artificial Neural Networks; Feed-Forward Back-Propagation*

Date of Submission: 19-12-2023 Date of Acceptance: 29-12-2023

I. Introduction

Rainfall is a natural climatic phenomenon resulting from atmospheric and oceanic circulation systems that cause an amount of rain to fall at a place during a given period [1]. It is one of the weather parameters whose prediction is challenging and demanding. A weather forecast is a scientific estimate of future weather conditions for a specified location and given time [2]. Rainfall forecasting is the most complicated and challenging operational task carried out by National Meteorological and Hydrological Services (NMHS) globally [3], [4]. It is challenging, demanding and complex due to the chaotic nature of the atmosphere and various dynamic environmental factors, that culminate in instabilities. Rainfall is a highly non-linear parameter that has significant implications for highly climate sensitive sectors such as the Agriculture, Water Resource Management, Energy, Disaster Risk Reduction and Tourism sectors, which form the bedrock of economic activities of Zambia[1], [2]. It affects human life and livelihood the most in countries where the majority of the population relies on rain fed agriculture [3].

Seasonal rainfall forecast is an outlook of the expected rainfall performance for a given rainy season. In the Southern African Development Community (SADC) region, Zambia inclusive, it is mostly generated using Simple Linear Regression Models (SLRM): where an empirical statistical forecasting model is developed to forecast seasonal rainfall. This model describes a linear relationship between an independent and a dependent variable. The statistical models based on regression analysis and eye ball inspection are discussed further in [5], [6]. The current approach for forecasting seasonal rainfall in Zambia assumes a direct correlation between Sea Surface Temperatures (SST) and station rainfall observations [7]. Yet atmospheric systems are not governed by only these two variables. As a result, this approach neglects the complex interactions among various atmospheric and oceanic processes that influence rainfall [8]. Therefore, a more comprehensive and robust method that potentially incorporates multiple predictors and accounts for the nonlinear dynamics of the climate system to forecast seasonal rainfall is needed. Additionally, the direct linear correlation between observed rainfall data and SST assumption is no longer reliable due to changing climate which introduces further uncertainties [8], [9]. Statistical rainfall forecasting methods have inherent limitations over long range rainfall forecasts, because correlations are assumed to remain constant for the duration of the forecast, yet they usually change with time and slowly lose significance [10] thus reducing the rainfall forecast accuracy. One of the challenges in forecasting seasonal rainfall is the high spatial variability of station point rainfall observations. This variability leads to more inaccuracy and uncertainty in the forecast, which lowers the forecast accuracy. Furthermore, some stages in the current seasonal rainfall forecasting process require eye ball inspection through expert knowledge. This makes the current forecasting method subjective and not easy to pass on through an educational process [6]. Therefore, the current seasonal rainfall forecasting method used in Zambia leaves room for improvement, and the incorporation of modern artificial intelligence techniques for high efficacy.

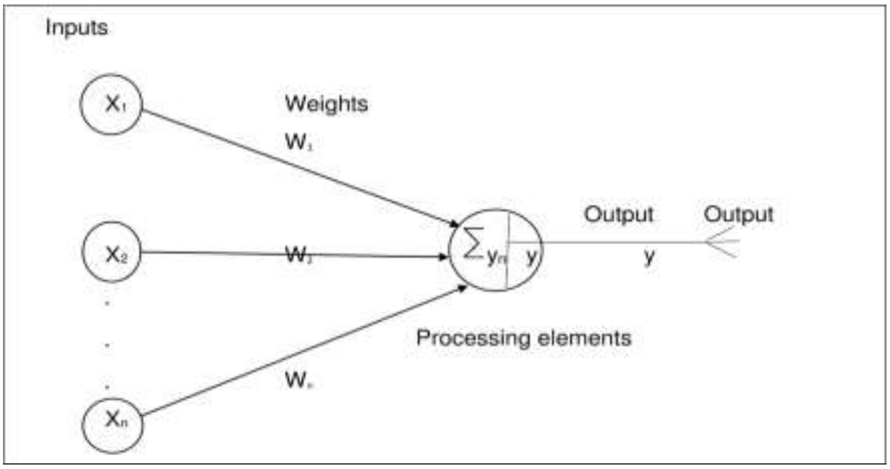
This research contributes to generating enhanced accurate seasonal rainfall forecasts in Zambia by incorporating other factors that influence rainfall. This will help stakeholders who rely on evidence based planning and decision making in all climate sensitive economic sectors especially for the farming season, where majority of the population rely on rain fed agriculture. Long range rainfall forecasts provide vital information that help nations and stakeholders to prepare for and reduce the potential negative impacts of climate change, hence accurate forecast play an integral part in this preparedness. Some stages in the current seasonal rainfall forecasting process depend on expert knowledge through visual analysis. This makes it difficult to transfer this knowledge through formal education processes [11]. The study also enhances objectivity by eliminating eyeball inspection which is inherent in the current process and consequently alleviates challenges related to knowledge transfer as discussed in [6].

This research is in response to increased interest and demand by stakeholders for accurate seasonal rainfall forecasts in Zambia. The call to diversify Zambia's economy to agriculture in the face of climate change has contributed to the demand for accurate seasonal rainfall forecasts in Zambia. The aim of this research is to forecast seasonal rainfall with enhanced accuracy using Artificial Neural Networks (ANN) and incorporating other data that influence rainfall.

II. Artificial Neural Networks (ANN)

Artificial Neural Networks are a branch of Soft Computing (SC) under Artificial Intelligence (AI) that tries to build intelligent systems based on approximate reasoning, exploiting tolerance for imprecision, uncertainty and partial truth [12]. Artificial Intelligence itself is a field of computer science that aims to create systems that can perform tasks that normally require human intelligence. Instead of exact algorithms, Artificial Intelligence relies on heuristic methods that can handle uncertainty and complexity with minimal human guidance [13]. ANNs are computational models that mimic the structure and function of the human brain [14]. An ANN consists of a large number of simple processing elements that are layered and interconnected with each other and consist of a set of connected cells called neurons that process information and learn from data similar to how the human brain does it. An artificial neuron is a mathematical function conceived to simulate the behaviour of a biological neuron in the brain. It is typically used to make up an artificial neural network [15], [16]. Given in Fig 1: below is a graphical presentation of an artificial neuron.

Figure 1: Artificial Neuron



Adapted from [16]

ANNs are non-linear data driven self-adaptive models where learning by example replaces programming in solving problems. Before Artificial Neural Networks can be used to perform any desired task, they must be trained to do so. This process is used to determine the arc weights [17]. The result of weighted and added inputs is transformed by a transfer function into output. The artificial neuron given above is a simple processing unit, where the net input is elucidated as

$$y_i n = x_1 w_1 + x_2 w_2 + \ldots + x_n w_n = \sum_{i=1}^{n} x_1 w_1, \tag{1}$$

where i represents the processing element. An activation function is applied to it to calculate the output. Weight represents the strength of the synapse connecting input and output neurons. A positive weight corresponds to an excitatory synapse while a negative weight corresponds to an inhibitory synapse. The true power and advantage of neural networks lies in their ability to:

1. Represent both linear and non-linear relationships
2. Be trained with known examples of a problem before being tested for their 'inference' capability on unknown instances of a problem: thus learn and adapt
3. Map input patterns to their associated output patterns
4. Generalize, thus predict new outcomes from past trends
5. Be robust and fault tolerant, by recalling full patterns from incomplete, partial or noisy patterns
6. Process information in parallel, at high speed and in a distributed manner [18]

Artificial Neural Networks are well suited for problems whose solutions require knowledge that is difficult to specify but for which there is enough data. These capabilities make ANN well suited for weather forecasting [19]

III. Related Technologies

Artificial Neural Networks (ANNs) are soft computing techniques that can approximate functions and recognize patterns. They have been applied to various domains for weather forecasting and have demonstrated their effectiveness and superiority [20]. The skill of rainfall forecasts produced using ANN methods in India [21], Thailand [22] and the west mountainous region of Iran [23] showed to be very accurate. ANNs are powerful rainfall forecasting tools that provide aggressive models over the existing rainfall forecast methods used in Zambia because an ANN can learn from data and adjust its parameters to improve its performance [6], [1] A. Chaturvedi, applied ANN methods using back propagation for rainfall prediction with minimal Mean Square Error (MSE) in Delhi – India. In another study, the performance of ANN methods with statistical linear regression methods for forecasting rainfall where compared in Thailand. The results showed that ANN methods had higher accuracy and adaptability than linear regression methods because they could learn from the data and adjust their parameters accordingly [22]. According to G. Shrivastava et al, the best method for forecasting rainfall in the long term is to use artificial neural networks (ANNs) that rely on a deterministic approach and a back-propagation algorithm. They claim that this technique can achieve high accuracy and reliability in recognizing and predicting patterns of precipitation [8].

