

## Research Article

# Seasonal rainfall prediction in Juba County, South Sudan using the feed-forward neural networks

David Lomeling\*, Salah Joseph Huria

Department of Agricultural Sciences, College of Natural Resources and Environmental Studies (CNRES), University of Juba, P.O. Box 82 Juba, South Sudan

(Received: November 19, 2019; Revised: July 2, 2021; Accepted: July 25, 2021; Published: August 13, 2021)

\*Corresponding Author: David Lomeling (Email: [dr.david\\_lomeling@gmx.net](mailto:dr.david_lomeling@gmx.net))

### ABSTRACT

Historical rainfall data from 1997-2016 of Juba County, South Sudan were used in a Feed-Forward Neural Network (FFNN) model to make future predictions. Annual rainfall data were aggregated into three seasons MAMJ, JAS and OND and later trained for best forecasts for the period 2017-2034 using the Alyuda Forecaster XL software. Best training of the time series was attained, once the *minimum error* of the weight  $\Delta\omega$  expressed as MSE or AE between the measured variable and predicted was achieved during gradient descent. The results showed that for MAMJ and JAS months, the number forecasts were over 85% whereas this was between 60-80% for OND months. The Seasonal Kendal (SK) test on future rainfall forecasts as well as the Theil-Sen slope showed negative monotonic trends in the mean values till the end of 2034 of all three seasons with MAMJ, JAS at OND at 100, 150 and 80 mm respectively. Rainfall forecast showed that the MAMJ months for the years 2019 to 2027 will be moderately wet except in April 2021 which will experience some severe wetness (*due to intensive rainfall*). Interdecadal severe drought with less than 60, 100 and 10 mm for MAMJ, JAS and OND respectively, is expected between 2028 to 2033 after almost two decades. Declining onset of MAMJ rains is expected to significantly affect the timing for land preparation and crop planting.

**Keywords:** Feed Forward Neural Network, rainfall forecasting, training set, test set, Theil-Sen slope estimator

### INTRODUCTION

Agriculture and more specifically crop production, is the mainstay of much of the rural population of Juba County of South Sudan. It is predominantly rain-fed and is contingent upon the frequency, intensity and magnitude of rainfall as well as on spatial-temporal variations. Until recently, rainfall onset was based on local farmers' calendar and traditional knowledge. Often, these were more of guess work with less predictability and reliability and especially now coupled with the effects of climate change. Although no prior studies on the seasonality of rainfall distribution in Juba County have been conducted, experiential evidences from farming communities over the last two decades report of clear deviations and decrease below mean values. Generally, the annual rainfall onset starts during the second to third dekad of April (*Ja'be*) and continues till June punctuated with a dry spell around July. Rainfall then continues from August till October, significantly decreasing toward November and December dry season (*Méling*). The rainfall pattern may be described as more or less bimodal in nature. Belated occurrences of onset rains during this period over the last two decades could be attributed to the prolonged impacts of *El Niño* within the East African region resulting into untimely availability of soil moisture (Lomeling *et al.*, 2016) and hence poor harvests of crops like cowpea, maize or peanuts. Understanding such erratic rainfall events and seasonal patterns is imperative in first understanding "meteorological drought" and the subsequent implications on

the "agricultural drought". The former is expressed entirely based of the degree of dryness (usually related to rainfall anomaly from the long-term mean) whereas the latter is based on temporal soil moisture deficit during crop phenology coupled with intensive actual evapotranspiration.

Rainfall predictions for South Sudan encompassing the study area, have in the last decade been issued by diverse regional and international institutions like the IGAD Climate Prediction and Applications Centre, (ICPAC); UN Food and Agriculture Organization (FAO); United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA). These predictions are, however, monthly with short decadal timescales and often in the form of probabilities relative to monthly or seasonal rainfall averages. Spatial and temporal rainfall patterns do not often correlate with soil moisture contents and dynamics. Li *et al.* (2016) showed that surface soil moisture dynamics generally follow rainfall patterns at the two gravel plain sites, whereas this was not the case soil moisture dynamics in the sand dune site. Therefore, depending on intensity of rainfall, soil structure, surface sealing and infiltration, clear distinctions between meteorological and agricultural droughts should be made and how both are interlinked.

In the last two decades, much research on rainfall prediction using the ANNs have been conducted in different parts of the world from monthly time series (Singh, 2018); seasonal (Hartmann *et al.*, 2016), daily (Devi *et al.*, 2016); hourly (Hung

et al., 2009); dekadad (Warsito et al., 2016); monthly (Singh, 2018). A comprehensive overview of ANNs use in temporal rainfall prediction has been reported by (Haviluddin et al., 2015). However, only a few similar studies have been conducted in Africa, in Ethiopia (Abbot and Marohasy, 2017), in Algeria (Elsanabary and Gan, 2014) and in West Africa (Benmahdjoub et al., 2013). Basically, the ANNs is a type of Machine Learning (ML), whereby a computer-based model fed with historical data in a time series is trained to identify specific patterns and the derived “intelligence” later used to predict future events.

In our study, we attempted to evaluate the relevance and accuracy of FFNN in forecasting “lumped” seasonal precipitation derived from historical data.

### ANN Architecture

The ANN model is based on a simplified and popularly used neural network architecture called multilayer perceptron network (MLPN) model (also known as *multilayer feedforward network*) consisting of an input, hidden and output layers as in Figure 1. The basic concept is training neural network so that the results of the goal function in the output layer are reflective of a sigmoid function. For a single pair of neurons in the hidden layer, random weights ( $w$ ) are assigned to each connection and the result in the output layer is a product of the initial value ( $x$ ) and the weights connected to this neuron. The accuracy of the neural networks during training can be improved by dividing a large neuron attribute number  $N$  into several discretized continuous subintervals of equal lengths. With value [1] for example, this number  $N$  can be divided into several paired subintervals (Badr et al., 2014). In our case, we illustrate that, if  $N=5$  then each would be subdivided into 0.2 subinterval and hence give 10 neurons. A further increase of  $N=10$  (or 20 neurons), the subintervals would be smaller at each 0.1. After training of the neurons, the approximated

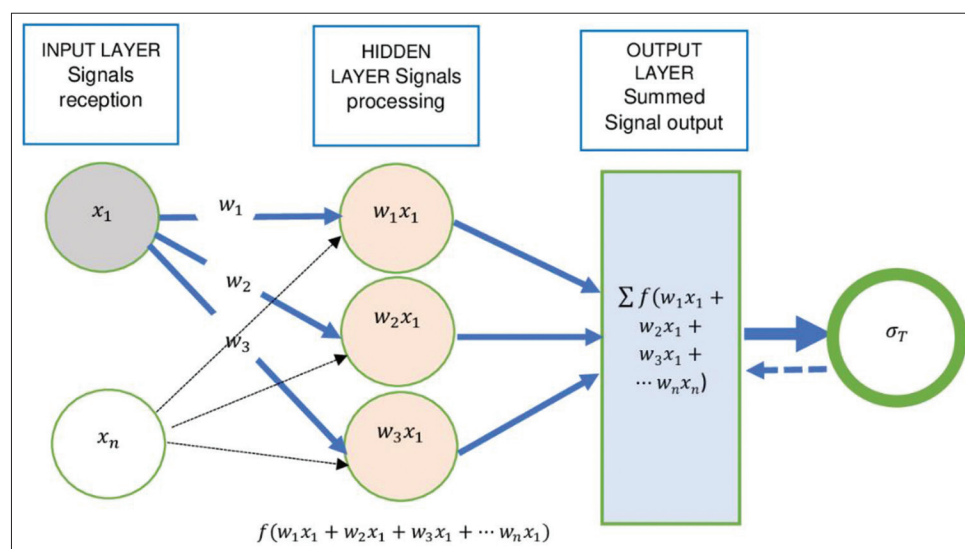
possible values in the output layer introduced by the non-linear activation function (*sigmoid function,  $f$* ) are anything between [0 or 1]. This process is repeated iteratively for the several weights connected to their respective neurons in the hidden layer and the resulting output approximated. For the input vector  $x_1$ , the magnitude in the output layer is the product of each paired neuron and their respective weighted values  $f(w_1 x_1 + w_2 x_1 + w_3 x_1)$  passed on upon activation through a non-linear sigmoid function  $f(x) = \frac{1}{1 + e^{-x}}$  with some bias  $b$  on each neuron.

The total output ( $\sum f x_i w_i$ ) is then the output of the entire ANN for that specific input ( $x$ ) and is compared to the target value. The difference is expressed as the measure of error ( $E$ ) between the computed and expected values. The process of back propagation from the output to the hidden layer ensues and continues iteratively depending on margin of error till a minimum error value is attained.

The optimal number of neurons in the hidden layer was obtained experimentally running the training process several times until a good performance was obtained or when no other changes were observed.

### METHODOLOGY

In general, there are five basics steps: (1) collecting data, (2) preprocessing data, (3) building the network, (4) training and (5) test performance of model. The basic flow in designing ANNs model is given in Figure 1. The daily rainfall data for Juba weather station as from the years 1983 to 2015 were downloaded from the US National Oceanic and Atmospheric Administration (NOAA). However, daily and consistent rainfall data were only recorded as from 1997 to 2016 and were used for this study. Data preprocessing involved aggregating the daily rainfall amounts to monthly



**Figure 1:** Simplified architecture of a three-layered ANN with a single neuron in the input, hidden and an output layers with backpropagation (dotted blue).

means of March-April-May-June (MAMJ), July-August-September (JAS) and October-November-December (OND). Due to the unpredictable onset of rains especially between mid to end of March of each season, the MAMJ was “lumped” together. Rainfall around mid-March prior to the onset of the rainfall season in April is characterized by drizzles and light rainfall showers. With these monthly rainfall data sets, neural networks were then created and later proceeded by training and forecasting.

The chosen rainfall data for each season were divided into two random groups, the training and test sets corresponding to 82% and 18% respectively. Networks were trained for a fixed number of epochs or iterations till a minimum error function was reached.

### The Seasonal Kendall (SK) Test

The rank-based nonparametric Seasonal-Kendall method was applied to the long-term rainfall to detect any statistically significant trends. In this SK test and for the null hypothesis ( $H_0$ ), assumed that there was no monotonic trend in precipitation amounts over time; and for the alternate hypothesis ( $H_1$ ), it assumed that there was either an increasing or decreasing monotonic trend over time.

Once the seasonal rainfall data from 1997-2015 for MAMJ, JAS and OND months were trained and future forecasts made using neural network, test for the presence of any monotonic trend in the seasonal rainfall during the entire period between 1997 to 2034 was conducted using the Seasonal Kendall (SK) test (Hirsch *et al.*, 1982; Gilbert, 1987; Helsel and Hirsch, 1995; Nielsen, 2015). The SK statistic for the  $i$ -th season  $S_i$  may be computed as:

$$S_i = \sum_{(k=1)}^{(n-1)} \sum_{(j=k+1)}^n \text{sgn}(x_{ij} - x_{ik}), i = 1, 2, 3, \dots, n \quad (1)$$

$$\text{sgn}(x_{ij} - x_{ik}) \begin{cases} 1 \text{ if } x > 0 \\ 0 \text{ if } x = 0 \\ -1 \text{ if } x < 0 \end{cases} \quad (2)$$

Where  $\text{sgn}(x_{ij} - x_{ik})$  is the indicator function for the month ( $i$ ) for the two respective years  $j$  and  $k$ . The variance  $\bar{S}$  of  $S$  for the entire series may be computed as:

$$\bar{S} = (n(n-1)(2n+5) - \sum_{i=1}^m t_i(t_i-1)(2t_i-5)) / 18 \quad (3)$$

where  $t_i$  denotes the number of ties with  $i$  tied values and  $m$  the number of tied groups of values. The presence of a statistically significant trend is evaluated using the  $Z$  statistic. A positive value of  $Z$  indicates an upward trend and a negative value indicates a downward trend. The value  $Z$  was computed as:

$$Z = \begin{cases} \frac{\bar{S}-1}{\sqrt{\text{Var}(\bar{S})}} & \text{if } \bar{S} > 0 \\ 0 & \text{if } \bar{S} = 0 \\ \frac{\bar{S}+1}{\sqrt{\text{Var}(\bar{S})}} & \text{if } \bar{S} < 0 \end{cases} \quad (4)$$

A positive monotonic trend is considered significant, if  $Z > 1.96$  or ( $< -1.96$ ) at  $p < 0.05$  or  $p < 0.01$  and vice versa. Incorporating the Theil-Sen slope estimator to the SK test gives a better understanding of the magnitude (*change of unit per time*) of the slope. Generally, the slope  $Q$  between two successive values in a time series is expressed as:

$$Q = \frac{x_j - x_k}{j - k}, j \neq k \text{ and } j > k \text{ for } i = 1, 2, \dots, N \quad (5)$$

Where  $x$  is the value at the  $j$  and  $k$ -th interval for  $n$  observations and  $N = n(n-1)/2$ . Significant trends at  $p < 0.05$  or  $0.01$  can then be computed with the confidence limits defined by  $M_1$  and  $M_2$ . Derivation of this index is referred to Salmi *et al.* (2002).