According to a 2012 study, ANNs are better at predicting rainfall than conventional methods based on statistics and numerical analysis. The study compared different ANN models with various input variables and found that they performed well in terms of accuracy and reliability [8]. ANNs has been usedas a suitable technique for the long-term climate variabilitylike seasonal rainfall forecast due to the fact that learning isaccomplished through training [24].ANN forecasting models predict weather by intelligently learning from a large amount of historical data. Mathematical or statistical weather models are very precise in calculation, but not in forecasting, because they cannot handle the irregular and complex patterns of atmospheric data that are not based on a function or a formula [25]. The use of ANN algorithms in forecasting rainfall becomes an attractive approach because ANNs can handle complex, nonlinear and flexible data. ANNs can also learn from data and create models without requiring any prior knowledge [8]. The ability of ANN to cope with non-linearity, speed of computation, learning capacity and prediction accuracy makes it a superior model of forecasting rainfall [26].

IV. Materials and Methods

This section includes an analysis of the current seasonal rainfall forecasting method and experimentation of forecasting seasonal rainfall using ANN.

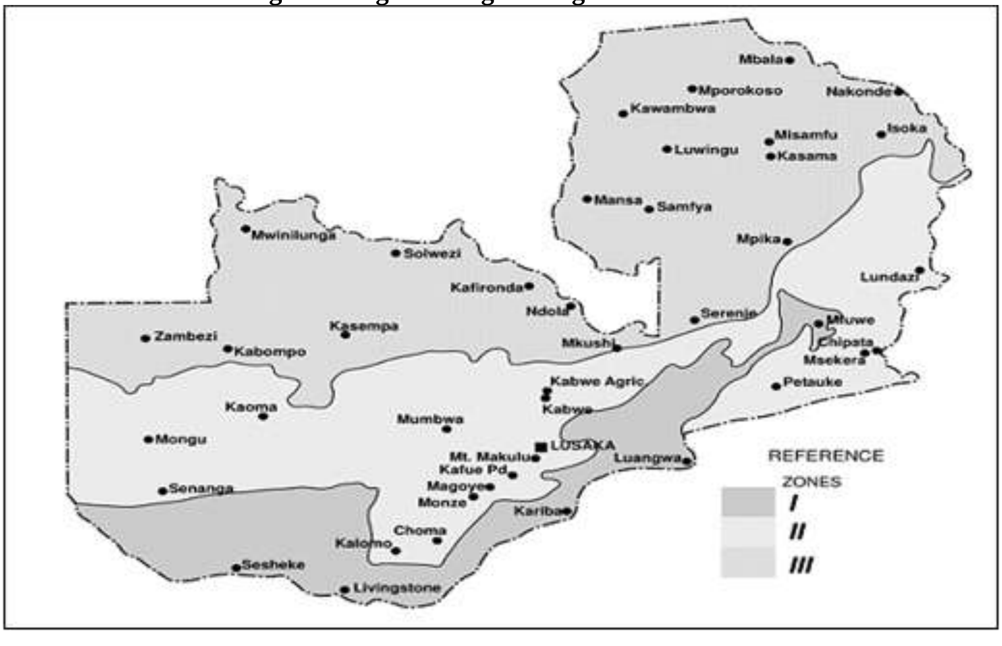
Current Method

To analyze the performance of ANNs we had to compare them to the current statistical method. We thus first used the current method and applied it to the Sea Surface Temperatures (SSTs) and station monthly rainfall observation data before applying ANNs to the data. In the following subsections, we explain the current method which begins with the identification of Homogeneous Rainfall Regions.

Identification of Homogeneous Rainfall Regions

To forecast seasonal rainfall, the country is divided into regions with similar rainfall patterns. Zambia has three homogeneous rainfall regions as shown in Fig. 2 below, [27].

Figure 2: Agro-ecological Regions of Zambia



Adapted from [27]

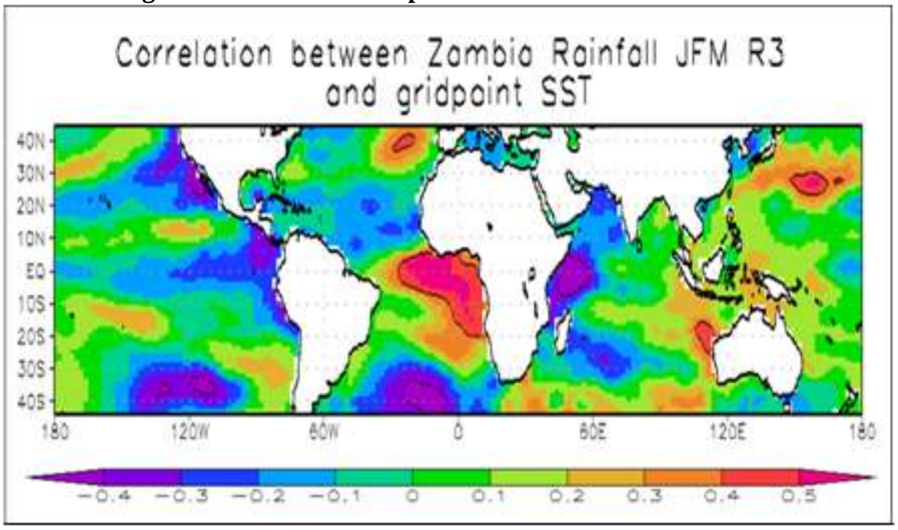
Downloading Sea surface temperatures (SST)

The next step after delineating the zones with homogeneous rainfall patterns in Zambia, is downloading Sea Surface Temperatures. The SSTs are downloaded from the International Research Institute for Climate and Society (IRI) website [28].

Correlation of station rainfall with Sea Surface Temperature data

The next step involves establishing a correlation between the SSTs and the station rainfall data. The data is then standardised to ensure consistency and avoid non-linear issues due to unequal variance of monthly rainfall. The correlation between Sea Surface Temperature and station rainfall is calculated for a specific period i.e. January; February; March (JFM) in a region.Given in Fig. 3, is a correlation map between rainfall for JFM in region 3 and SSTs.

Figure 3: Correlation Map - JFM and SST for R3



Adapted from [6]

From the given correlation map, basins that correlate ±0.3 or higher are selected through visual inspection by estimating the coordinates for each basin one at a time. This is done by zooming in one at a time to estimate the basin coordinates. Selected basin coordinates are recorded and assigned an identity to each basin.

Regression Analysis

Finally, a Simple Linear Regression Model (SLRM) is used to create an empirical statistical forecasting model. This model assumes a linear relationship between two variables. The independent variable X is the SST basins and the dependent variable Y is the rainfall. A regression analysis output for JFM R1 is shown in Tables 1 - 5 below.