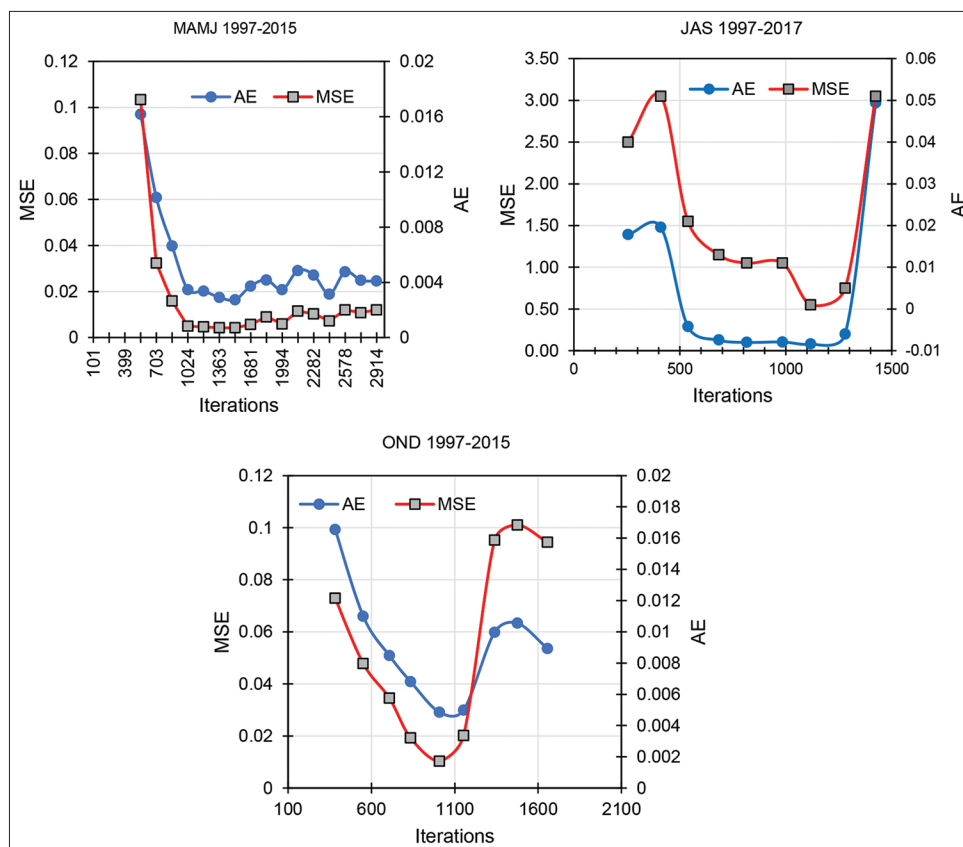
The presence of a statistically significant trend is evaluated using the  $Z$  statistic. A positive value of  $Z$  indicates an upward trend and a negative value indicates a downward trend.

### Data pre-processing and network training

In order to enhance a faster convergence, the monthly input variables were normalized relative to the seasonal averages. Data normalization is the process of “scaling down” the individual raw input data relative to their mean value. The normalized values consistent with the sigmoid activation function would range between 0 and 1. Since probability is between 0 and 1, the normalized values would learn faster and give better predictions during training. The Alyuda ForecasterXL basically splits the data into two sets (1) training and validation sets (2) training set. During training, the weights of the neural network were adjusted whereas the validation increases the accuracy by minimizing the error function ( $E$ ) during iteration. The training stopped once the error function reached a global minimum (Figure 2). Finally, the performance of the network was evaluated on the test data set which had not been involved in the training process. In this study, the neural network was trained with 76, 56 and 55 datasets for the MAMJ, JAS and OND months respectively.

### Model performance

The performance of the neural network was best done by using the linear regression coefficients ( $r^2$ ) of the actual and forecasted data during training. Hereby, the regression coefficients for each season for the test period 1997-2015 were calculated for the entire dataset, as well as the best



**Figure 2:** Error development and convergence to global minimum during training of MAMJ, JAS and OND of historical data between 1997 to 2015.

model predictive performance in terms of good and bad forecasts (*expressed relative to 100% highest accuracy*) for the training ( $P_{train}$ ) and test ( $P_{test}$ ) data respectively. As in our case,  $r^2 \geq 0.9$  and the accuracy for the training set was  $P_{train} \geq 90\%$  while for test set was  $P_{test} \geq 70\%$  and were considered as good model performance indicators within the error of tolerance.

For each of the seasons, we started from a network with one input and output layers in and choosing randomly between 1 and 4 hidden layers. For all the seasons, the error tolerance for both training and test sets was set at 10 and 30% respectively. The training was run several times for each case till the MSE, AE or tolerance error was low and the percentage of good forecasts or the highest correlation between actual and forecasted data attained. (Table 2) shows the relationships between the forecasted and observed values of the training and data sets. However, more emphasis was laid on the number of hidden layers during each training for each season. The best neural performance was with a single hidden layer for JAS and OND months. The predicted rainfall amounts therefore, varied between 0.6-3.0 mm from the mean. The error tolerance during training for all the months was necessary to ensure that no overfitting occurred. This is when the number or percentage of bad forecasts, the model performs less well on the test set as in MAMJ and OND. Hereby, the test set becomes too adapted to the training data and may make unwanted generalizations. On

the other hand, as in the months of JAS, the model was assumed to perform better once the error margin in the test set are less than those in the training set. The model therefore is assumed to have learnt better and so make better predictions. Although the actual and predicted data during training seemed to give high correlation coefficients ( $r^2$ ), it still showed significant amount of error as in (Figure 2) for the months of MAMJ in the years 2005-07, 2012 and 2014-15.