Table 1: Regression Analysis Output

Table 1: Regression Analysis Output

Dependent Variable	RAINFALL
N	57
Multiple R	0.309
Squared Multiple R	0.095
Adjusted Squared Multiple R	0.079
Standard Error of Estimate	0.960

Table 2: Regression Coefficients = B (X1X)-1 X1Y

Effect	Coefficient	Std. Error	Std. Coefficient	Tolerance	t	p-value
CONSTANT	0.005	0.127	0.000	-	0.043	0.966
NEPAC	0.309	0.128	0.309	1.000	2.408	0.019

Table 3: Analysis of Variance

Source	SS	df	Mean Squares	F-Ratio	p-value
Regression	5.341	1	5.341	5.799	0.019
Residual	50.659	55	0.921		

Table 4

Durbin-Waston D Statistic	1.948
First Order Auto correl ation	0.038

Table 5: Information Criteria

AIC	161.037
AIC (Corrected)	161.490
Schwarz's BIC	167.166

Statistical Techniques

P-value indicates whether the relationship between independent and dependent variables is significant, below 0.05 is significant while over 0.05 is not significant. F-ratio is a comparison ratio of the model to its error. R-value indicates model confidence level. Basins meeting requirements based on a low P-value of 0.059 or lower, higher F-Ratio and an R-value of 0.3 or higher are used to calculate rainfall. To calculate the forecast rainfall value for any region, in the given three months i.e. JFM, the following formula is used.

Y = mx + c where; Y is expected rainfall, M is coefficient for the chosen basin, X is the SST for the month of July and C is the constant. If more than one basin meets the requirements, forecast rainfall value is found using the same equation but the constant is only added once [7], such as;

Y = c + (coefficient*SST) + (coefficient*SST) + (coefficient*SST).

Forecast classification

Once a forecast value has been calculated, it is classified into one of three categories as either Below Normal, Normal or Above Normal. The current method for classifying seasonal rainfall using Microsoft Excel is as follows. First, the rainfall data from 1981 to the latest year are standardized and sorted in ascending order according to years. Then, the data is classified by dividing it into three groups with an equal number of records, called terciles. If there are extra records, they are added to the middle group. Thus all rainfall data is classified as either Below Normal, Normal or Above Normal. Once a rainfall forecast value is calculated using the approach described in the previous section, it is matched with the closest value in the data classified using eyeball inspection. Algorithm 1 shows the steps of this method [7].

Algorithm 1 Seasonal Rainfall Classification - Current Method

---

1: **procedure** Manually using excel
2: Arrange Rainfall values for the agreed normal period to the Latest Year
3: Sort the rainfall data in ascending order by rainfall column
4: Divide Number of records into 3 groups
5: Classify data as Below Normal, Normal and Above Normal.
6: Assign remainders if any to middle group
7: Analyse rainfall values and place calculated rainfall value next to closet number
8: The group on which calculated value is placed is the forecast category
9: Forecast can be Above Normal rainfall; Normal rainfall; or Below Normal rainfall

---

The classified seasonal rainfall forecast is given for a three months' period, for example; October November December (OND).

Artificial Neural Network Foresting Model

In this experimentation, already identified rainfall homogeneous regions in the current method, given in Fig: 2 above were also used in the ANN approach.

Data

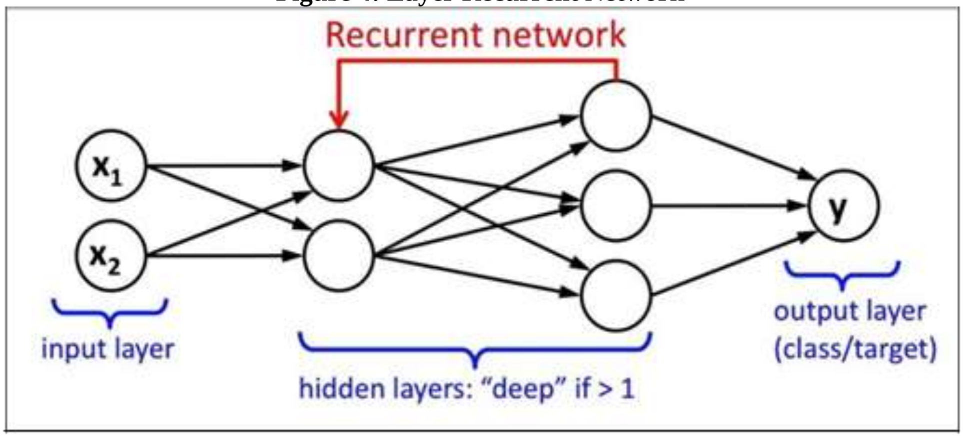
To forecast seasonal rainfall in Zambia using Artificial Neural Networks (ANN), we used Sea Surface Temperatures, Indian Ocean Dipole (IOD) downloaded from the IRI website [28] and station observation data for rainfall, temperature (maximum and minimum) and wind speed in Zambia. The station observation data for 38 stations is from 1961 to 2016 and was collected from the Zambia Meteorological Department. Basins with a more than ±0.3 correlation are selected using an artificial neural network for pattern recognition to estimate coordinates for the basin areas.

Analysis of ANN Techniques

There are several ANN techniques in use for weather forecasting, but only three artificial neural network techniques were analyzed to select a suitable one for generating seasonal rainfall forecasts in Zambia. We therefore began by selecting an appropriate ANN and then applying it to the data used with the current model. Three ANNs were considered, Layer Recurrent, Feed-forward back-propagation and Cascade-forward back-propagation.

Layer Recurrent Network (LRN): Layer Recurrent Network is an ANN technique where connections between nodes form a directed graph along a sequence, thereby exhibiting temporal dynamic behaviour for a time sequence. It is similar to a feed-forward network, except each layer has a recurrent connection with a tap delay associated with it in the layer recurrent network. LRNs can use their internal memory to process arbitrary sequences of inputs, with a feedback loop [29] and has at least one feed-back connection, where activation can flow round in a loop. Thus sequence recognition or prediction of the network is enabled [30]. It takes into account past values making it the best model for regression. LRNs can compute anything with correct set of weights [31] and perform the same task for every element of a sequence with the output being dependent on the previous computations [32]. This allows the network to have an infinite dynamic response to time series input data [33]. Given in Fig: 4 is a layer recurrent network diagram.

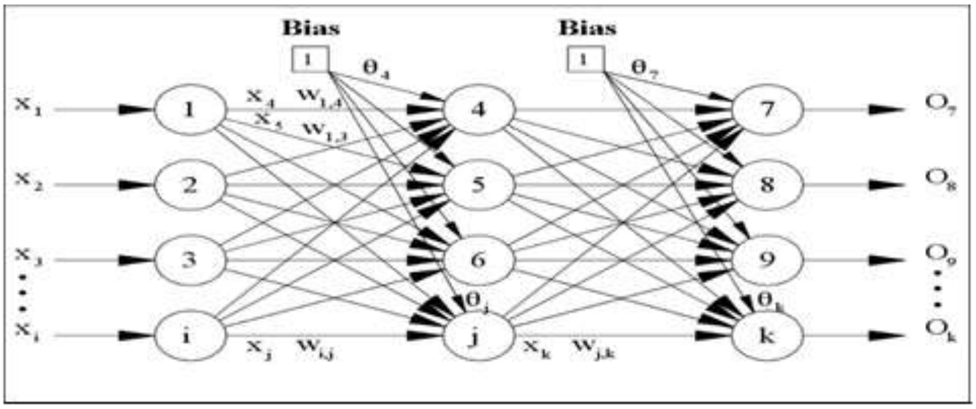
Figure 4: Layer Recurrent Network



Adapted from [34]

Feed-Forward Back-propagation (FFBP): Feed-forward neural network is a fully connected network model that maps the input data sets into the corresponding output sets. Information flows in one direction along connecting pathways, from the input layer via hidden layers to the final output layer. There is no feedback or loops [35], [36]. Back-propagation algorithm is the traditional training method for feed-forward during which the neurons adapt their weights to acquire new knowledge, hence the name feed-forward back-propagation [37]. In Feed-forward back-propagation, connections are fed-forward and whatever is given as input moves to the next layer. Learning occurs during the training phase in which each input pattern from the training set is applied to the input layer and then propagated forward. The pattern of activation arriving at the output layer is then compared with the associated output pattern to calculate an error signal [35]. Error signal for each such target output pattern is then back-propagated from the output layer to the input neurons to adjust the weights in each layer of the network [12]. This works towards error correction and quickly learns complex relationships [29]. Given in Fig: 5 is a Feed-Forward Back-Propagation network diagram.