Assessing neural accuracy was tested by using both the Mean Standard Error (MSE) and Absolute Error (AE) during training. AE is the absolute difference between the predicted and observed values. Training using a single input variable ( $x_{ij}$  measured rainfall amount for the  $i$ -th training case at the  $j$ -th network output) for ( $n$ ) observations in a time series was conducted and the best forecast or prediction ( $\hat{x}_{ij}$ ) after each iteration was estimated by minimum *error function* denoted by the AE or MSE as:

$$AE = x_{ij} - \hat{x}_{ij} \tag{6}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_{ij} - \hat{x}_{ij})^2 \tag{7}$$

In both cases, the error function ( $E$ ) is directly dependent on the weight component ( $W$ ) which in turn influences the learning rate ( $\eta$ ). This is updated or changes iteratively during gradient descent as:

**Table 1: Training parameters and network structures showing error estimation during training**

	MAMJ		JAS		OND	
	Training set	Test set	Training set	Test set	Training set	Test set
Nr. of data set	63	13	47	9	46	9
AE	0.005	0.005	0.003	0.006	0.015	0.02
MSE	4.91E-05	2.35E-05	2.71E-05	7.42E-05	0.0003	0.0004
Error tolerance (%)	10	30	10	30	10	30
Nr. of good forecasts (%)	61(98%)	13(100%)	46(98%)	9(100%)	28(61%)	7(78%)
Nr. of bad forecasts (%)	2(3%)	0(0%)	1(2%)	0(0%)	18(39%)	2(22%)
r <sup>2</sup>	0.994		0.996		0.997	
No. of hidden layer (s)	2		1		1	
No. of input layer (s)	1		1		1	
No. of output layer(s)	1		1		1	
Best at iteration number	1363		1625		1008	
Learning rate $\eta$ till global minimum	0.0040		0.0021		0.0062	

$$\dot{\mathbf{A}}\mathbf{W} = \eta \frac{\partial \mathbf{E}(\mathbf{W})}{\partial \mathbf{W}} \quad (8)$$

The smaller the error function the better the prediction during the training process. A minimum of five training runs were done on the same data set to obtain the best MSE. Thereafter, the neural network was perceived to have learned and could then be used for making predictions for unknown data. Training parameters like number of hidden layers, stopping condition, iterations number, learning rate and generalization loss were estimated on trial-and-error basis for each dataset.

Figure 3 compares the neural accuracy in terms of MSE for both actual and predicted values during training. The results show that both linear and nonlinear estimations had excellent abilities to forecast the seasonal rainfall amounts in the time series between 1997-2015.

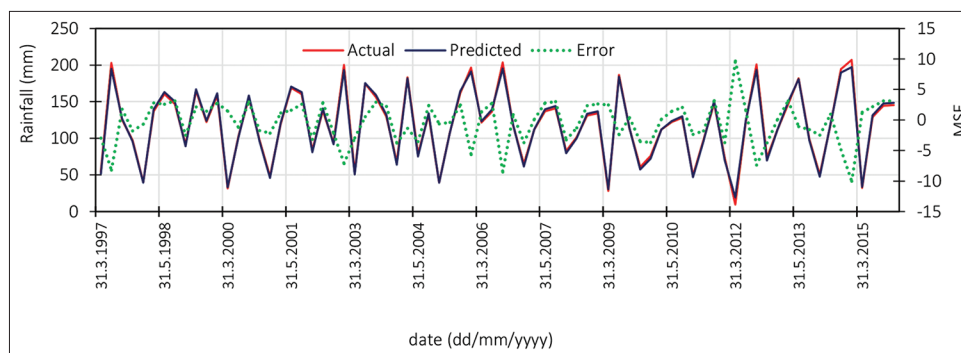
The average regression coefficient in the linear method was about ( $r^2=0.99$ ). Although the actual and forecasted data during training seemed to give a high correlation and ( $r^2$ ), it still showed some amount of error or noise for most part of the seasonal dataset trained. These errors were within the tolerance range put forth by the software. Hereby, neural model for rainfall forecasting may be assumed to be probabilistic and containing both deterministic as well as random error components. Therefore, linear representation of actual and predicted values as well as nonlinear methods presented here (Figure 3 and 4) during training of datasets can be regarded as good tools in ensuring neural accuracy in forecasting seasonal rainfall patterns.

## RESULTS AND DISCUSSION

### Neural network performance

The trained JAS with smaller dataset ( $n=56$ ) and single hidden layer appeared to outperform the MAMJ dataset with

larger dataset ( $n=76$ ) and two hidden layers demonstrating the difference in performance as influenced by data size. The variances for MAMJ ( $\sigma^2 = 0.252$ ), JAS ( $\sigma^2 = 0.332$ ) and OND ( $\sigma^2 = 0.345$ ) were 0.01, 0.006 and 0.07 respectively. Comparing the variance effects on all datasets, there was a notable difference on learning especially of JAS and OND datasets with similar data size. The MAMJ and OND were characterized by high standard deviation ( $\sigma = 0.266$ ) and ( $\sigma = 0.1$ ) respectively, whereas for JAS, this was  $\sigma = 0.076$ . However, all training sets achieved high relative coefficients ( $r^2 \geq 0.99$ ) with number of good forecasts over 60%. Although the neural network is suitable for characterizing non-linear relations, the findings here also show the capability of neural network in characterizing linear processes. Similar finding was reported by Zhang (1998). Figure 2 shows a plot of MSE and AE vs iterations during training for the MAMJ, JAS and OND datasets for the years 1997-2015. The accuracy of both error function estimates showed steep gradients prior to 1000-th iteration till to convergence at global minima. During training of the MAMJ dataset for example, there was a sharp decrease of the MSE from about 0.016 to as low as 0.0007 while for the AE this was between 0.1 to 0.016. It is seen that both error functions were large at low iteration values decreasing till convergence and subsequently increasing with further iterations. The AE and MSE during training for OND was ten-fold larger that of either MAMJ or JAS. The learning rate ( $\eta$ ) as measured by number of iterations to reach global minimum is fastest for OND at 1008 than, for JAS and MAMJ at 1625 and 1363 respectively. Low iteration number for OND would suggest, that the stochastic gradient descent algorithm effected larger step size parameter with large errors. This accounted for faster and poor learning rates and therefore, poor generalization. Conversely, smaller step sizes with smaller gradients resulted into larger number of iterations and comparatively lesser errors and better generalization. Better generalization was manifested by the comparatively higher percentage of good forecasts in the training sets of both MAMJ and JAS. Further, the rate of change of the error function  $\Delta \mathbf{W}$



**Figure 3:** Error estimation between the actual and predicted mean rainfall amounts during training of time series data between 1997 to 2016.

as in Eq. (8) had a significant effect on the performance and accuracy of the ANN.