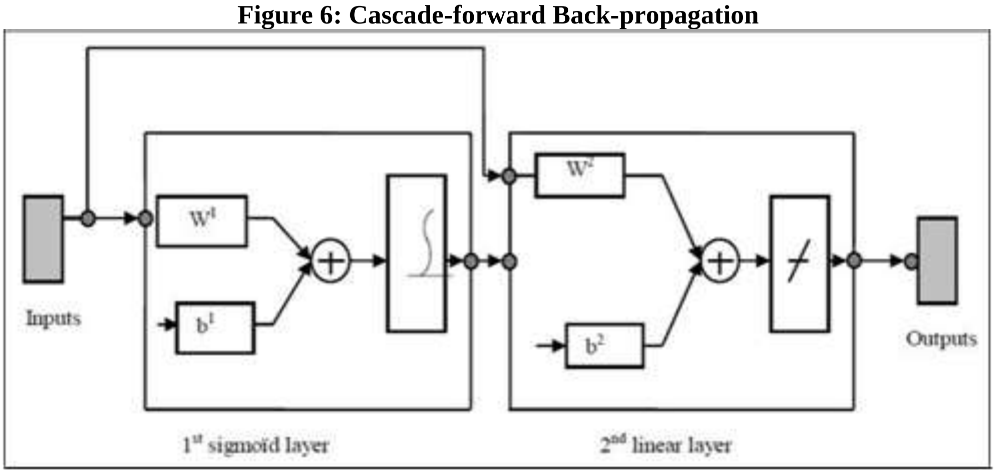
Figure 5: Feed-Forward Back-Propagation



Adapted from [12]

Cascade-Forward Back-propagation (CFBP): Cascade-forward back-propagation is an ANN technique where connections are fed-forward and whatever is given as input moves to the next layer, but includes a connection from the input and every previous layer to the following layers. Though similar to feed-forward back-propagation neural networks in using back-propagation algorithm for weight updating, CFBP includes a weight connection from input layer to each layer and from each layer to the successive layers [29]. CFBP operates in forward input signal and backward error signal [38] and can accommodate nonlinear relationship between input and output by not eliminating the linear relationship between the two [39]. In cascade-forward back-propagation, the neurons of one layer are involved in computation and weight updating of all the layers ahead [40]. Given in Fig: 6 is a Cascade-forward Back-propagation network diagram.

Figure 6: Cascade-forward Back-propagation



Adapted from [38]

Selection of suitable ANN

Each of the three neural network techniques were given prepared data whilst changing the number of neurons, training function and adaptive learning function. Table 6 tabulates results from the three different neural networks techniques based on the training functions, adaptive learning functions and Mean Square Error (MSE).

Table 6: Table caption Testing cases for three Neural Network

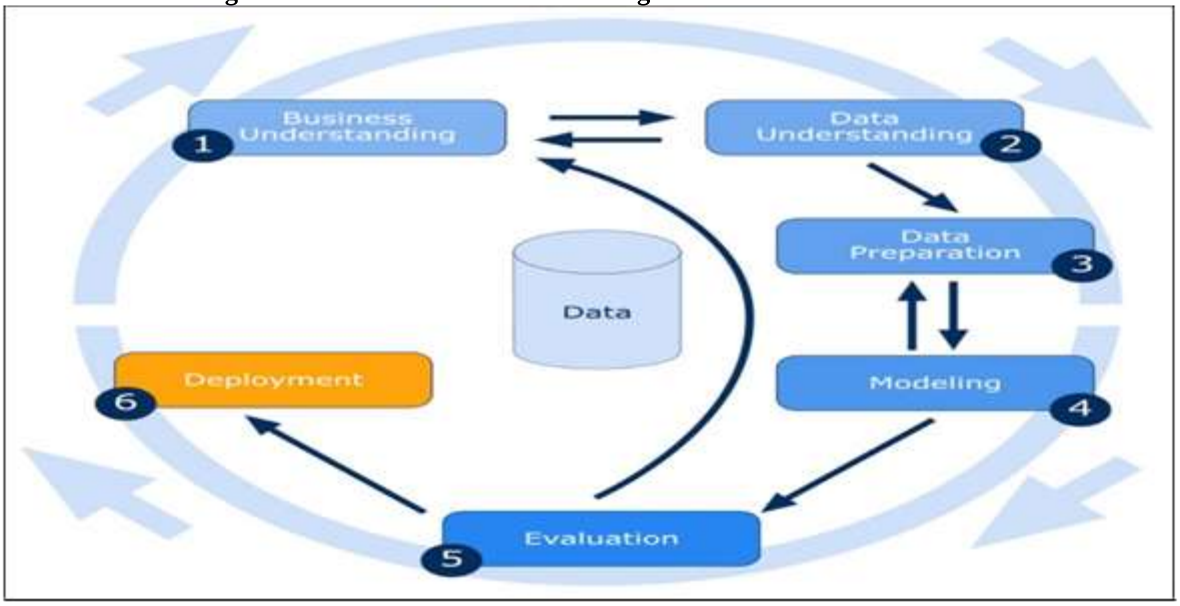
Case No.	Training Function	Adaptive Learning Function	Neurons No.	LRN MSE	FFBP MSE	CFBP MSE
Case1	TRAINLM	LEARNGDM	10	8.66	**6.94**	7.90
Case2	TRAINLM	LEARNGD	10	8.90	10.39	8.48
Case3	TRAINLM	LEARNGDM	20	9.50	7.12	**7.63**
Case4	TRAINLM	LEARNGD	20	**8.31**	7.33	7.93
Case5	TRAINRP	LEARNGDM	10	9.99	9.93	9.97
Case6	TRAINRP	LEARNGD	10	9.70	9.55	9.95
Case7	TRAINRP	LEARNGDM	20	9.60	9.35	9.38
Case8	TRAINRP	LEARNGD	20	9.72	9.49	9.64

From the testing cases in table 6, TRAINLM (training functions) and LEARNGDM (adaptive learning functions) gave the smallest Mean Square Error (MSE) for feed-forward back-propagation technique. This was the basis of selecting feed-forward back-propagation as the artificial neural network technique used in this study. Also, it was selected due to its principles of simplicity and effectiveness [3], [22]. Feed-forward back-propagation neural network uses minimum of the error function in weight space using the method of gradient descent. The technique of artificial neural network used in this research is feed-forward back-propagation. This network used sigmoid activation function which is usually used to predict the probability as an output [41].

CRISP-DM ANN Forecasting Seasonal Rainfall

Forecasting Seasonal Rainfall using ANN was done by following the Cross-Industry Standard Process for Data Mining (CRISP-DM) model [42]. The six phases of the CRISP-DM model were employed as shown in Fig: 7 below.

Figure 7: CRISP-DM ANN Forecasting Seasonal Rainfall in Zambia



Adapted from [42]

Business Understanding - In order to understand the problem space, a situational analysis was conducted to understand the workflows associated with forecasting seasonal rainfall.