Table 1 shows the training parameters and accuracy according to the  $r^2$ , number of good forecasts, hidden layer(s), MSE, and AE. The network demonstrated better performance rate for MAMJ and JAS when using two and one hidden layers respectively. The results indicate that model performance in terms of the number of good forecasts (98%) and approximation during validation for both seasons was independent of the number of hidden layers. In effect, one hidden layer performed just as good as two layers. Similar results on neural performance with a single hidden layer were reported by (Sonntag, 1992; Christiansen *et al.*, 2004; Mahmoud *et al.*, 2007; Nakama, 2011; Lolli *et al.*, 2016). Although both JAS and OND trained datasets had each one hidden layer and almost equal data size, the latter gave a low number of good forecasts at 61% and high number of bad forecasts (39%).

Generally, better accuracy was shown by both MAMJ (2 hidden layers) and JAS (one hidden layer) trained data with a ten-fold less error than that of OND. However, training in terms of number of iterations needed for convergence, percentage of good and bad forecasts was observed in JAS dataset with one hidden layer, while the second best was MAMJ with two hidden layers. On the other hand, OND showed the highest inaccuracy with one hidden layer. This could be due to the inability to learn from a small dataset, although the learning logistic regression algorithm for JAS with similar data size seemed to work well. Similar observations were reported by Forman and Cohen (2004), Shaikhina and Khovanova (2017).

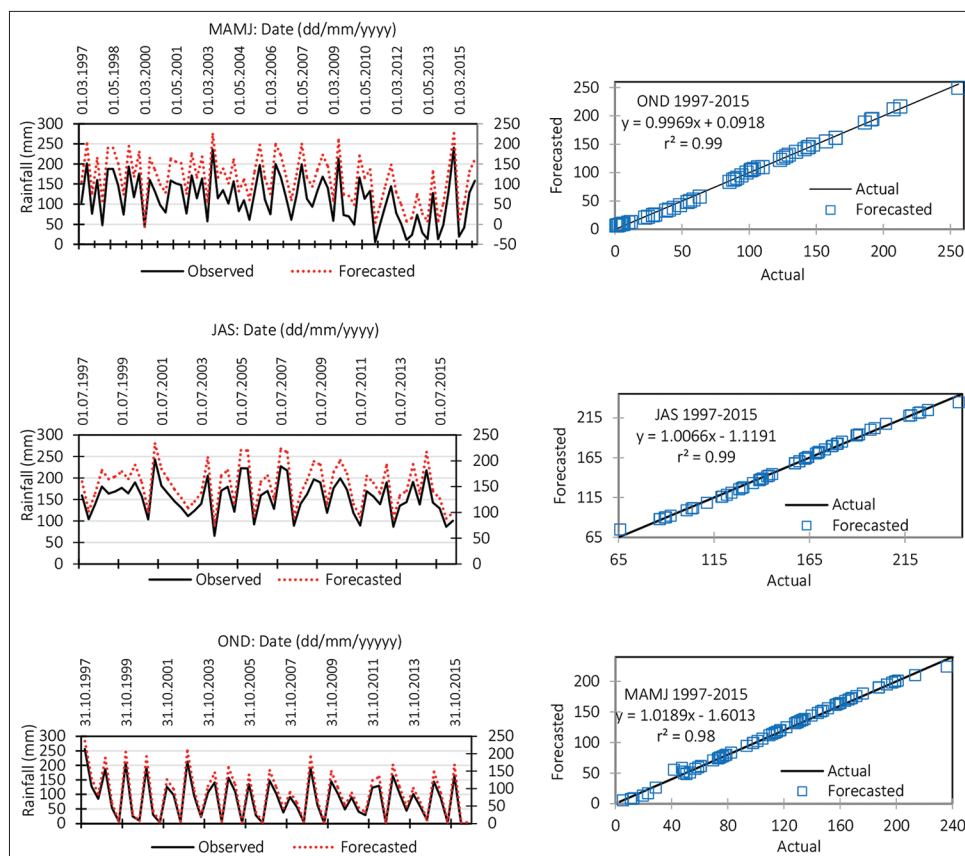
Such conflicting generalizations in terms of the number of hidden layers for MAMJ-JAS as well as for JAS-OND on model performance and accuracy indicate striking instability especially for smaller datasets. For instance, using one hidden layer, the OND dataset had a learning rate of 0.0062 and reached the global minimum at lower iterations than JAS at 0.0021. After that, the error functions AE and MSE started to increase indicating that the model was getting over-fitted. Moreover, the MAMJ dataset with two hidden layers had a learning rate at 0.004 (Table 2) and was comparatively lower

than that of OND but greater than that of JAS dataset with one hidden layer. Conventionally, neural learning in such an architecture with single hidden layer and finite number of neurons that approximate *continuous functions* may be described as “*shallow*” whereas with two or more hidden layers as “*deep*”. In our study, the single hidden layer as compared to two hidden layers had the best predictions during training.

Therefore, one can argue, that the learning rate during gradient descent is inversely related to the number of iterations in reaching a global minimum. Judging by the rule-of-thumb in estimating the number of neurons in the hidden layer(s), our study showed that this was between 105 and 210 neurons for one and two hidden layers respectively for MAMJ dataset, whereas these were 99 and 100 neurons for OND and JAS datasets respectively. Despite such striking inconsistency between the JAS and OND datasets with the single hidden layer, the accuracy and generalization performance of the two-layer feedforward neural network model was satisfactory. With the error tolerance (%) as indicator for overall performance, the results demonstrate that, this model was able to achieve remarkable performances on predictive tasks with limited data size as in MAMJ and JAS datasets, but unable to perform well on smaller datasets as in OND.