Data Understanding - Data sources that were used for collecting input features were identified and mechanisms put in place to reliably collect the data. The identified data are historical station observation data (Rainfall, Maximum Temperature, Minimum Temperature, Relative Humidity and wind speed) which was collected from ZMD in CSV format. Identified oceanic data are SST and Indian Ocean Dipole (IOD) which were downloaded from IRI website in netcdf format.

Data Preparation - The five historical station observation data-sets and two oceanic data-sets were cleaned and normalized to maintain uniformity of data-sets.

Modelling Data, from 1961 to June 2010 was given to the model to learn to find the desired output which was provided. The same data sets, but from July 2010 to June 2011, was given to the trained FFBP model to predict rainfall for the following season. Feed-forward back-propagation model was used for training the network.

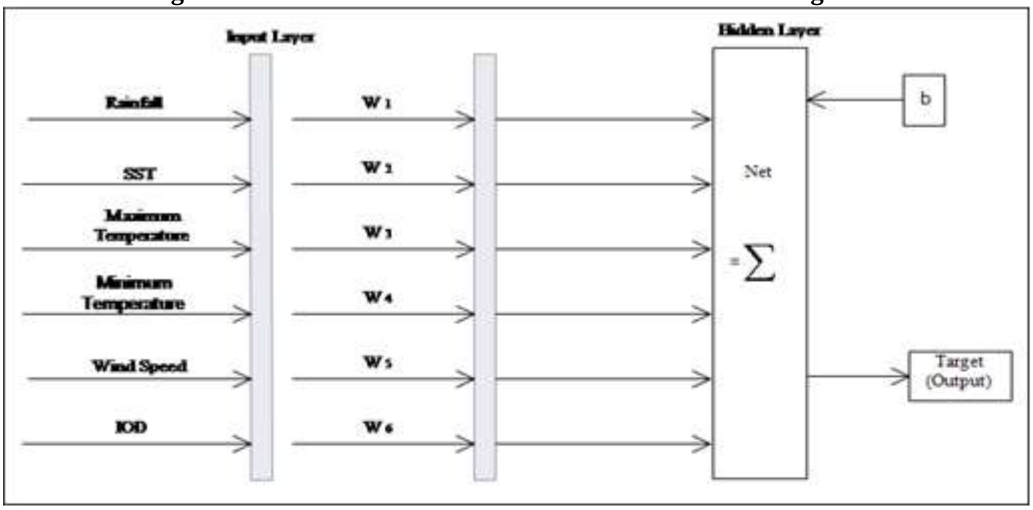
Evaluation - 70 percent of the data-sets were used to train the model, while 15 percent was used for testing the model and other 15 was to evaluate the model. At testing, model performance and optimization were evaluated. Root Mean Square Error model evaluation technique was used to evaluate the model.

Deployment - The forecast model was not deployed. The model was built as a Proof of Concept and was designed to demonstrate the feasibility of the proposed idea of forecasting seasonal rainfall in Zambia using ANNs to solve a business need of accurate seasonal rainfall forecasts.

Proof of Concept Model (PoC Model)

Based on the selected ANN technique above, feed-forward back-propagation neural network technique was used to develop a Proof of Concept Model (PoC Model) to generate seasonal rainfall forecasts in Zambia. The proof of concept was designed to demonstrate the feasibility of the proposed idea of forecasting seasonal rainfall in Zambia using ANN in order to solve a business need of accurate seasonal rainfall forecasts. Six different data sets were given to FFBP neural network as input data and expected output data was also given as targets. The network iterated several times in learning how to get the provided output. A proof of concept model was developed to generate seasonal rainfall forecasts in Zambia and its principle driving function is illustrated in Fig: 8 below.

Figure 8: PoC Model for Seasonal Rainfall in Zambia using ANN



Experiments

Experiments to predict seasonal rainfall were undertaken using data from July 2010 to June 2016, one season at a time. The experiment used a model trained using past data to predict rainfall for the next season. According to the Zambia Meteorological Department, a seasonal rainfall is considered to be from July to June the following year. Given in Table 7 is summary of the data used to train the network.

Table 7: Experimental Data

Data Name	**From**	**To**	**Data Type**	**Data Units**
Rainfall	July 1961	2017	Monthly Total	mm
Maximum Temperature	July 1961	2017	Monthly Mean	$°C$
Minimum Temperature	July 1961	2017	Monthly Mean	$°C$
Wind Speed	July 1961	2017	Monthly Mean	$_m2$
SST	July 1961	2017	Monthly Mean	$°C$
IOD	July 1961	2017	Monthly Mean	Nm^{-2}

Forecast Classification

Using FFBP neural network, seasonal rainfall forecasts were generated and classified into meaningful information using an algorithm. Matlab was used to automatically classify the monthly rainfall numbers into information as Below Normal, Normal or Above Normal rainfall. Given in Algorithm 2 is the automatic classification algorithm used to classify the monthly rainfall values.

Algorithm 2 Seasonal Rainfall Classification - Proposed Approach

1: **procedure** Automatically done using Matlab
2: Read excel file containing rainfall values from 1981 for same month into Matlab

3: Sort records in ascending order by Rainfall column
4: Count the number of records
5: Divide Number of records into 3 group
6: Group data to Below Normal, Normal and Above Normal.
7: Put any reminder to middle group
8: Assign training inputs and targets.
9: Create and Train the Feed Forward Network
10: Select group for the predicted value.
11: Return predicted value and group to which it belongs.

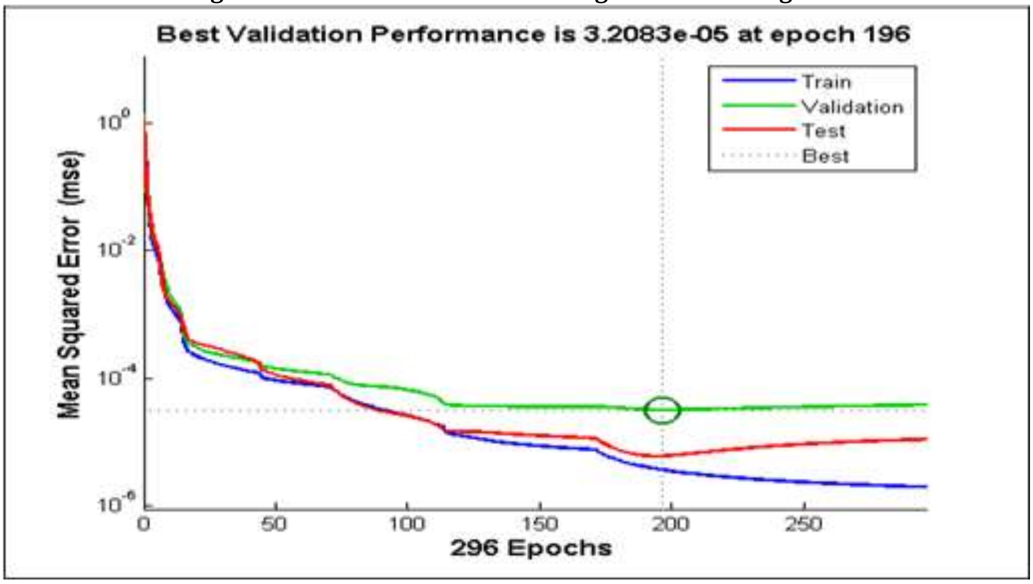
V. Results

Results generated from artificial neural network approach being presented are performance analysis, regression analysis outputs which give the confidence level of the model, predicted rainfall values, forecast classification and forecast period.

Performance Analysis

Performance analysis result is given in a graphical form and shows the current status of the training process. On the performance graph, X - axis indicates the number of iterations (Epochs) which the model went through before it learned to give the given targets, while Y - axis represents the Mean Square Error (MSE) for each iteration [43]. Given in Fig: 9 is the best validation performance graph for the proposed approach (feed-forward back-propagation neural network) for region I.