### Trends in mean seasonal rainfall and SK

Trained rainfall data for the MAMJ, JAS and OND months from 1997-2015 were used to forecast the mean rainfall over the period 2016-2034. Obtained results are shown in Figure 4. The SK test ( $S = -31.7$ ;  $Z = -0.774$  and  $Q = -0.029$ ) showed a negative monotonic trend and statistically significant at  $p < 0.01$ . The mean rainfall at the start of measurement for the MAMJ in 1997 was about 125 mm with about 5-10 mm reduction in 2015. This was a mean rainfall reduction of approximately 0.278 to 0.556 mm/year. Model projections from 2016-2034 forecasted a near 18% decrease in mean rainfall to about 100 mm. The total MAMJ rainfall reduction for Juba county between 1997 to 2034 is projected to be close to 32 mm. The JAS months also showed a decrease in the mean rainfall amount towards the end of 2034 forecasting period. The SK test ( $S = -11.71$ ;  $Z = -1.901$ ;  $Q = -0.234$ ) also showed a negative monotonic trend and statistically significant at  $p < 0.01$ .



**Figure 4:** Scatter plots showing the correlation between the actual and forecasted data during training of neural network.

This was a slight 6-7% decrease from about 160 mm in 1997 to about 100 mm in 2034. The Theil-Sen slope  $Q$  forecasted an annual drop of 0.2 mm/year which on average would be close to 7.64 mm as from 1997 to 2034. The OND also showed similar negative monotonic trend with SK test ( $S = -13.7$ ;  $Z = -0.774$ ;  $Q = -0.220$ ). There was a 21% decrease from about 95 mm in 1997 to about 75 mm forecast in 2034. The estimated Theil-Sen slope  $Q$  value of 0.100 mm/year would be about 3.8 mm in 2034. Trend analysis in of all the seasons revealed a general decrease of rainfall in Juba County with highest decrease during the MAMJ and lowest in OND (Figure 5). The results of the mean seasonal rainfall time series by the Mann-Kendall test and Sen's slope estimator both confirm and reinforce this monotonic negative trend.

### Changing rainfall patterns and impacts on crop production

Figure 6 shows the anticipated decline in the amount of mean rainfall at the onset of rain during the MAMJ. The onset rains varied between the 4<sup>th</sup> dekad of February and 1<sup>st</sup> dekad of March with daily rainfall values generally below 4.0 mm level. The March rainfall amounts locally termed as '*doko kulunyit* (that which carries away grass cinders after burning) are barely enough for any effective land preparation and planting. Thus, most farmers tend to shift their land preparation and planting dates toward the 3<sup>rd</sup> and 4<sup>th</sup> dekad of April. Most farmers

plant cowpeas (*ngete*), amaranth (*kwedekwede*), jute mallow (*mulukhiya/khudra*), okra (*bamia*) whose short growing and maturity periods (from 21 to 70 days) often offers best food security options prior to the onset of the longer rainy JAS season. Increasing inter-seasonal rainfall variability with declining mean rainfall amounts during MAMJ is forecasted to continue, thus much crop production will have to be shifted toward the 4<sup>th</sup> dekad of April or 1<sup>st</sup> dekad of May while for maize, sorghum, sesame will have to be grown during the JAS to OND season. Mean onset rainfall amounts in 2018 is expected to be around 65 mm with a 25% probability. With declining amounts of the onset rains, there is need to intensify inter-cropping of fast and slow growing crops during the MAMJ-JAS seasons as much time, energy and water resources can effectively be utilized. These findings corroborate similar studies by Rowell et al. (2015) on declining rainfall trend in the March-May rains within the East African region.

### Future seasonal rainfall projections

Figure 6 shows the observed and forecasted trend from 1997-2034 for the onset rains during the months of March and April. The results showed that there was a decreasing trend (red line) of about 8.8 mm per decade between 1997 to 2017 and about 5.4 mm between 2018-2042 accounting to about 14.2 mm reduction over this entire period. For the onset rainfall, this decrease in the next 60 years is forecasted to remain slightly below the average normal range. The

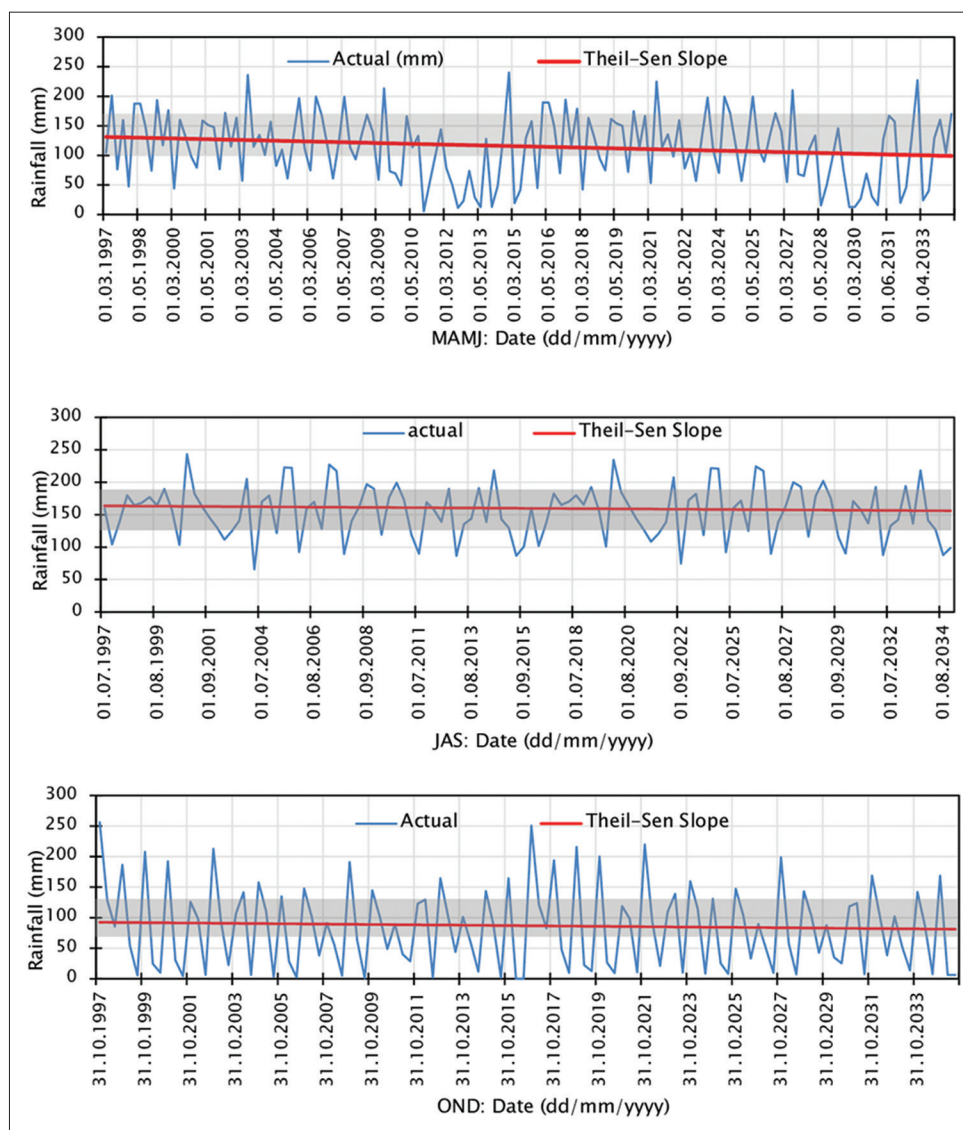


Figure 5: Trendline showing the Theil-Sen slope between 1997-2034 for the MAMJ, JAS and OND months.