Figure 9: Best Performance Training Process for Region I

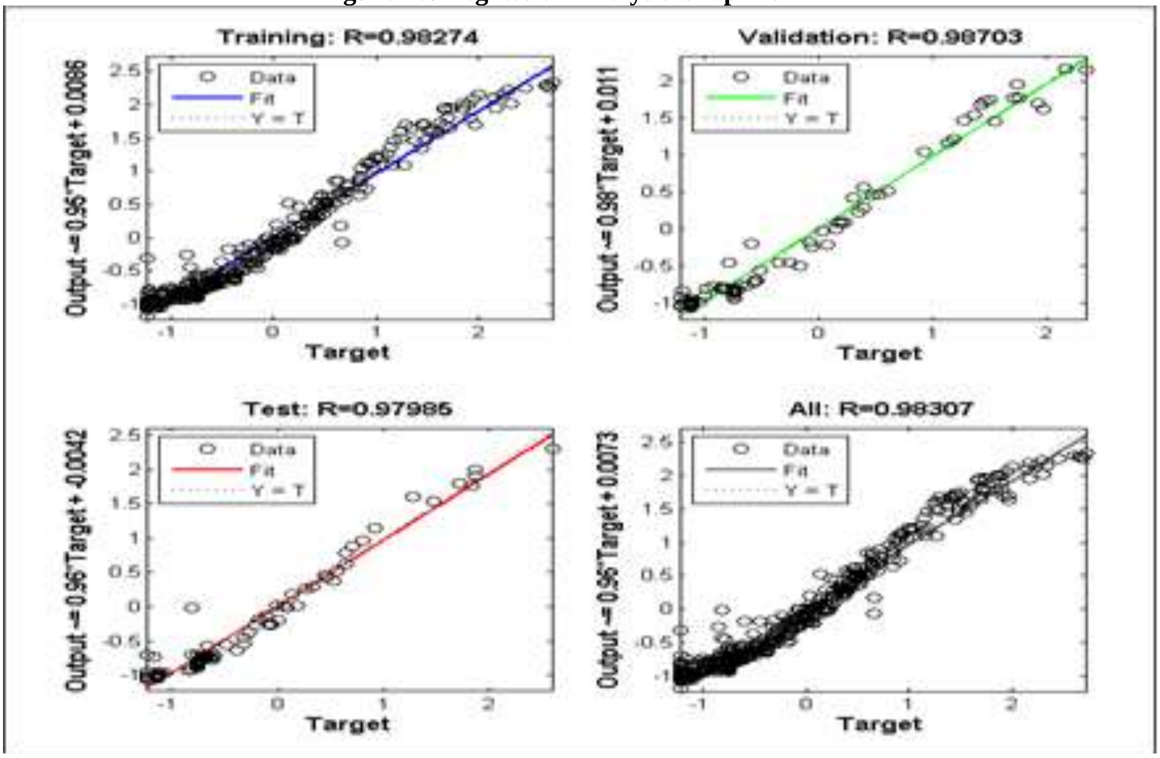


From the best performance graph given above in Fig: 9 above, best performance was reached after 196 iterations (Epochs). Performance graphs are computed for every iteration in the training process and the graph in which all three results of training, validation and testing coincide at almost all points is chosen to be the best performance. When best performance has been achieved, training is stopped and no further iteration proceeds. It may predict results wrongly should further training be done [43].

Regression Analysis

The in Fig: 10 regression graph gives an idea of how accurate the output from the model (feed-forward back-propagation in this case) is compared to the target values. The graph four plots for the regression analysis. These are Training, Validation,Testing and All (average of the three plots) for regression analysis. When inputs = targets, it is represented with dashed lines on the regression analysis plot thus a perfect result. Solid lines represents the best fit linear regression line between outputs and targets. Given in Fig: 10 is the regression analysis output graph for region 1. The combination of R value of training, validation and testing for region 1 was 0.983.

Figure 10: Regression Analysis Graph for R1



Predicted Rainfall

After the network learned the relationship between input and target data (training data), sample data was presented to the network as input, to now generate targets (forecast) based on what the network learned. The same input data parameters (rainfall, wind speed, maximum and minimum temperature, Indian Ocean Dipole and Sea Surface Temperature) that were used to train the neural network, but now from July 2010 to June 2016 were given to the network one season at a time to predict rainfall for following season. After the neural network had learned the relationship between inputs and target, sample data from July 2010 to June 2011 was used to predict rainfall for the July 2011 to June 2012 season. Given in Table: 8 are monthly rainfall forecast values for 2012/2013 season for region 1.

Table 8: Table caption Rainfall Forecast for Region 1

Year	Month	Predicted
2012	Jul	-0.847
2012	Aug	-0.850
2012	Sep	-0.850
2012	Oct	-0.759
2012	Nov	-0.236
2012	Dec	0.900
2013	Jan	0.761
2013	Feb	0.658
2013	Mar	0.076
2013	Apr	-0.569
2013	May	-0.846
2013	Jun	-0.850

Forecast Classification

Generated seasonal rainfall forecast values are just numbers until they are classified into meaningful information. Matlab was used to automatically classify the predicted monthly rainfall values into information as Below Normal, Normal or Above Normal rainfall. The classification was done by comparing the rainfall forecast values with actual values of the same month from the minimum climatology normal period of 30 years for each region. 1981 to 2010 was used as the minimum climatology normal such that the rainfall forecast value for December was compared with December rainfall actual values from 1981 to the latest year's December value. Algorithm 2 was used to automatically classify the monthly rainfall values. Classification of the 2012/2013 seasonal rainfall forecast for Region 1 is given in Table: 9.

Table 9: Classification of Rainfall Forecast for 2012/2013 Season R1

Year	Month	Forecast Classification
2012	Jul	Normal
2012	Aug	Normal
2012	Sep	Normal
2012	Oct	Below Normal
2012	Nov	Below Normal
2012	Dec	Normal
2013	Jan	Normal
2013	Feb	Normal
2013	Mar	Normal
2013	Apr	Below Normal
2013	May	Normal
2013	Jun	Normal

Forecast Period

A forecast period is the duration for which a generated forecast remains valid. Generally, forecast period ranges from hours, days, months, seasons to years [44], [45], [46]. From the rainfall prediction table given in table 8, seasonal rainfall forecast was generated on a monthly basis.

VI. Evaluation and Discussion

ANN results were evaluated by comparing them with results from the current seasonal rainfall forecasting method and actual rainfall values collected data from ZMD for particular seasons. This was done in order to ascertain which of the two forecast generation methods (current and ANN) gave an enhanced seasonal rainfall forecast skill. Comparison was based on the regression analysis output, length of forecast period and method objectivity. The R value from regression analysis output of the current forecasting method given in Tables: 1 - 5 was compared to the R value from regression analysis for the same region in Fig: 9 from the ANN model. This comparison is given in Table: 10 below.

Table 10: R-value Comparison

Current Method	ANN Model
0.309	0.983

Multiple R is the absolute value of the correlation coefficient. In the current forecasting method, multiple R gives the regression analysis value, which was 0.309. In the ANN model, regression analysis process was analyzed when training the network, during validation, testing and an average combination of all three. From the example given in Fig: 10 for region I, the regression analysis for all is 0.983. The R values during training, validation or testing are all above 0.9. R value of close to one indicates a higher linear relationship between inputs and targets, which means a higher confidence level of the model or method [47]. Therefore, ANN model gave a higher confidence level compared to the current seasonal rainfall forecasting method.

The forecast period in the current forecasting method is three months i.e. January February March for each region. In the ANN model, the forecast period is a month as shown in the Table: 4. According to Ashok Kumar, long range forecasts like three months' period in the current method, become less accurate as the difference in time between the present moment when the forecast is generated and the time for which the forecast is given increases. A shorter forecast period like monthly forecast has enhanced accuracy [10]. Both the current forecasting method and proposed gave predicted rainfall values in indices format because standardized data values were used as inputs, targets and sample data.

Selection of basins with a ±0.3 correlation between station rainfall data and SST is done through eye ball inspection by zooming to each basin, one at a time whilst noting the coordinates in the current forecasting method. Further, rainfall values from 1981 to year before forecast season are sorted, grouped into the three terciles. Calculated rainfall value is compared to sorted data using eye ball inspection and placed close to one it is nearest to. Thus determining the category to be either below normal, normal or above normal rainfall [48]. Use of eye ball inspection makes the whole process subjective to the person's judgment.

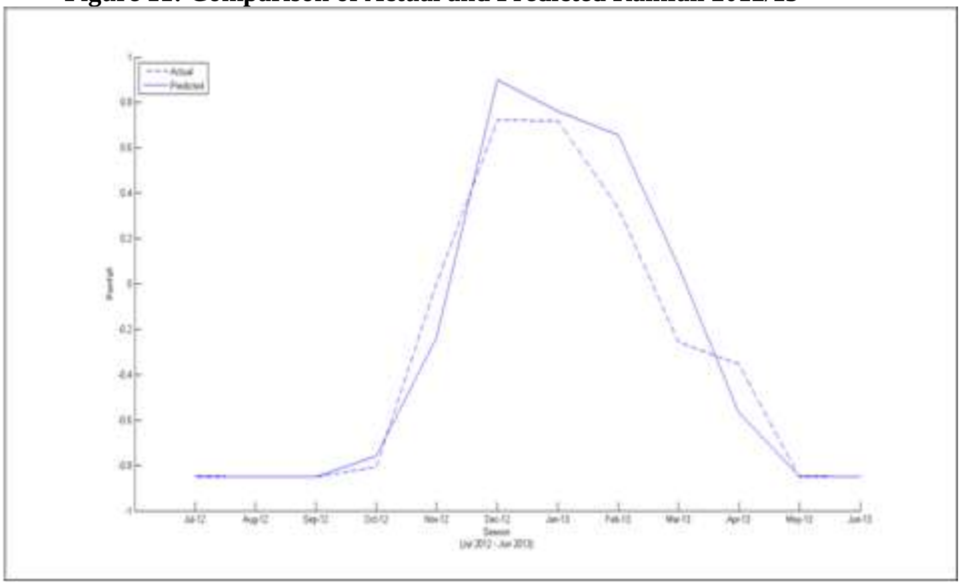
In the proposed approach, an algorithm was used to pick category of the generated rainfall values as either below normal, normal or above normal rainfall automatically. Furthermore, selection of basin coordinates with a more than ±0.3 was done using ANN for pattern recognition in the proposed approached. This automation of selection of basins and classifying of predicted rainfall values made the proposed approach objective and easy to pass on knowledge [48]. Generated seasonal rainfall values were compared with actual rainfall values for the same season and region. Comparison was done both in table and graph formats. Give in table 11 is the comparison of actual and ANN predicted values for 2012/2013 season.

Table 11: Comparison of Actual and Predicted Rainfall values for 2012/2013 Season

Year	Month	Actual	Predicted
2012	Jul	-0.849	-0.847
2012	Aug	-0.849	-0.850
2012	Sep	-0.849	-0.850
2012	Oct	-0.806	-0.759
2012	Nov	0.004	-0.236
2012	Dec	0.721	0.900
2013	Jan	0.720	0.761
2013	Feb	0.335	0.658
2013	Mar	-0.257	0.076
2013	Apr	-0.354	-0.569
2013	May	-0.849	-0.846
2013	Jun	-0.849	-0.850

For easy comparison between monthly actual rainfall values and monthly rainfall forecast values (from FFBP), a graph is given in Fig: 11 for 2012/2013.

Figure 11: Comparison of Actual and Predicted Rainfall 2012/13



From the comparison graphs given in Fig: 11, feed-forward back-propagation model was able to pick the increases and decreases of rainfall or follow the pattern with acceptable accuracy. Except for January 2013 where the predicted rainfall was very high compared to what was actually recorded.

VII. Conclusion

The conclusion is presented in line with the results obtained after evaluating the forecast model and research aims. Research results have revealed that forecasting seasonal rainfall in Zambia using artificial neural networks enhanced the forecast accuracy skill. The results and discussions in this study have clearly revealed that artificial intelligence tools like artificial neural network techniques:

1. Offer an enhanced seasonal rainfall forecast accuracy level because more factors that influence rainfall were incorporated than in the current forecasting method. Seasonal rainfall forecast skill was further enhanced, due to the artificial neural network's ability to acquire knowledge through a learning (training) process and use of the inter-neuron connection strength to store knowledge and apply that knowledge to predict when given a new data set.
2. Are data-driven, self-adaptive methods that learn from examples and don't need restrictive assumptions about the form of the basic model. ANN discovers the relationships among data which may be too complex to define.
3. Are a family of massively parallel architectures capable of learning and generalizing from examples and experience to produce meaningful solutions to problems, even when input data may contain errors and are incomplete. This makes ANN a powerful tool for solving problems like forecasting long range rainfall.
4. For use of ANN in seasonal forecasts, feed-forward back-propagation neural network gave a minimized error. FFBP neural network was selected for use because of its principles of simplicity and effectiveness after analyzing three different ANN models. Further, FFBP gave a smallest MSE because it is based on error correction learning rule.
5. The ANN model gave higher confidence level (regression analysis R-value) compared to the current method. This is because artificial neural networks ability to learn and adapt to the environment used during training.
6. Eye ball inspection stages were eliminated which made ANN model objective, thereby enhancing forecast skill.
7. Reduction of the forecast period from three months in the current method to one month in the ANN model, further improved the accuracy

References

[1] DeepaBisht, Mahesh C Joshi, And Ashish Mehta. Prediction of Monthly Rainfall of Nainital Region Using Artificial Neural Network (Ann) And Support Vector Machine (SVM). IJARIIE. 2015 1(3):400–406.

[2] John Abbot and Jennifer Marohasy. Application of Artificial Neural Networks to Rainfall Forecasting in Queensland, Australia. Advances in Atmospheric Sciences. 2012 (2943):717–730.

[3] Kumar Abhishek, Abhay Kumar, Rajeev Ranjan, And Sarthak Kumar. A Rainfall Prediction Model Using Artificial Neural Network. IEEE Control and System Graduate Research Colloquium (ICSGRC). 2012 4:82–87.

[4] Lubna Shaikh and KirtiSawlani. A Rainfall Prediction Model Using Artificial Neural Network. International Journal of Technical Research and Applications. 2017 5(2):45–48.

[5] Jignesh Patel and Falguni Parekh. Forecasting Rainfall Using Adaptive Neuro-Fuzzy Inference System (ANFIS). IJAIEM. 2014 3(6):262–269.

[6] Mzyece L., Nyirenda M., Kabemba, K. M. And Chibawe G. Forecasting Seasonal Rainfall in Zambia – An Artificial Neural Network Approach. Zambia Information Communication Technology (ICT) Journal. 2018 2(1):16–24.

[7] Nsandisa Sunshine, Mduduzi; Oludhe, Christopher; Faka. Manual on the Development of Statistical Seasonal Forecasting Models SADC-CSC. Technical Report, SADC. 2012.

[8] GyaneshShrivastava, Sanjeev Karmakar, Manoj Kumar Kowar, and PulakGuhathakurta. Application of Artificial Neural Networks in Weather Forecasting: A Comprehensive Literature Review. International Journal of Computer Applications. 2012 51(18):17–29.