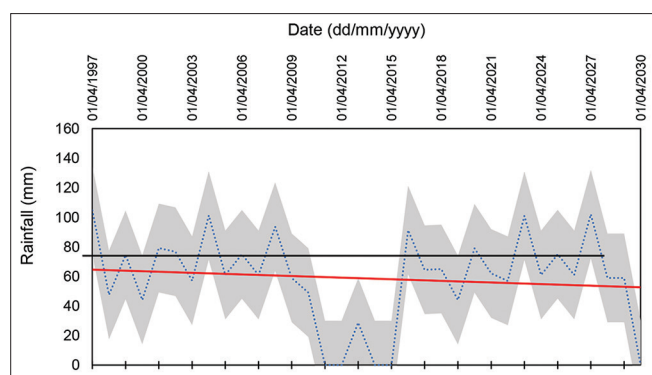


Figure 6: Observed and projected mean rainfall amounts during onset at first dekad of March (70 days from start of each new year).

reason is due to the low mean rainfall during OND coupled with high daily temperatures around this period that often continues into the MAMJ months prior to the onset of the first rainfall.

Manufacturing industries and large-scale agricultural farming are practically non-existent in South Sudan and therefore, CO<sub>2</sub> or methane emissions due to “*anthropogenic compulsions*” are unlikely to be the causes for spatially regionalized temperature increase and so changing rainfall patterns. However, the increased burning of fossil fuel, indiscriminate cutting down of forest trees as cheap energy source (Lomeling *et al.*, 2016) over the last 50 years suggest a possible anthropogenic cause for the increase of dry events and thus, reduced rainfall mean over Juba County. The results of our study on the negative trend of the onset rainfall also suggests that there’s an overall shift of the soil moisture drought from near to normal towards more moderate drought events. Conversely, such drought events are coupled with occasional severe wetness characterized by flooding events especially during September-October months as anthropogenic compulsions and the unpredictable effects of climate change continue to increase.

No conclusive reasons are attributable to the reduction in the mean seasonal rainfall amounts. However, the effects of global



warming exacerbated by *El Niño Southern Oscillation* (ENSO) (Fer *et al.*, 2017) on rainfall patterns at the regional level may have occurred, but this could not be statistically identified and verified within the available historical rainfall data and time series.

## CONCLUSION

Time series rainfall data from 1997 to 2015 were trained, tested and used to make 3-months ahead forecast. The performance of the FFNN model based on the MSE, degree of tolerance as well as the number of good forecasts during training and testing suggested that this model was accurate and therefore a versatile tool in the seasonal rainfall forecasts. Rainfall projection to year 2034 using the FFNN showed that there was negative monotonic trend significant at  $p < 0.01$  for the MAMJ, JAS and OND months with rainfall amounts varying between 5-12% below seasonal averages. There was also decreasing trend of the average onset rainfall amounts with much events occurring towards the end 3<sup>rd</sup> and 4<sup>th</sup> dekad of April and in other cases until the 1<sup>st</sup> dekad of May. This may significantly affect the timing for land preparation and subsequently planting. Future rainfall projection also showed a decreasing trend in all the seasons or months with values forecasted to remain within the near normal range for JAS and OND months while for MAMJ forecasted to have moderate drought in the next 100 years. Rainfall amounts during these seasons are expected to be slightly below the seasonal averages at less than 60, 100 and 10 mm for MAMJ, JAS and OND respectively. National and state governments as well as development partners will be urged to prepare contingency and intervention plans that could quickly and timely be implemented to avert any disruptions to crop production.

However, challenges by the application of FFNN model in projecting spatial and temporal rainfall patterns especially on shorter hourly and daily time scales persist. Understanding rainfall variability and intensity on hourly and daily basis within Juba County would increase the capacities and readiness of all stakeholders to timely and adequately respond to uncertainties arising from erratic rainfall patterns due to climate change. This paper recommends further studies to investigate whether such seasonal projections of rainfall can be corroborated with empirically measured rainfall amounts from several spatially placed stations within the county.

## ACKNOWLEDGEMENT

The authors would indeed like to express their sincere gratitude to the National Oceanic and Atmospheric Administration (NOAA) of the US Department of Commerce for its provision and allowing free access to its servers with weather data on South Sudan.