[9] Kevin E. Trenberth. Changes in Precipitation with Climate Change. Climate Research. 2011.47(1-2):123–138

[10] Ashok Kumar. Statistical Models for Long-Range Forecasting of Southwest Monsoon Rainfall Over India Using Step Wise Regression and Neural Network. Atmospheric and Climate Sciences. 2012 02(03):322–336.

[11] Gina Ziervogeland Rebecca Calder. Climate Variability and Rural Livelihoods: Assessing The Impact of Seasonal Climate Forecasts in Lesotho. Society. 2003 403– 417.

[12] Puja Gupta and Neha Kulkarni. An Introduction of Soft Computing Approach Over Hard Computing. International Journal of Latest Trends in Engineering and Technology. 2013 3(1):254–258.

[13] Daniele Chella, Antonio; Iocchi, Luca; Macaluso, Irene; Nardi. Artificial Intelligence and Robotics. International Journal of Computer Application. 2013.

[14] Singh VibhashShivani; Sisodia. Application of Soft Computing Techniques: Fuzzy Logic and Genetic Algorithms. International Journal of Scientific Research Engineering and Technology (IJSRET). 2012 1(5):309–312.

[15] B. Yegnanarayana. Artificial Neural Networks. Prentice-Hall of India, 11th Edition, New York.; 2005.

[16] S. N.; Deepa S.N. Sivanandam. Principles of Soft Computing. Wily, 2nd Edition.; 2011.

[17] A D Dongare, R RKharde, And Amit D Kachare. Introduction to Artificial Neural Network. International Journal of Engineering and Innovative Technology (IJEIT). 2012 2(1):189–194.

[18] ParveenEr Kumar and Pooja Er Sharma. Artificial Neural Networks-A Study. International Journal of Emerging Engineering Research and Technology. 2014 2(2):143–148.

[19] C Devi, B Reddy, and K Kumar. Ann Approach for Weather Prediction Using Back Propagation. International Journal of Engineering Trends and Technology. 2012 3(1):19–23.

[20] Nazim Osman BusharaandAjith Abraham. Computational Intelligence in Weather Forecasting. Journal of Network and Innovative Computing. 2013 1:320–331.

[21] Ankit Chaturvedi. Rainfall Prediction using Back-Propagation Feed Forward Network. International Journal of Computer Applications. 2015 119(4):1–5.

[22] N Q Hung, M S Babel, S Weesakul, and N K Tripathi. An artificial neural network model for rainfall forecasting in Bangkok, Thailand. Hydrology and Earth System Sciences. 2008 13(8):1413–1425.

[23] F Mekanik, T S Lee, and M AImteaz. Rainfall Modeling using Artificial Neural Network for a mountainous region in West Iran. In 19th International Congress on Modelling and Simulation. 2011 :12–16.

[24] S C Michaelides, C S Pattichis, and G Kleovoulou. Classification of rainfall variability by using artificial neural networks. International Journal of Climatology. 2001 21(11):1401–1414.

[25] Dayanand G Savakar and Rani Channamma. University. Rainfall Prediction based on Rainfall Statistical Data. Inter-national Journal on Recent and Innovation Trends in Computing and Communication. 2016 4(3):270–275.

[26] Rohit R Deshpande. On The Rainfall Time Series Prediction Using Multilayer Perceptron Artificial Neural Network. International Journal of Emerging Technology and Advanced Engineering. 2012 2(1):148–153.

[27] Zambia Meteorological Department. Revised Agro-Ecological Regions of Zambia. Zambia Meteorological Department (ZMD). Computer Unit, 2006.

[28] National Oceanic and Atmospheric Administration. Extended Reconstruction Sea Surface Temperature. http://iridl.ldeo.columbia.edu/ SOURCES/.NOAA/.NCDC/.ERSST/.version3b/2007. [Online: accessed 1-November-2017].

[29] SumitGoyal and Gyandera Kumar Goyal. Cascade and Feedforward Backpropagation Artificial Neural Network Models for Prediction of Sensory Quality of Instant Coffee Flavoured Sterilized Drink. Canadian Journal on Artificial Intelligence, Machine Learning and Pattern Recognition. 2011 2(6):01–06.

[30] A. John Bullinaria. Biological Neurons and Neural Networks. Artificial Neurons. 2015.

[31] WeidiXie, Alison Noble, and Andrew Zisserman. Layer Recurrent Neural Networks. In Layer Recurrent Neural Networks. 2017 :1–17.

[32] Denny Britz. Recurrent Neural Networks Tutorial, Part 1 – Introduction to RNNs. http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/, 2015. [Online: accessed 13-August-2018].

[33] Ankita Sharma and GeetaNijhawan. Rainfall Prediction Using Neural Network. International Journal of Computer Science Trends and Technology. 2015 3(3):65–69.

[34] Leonardoaraujosantos. Artificial Neural Networks.https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/recurrentneural-networks.html/. 2011. [Online: accessed 12-June-2018].

[35] J T Lalis, B D Gerardo, and Y Byun. An Adaptive Stopping Criterion for Backpropagation Learning in Feedforward Neural Network. International Journal of Multimedia and Ubiquitous Engineering. 2014 9(8):149–156.

[36] Murat H Sazli. A brief review of feed-forward neural networks. Commun. Fac. Sci. Univ. Ank. Series A2-A3. 2014 50(1):10–17.

[37] XinghuowyYu, Mehmet Efe, and OkyayKaynak. A general backpropagation algorithm for feedforward neural network learning. IEEE Computational Intelligence. 2002 13(1):9–13.

[38] SupriyaNarad and PallaviChavan. Cascade Forward Back-propagation Neural Network Based Group Authentication Using (n, n) Secret Sharing Scheme. Procedia - Procedia Computer Science. 2016 :185–191.

[39] HasbiWarsito, Budi and Santoso, Rukun and Suparti and Yasin. Cascade Forward Neural Network for Time Series Prediction. In Forward Neural Network for Time Series Prediction Series 1025. 2018 6(1):19860–19866.

[40] Pranay Subramanian, deviRenuga; Arulmozhivarman, Pachiyappan.; Venkatesh, C.; Agarwal. Performance comparison of artificial neural network models for daily rainfall Performance Comparison of Artificial Neural Network Models for Daily Rainfall Prediction. International Journal of Automation and Computing. 2016 7(3):331–384.

[41] K Amardeep, Rashmi and Thippeswamy. Training Feed Forward Neural Network with Back-propagation Algorithm. International Journal of Engineering and Computer Science. 2017 6(1):19860–19866.

[42] Wirth,Rudiger and HippJochen CRISP-DM: Towards a standard process model for data mining. Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining. 2000.

[43] AMuthuchuda. Artificial Neural Networks - training, testing and validation. http://shodhganga.inflibnet.ac.in/bitstream/10603/35997/6/.pdf. 2015. [Online: accessed 13-November-2018]

[44] SanamNarejo and Eros Pasero. Meteo-nowcasting using Deep Learning Architecture. International Journal of Advanced Computer Science and Applications (IJACSA). 2017 68(8):16–23.

[45] Salem Ibrahim and Gamal S El Afandi. Short-Range Rainfall Prediction over Nigeria Using the Weather Research and Forecasting Model. Open Journal of Atmospheric and Climate Change. 2014 1(2):2374–3808.

[46] Ganesh P Gaikwad and Prof V B Nikam. Different Rainfall Prediction Models and General Data Mining Rainfall Prediction Model. International Journal of Engineering Research and Technology(IJERT). 2013 2(7):115–124.

[47] MathWorks Documentation. Analyze Shallow Neural Network Performance After Training. http://nl.mathworks.com/help/deeplearning/ug/analyze-neural-network-performance-after-training.html;jsessionid= cb11a75a1d3d02f8196ce498. 2018. [Online: accessed 13-November-2018].

[48] Lillian Mzyece. MSc Dissertation; Forecasting Seasonal Rainfall in Zambia using Artificial Neural Networks. Lusaka, 2019.