## REFERENCES

- Abbot, J., and Marohasy, J. (2017). Skilful rainfall forecasts from artificial neural networks with long duration series and single-month optimization. *Atmospheric Research*, 197, 289-299. <https://doi.org/10.1016/j.atmosres.2017.07.015>
- Badr, H. S., Zaitchik, B. F., and Guikema, S. D. (2014). Application of statistical models to the prediction of seasonal rainfall anomalies over the Sahel. *Journal of Applied Meteorology and Climatology*, 53(3), 614–636. <https://doi.org/10.1175/JAMC-D-13-0181.1>
- Benmahdjoub, K., Amour, Z., and Boulifa, M. (2013). Forecasting of rainfall using time delay neural network in Tizi-Ouzou (Algeria). *Energy Procedia*, 36, 1138–1146. <https://doi.org/10.1016/j.egypro.2013.07.129>
- Christiansen, N. H., Voie, N. H., Winther, P. E. T., and Høgsberg, O. J. (2014). Comparison of Neural Network Error Measures for Simulation of Slender Marine Structures. *Journal of Applied Mathematics*, 759834. <https://doi.org/10.1155/2014/759834>
- Devi, S. R., Arulmozhivarman, P., Venkatas, C., and Pramay, A. (2016). Performance comparison of artificial neural network models for daily rainfall prediction. *International Journal of Automation and Computing*, 13(5), 417-427. <https://doi.org/10.1007/s11633-016-0986-2>
- Elsanabary, M. H., and Gan, T. Y. (2014). Wavelet analysis of seasonal rainfall variability of the Upper Blue Nile Basin, its teleconnection to global sea surface temperature, and its forecasting by an artificial neural network. *Monthly Weather Review*, 142(5), 1771–1791. <https://doi.org/10.1175/MWR-D-13-00085.1>
- Fer, I., Tietjen, B., Jeltsch, F., and Wolff, C. (2017). The influence of El Niño–Southern Oscillation regimes on eastern African vegetation and its future implications under the RCP8.5 warming scenario. *Biogeosciences*, 14, 4355–4374. <https://doi.org/10.5194/bg-14-4355-2017>
- Forman, G., and Cohen, I. (2004). Learning from little: comparison of classifiers given little training. In J. F. Boulicaut, F. Esposito, F. Giannotti and D. Pedreschi (Eds.), *Knowledge Discovery in Databases: PKDD 2004. Lecture Notes in Computer Science* (Vol. 3202), Berlin, Heidelberg: Springer. [https://doi.org/10.1007/978-3-540-30116-5\\_17](https://doi.org/10.1007/978-3-540-30116-5_17)
- Gilbert, R. O. (1987). *Statistical Methods for Environmental Pollution Monitoring*. New York: Wiley.
- Hartmann, H., Snow, J. A., and Su, B. (2016). Seasonal predictions of precipitation in the Aksu-Tarim River basin for improved water resources management. *Global and Planetary Change*, 147, 86–96. <https://doi.org/10.1016/j.gloplacha.2016.10.018>
- Haviluddin, M., Hardwinarto, S., and Sumaryono, M. A. (2015). Rainfall Monthly Prediction Based on Artificial Neural Network: A Case Study in Tenggara Station, East Kalimantan – Indonesia. *Procedia Computer Science*, 59, 142 – 151. <https://doi.org/10.1016/j.procs.2015.07.528>
- Helsel, D. R., and Hirsch, R. M. (1995). *Statistical Methods in Water Resources*. New York: Elsevier.
- Hirsch, R. M., Slack, J. R., and Smith, R. A. (1982). Techniques of Trend Analysis for Monthly Water Quality Data. *Water Resources Research*, 18(1), 107-121. <https://doi.org/10.1029/WR018i001p00107>
- Hung, N. Q., Babel, M. S., Weesakul, S., and Tripathi, N. K. (2009). An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrology and Earth System Sciences*, 13(8), 1413–1425. <https://doi.org/10.5194/hess-13-1413-2009>
- Li, B., Wang, L., Kaseke, K. F., Li, L., and Seely, M. K. (2016). The Impact of Rainfall on Soil Moisture Dynamics in a Foggy Desert. *PLoS ONE*, 11(10), e0164982. <https://doi.org/10.1371/journal.pone.0164982>
- Lolli, F., Gamberini, R., Regattieri, A., Balugani, E., Gatos, T., and Gucc, S. (2016). Single-hidden layer neural networks for forecasting intermittent demands. *International Journal*

- of *Production Economics*, 183(A), 116-128. <https://doi.org/10.1016/j.ijpe.2016.10.021>
- Lomeling, D., Modi, A. L., Moti, S. K., Kenyi, M. C., Silvestro, G. M., and Yieb, J. L. L. (2016). Comparing the Macro-aggregate Stability of Two Tropical Soils: Clay Soil (*Eutric Vertisol*) and Sandy Loam Soil (*Eutric Leptosol*). *International Journal of Agriculture and Forestry*, 6(4),142-151.
- Mahmoud, O., Anwar, F., and Jimoh, E. S. M. (2007). Learning Algorithm Effect on Multilayer Feed Forward Artificial Neural Network Performance in Image Coding. *Journal of Engineering Science and Technology*, 2(2), 188–199.
- Nakama, T. (2011). Comparisons of Single- and Multiple-Hidden-Layer Neural Networks. In D. Liu, H. Zhang, M. Polycarpou, C. Alippi and H. He. (Eds.), *Advances in Neural Networks – ISNN 2011*, Lecture Notes in Computer Science (Vol. 6675), Berlin, Heidelberg: Springer. <https://doi.org/10.1007/978-3-642-21111-9>
- Nielsen, M. A. (2015). *Neural Networks and Deep Learning*. Determination Press.
- Rowell, D. P., Booth, B. B. B., Nicholson, S. E., and Good, P. (2015). Reconciling past and future rainfall trends over East Africa. *Journal of Climate*, 28(24), 9768–9788, <https://doi.org/10.1175/JCLI-D-15-0140.1>
- Salmi, T., Maatta, A., Anttila, P., Ruoho-Airola, T., and Amnell, T. (2002). *Detecting Trends of Annual Values of Atmospheric Pollutants by the Mann-Kendall Test and Sen's Slope Estimates -The Excel Template Application MAKESENS*. Helsinki, Finland: Finnish Meteorological Institute.
- Shaikhina, T., and Khovanova, N. A. (2017). Handling limited datasets with neural networks in medical applications: A small-data approach. *Artificial Intelligence in Medicine*, 75, 51-63. <https://doi.org/10.1016/j.artmed.2016.12.003>
- Singh, P. (2018). Indian summer monsoon rainfall (ISMUR) forecasting using time series data: A fuzzy-entropy-neuro based expert system. *Geoscience Frontiers*, 9(4), 1243-1257. <https://doi.org/10.1016/j.gsf.2017.07.011>
- Singh, P. (2018). Indian summer monsoon rainfall (ISMUR) forecasting using time series data: A fuzzy-entropy-neuro based expert system. *Geoscience Frontiers*, 9(4), 1243-1257. <https://doi.org/10.1016/j.gsf.2017.07.011>
- Sonntag, E. D. (1992). Feedback stabilization using two-hidden-layer nets. *IEEE Transactions on Neural Networks*, 3(6), 981–990. <https://doi.org/10.1109/72.165599>
- Warsito, B., Gernowo, R., and Sugiharto, A. (2016). Rainfall prediction by using wavelet general regression neural network. *International Journal of Applied Mathematics and Statistics*, 54(3), 32–41.
- Zhang, G. (1998). *Linear and nonlinear time series forecasting with artificial neural networks*. Doctoral Dissertation, Kent State University, Kent, OH.