

Time Series Forecasting with Statistical Neural Network using Continuous Wavelet Decomposition

C. G. Udomboso^{a*} and G. N. Amahia^b

^{a,b}Department of Statistics, University of Ibadan, Ibadan, Nigeria

We propose a forecasting model based on the combined efficiency of the artificial neural network and wavelet transform in modeling time series data. The data used were decomposed into continuous wavelet signals on a scale of 10. Each of the decomposed series was subjected to correlation test with the original data. We compare the new model's performance with the conventional time series regression model (TSRM) and the wavelet neural network (WNN) forecasting model. The WNN model performed better than the TSRM. The analysis also showed that except in extremely rare cases, all the wavelet series performed optimally compared to the original data.

Keywords: Wavelet; neural network; activation function; forecasting.

1. Introduction

The usage of wavelet transforms in statistical models has come a long way. These infusions have made results of findings appealing. For example, Pandhiani and Shabri (2013) explored the use of least square support vector and wavelet technique in monthly stream flow forecasting, using data from the Jhelum and Chenab rivers in Pakistan, over a 30-day period. Results show that using wavelet alongside the other methods discussed was more accurate. Recently, Mishra et al (2015) conducted a review of several studies of time series data mining for real time hydrological forecasting in an attempt to solve flooding problems. In their review, they identified such tools and techniques as regression analysis, clustering, artificial neural network (ANN), support vector machines, Genetic Algorithm, fuzzy logic, and rough set theories.

In the same vein, Adamowski and Chan (2011) proposed a new method based on coupling discrete wavelet transforms and ANNs for groundwater level forecasting applications from data obtained from the Chateau-guay watershed in Quebec, Canada. Their method was found to be more accurate than the regular ANN and autoregressive integrated moving average. Moreover, Krishna et al (2011) conducted a time series modeling of river flow from daily flow data of the Malaprabha river basin in Karnataka state of India using wavelet neural network (WNN). The WNN was found to provide a better fit than the individual ANN and autoregressive models. Satyajai Rao et al (2014), in their case, proposed a wavelet based neural networks analysis for daily stream flow forecasting in three river basins in India, namely, Kollur, Sitanadi, and Verahi. Apart from the effectiveness of the WNN over the ANN, the WNN model performance was also evaluated separately for low, medium, and high runoff values and found to be suitable for various runoff patterns from catchment.

In another attempt, Kilby and Prasad (2013) used the continuous wavelet analysis and classification of surface electromyography signals for normal muscle activity (SEMG). Their method has been shown to be sound and successful for the basis of implementation for developing and intelligent SEMG signal classifier.

We note here that these works were conducted without reference to forecasting for the purpose of predicting future occurrence of hydrological problems. Moreover, much work has not been done in time series analysis using a combination of wavelet and ANN. Hence, this study undertakes to bridge the gap, and contribute to knowledge on the application of

*Corresponding author. Email: cg.udomboso@gmail.com

wavelet and ANN in forecasting time series data. Forecasting or prediction is the process of estimating the future in order to take decisions. This is based on the analysis of some factors that are believed to influence the future values, or on the study of the past data behavior over time. Modeling and forecasting have applications in several time series areas such as marketing, finance, telecommunications, organizational behavior, sales history, foreign currency risk or stock market volatility, network traffic prediction, and so on.

Qiu and Rong (2008) noted numerous existing models for time series forecasting, which can be grouped into four categories. These include:

- (1) case-based reasoning (CBR): is a means for solving a new problem by using or adapting solutions to old problems its essence in analogy. The basic principle of CBR is that similar problems have similar solutions.
- (2) rule based forecasting: its application depends upon features of the time series
- (3) statistical models: exploit historical data. They contain early traditional models such as the single regressive model, exponential smoothing, ARIMA model.
- (4) soft computing models: such as neural networks and their amelioration or mixture with other methods.

It was shown by Armstrong and Adyaa (2000), and Mitra and Mitra (2006), that among these forecasting models, Artificial Neural Network (ANN), whose statistical form is known as the Statistical Neural Network (SNN) have been shown to produce better results. Also, Tan (1993) proved the advantage of the artificial neural networks over traditional rule-based systems. Several distinguishing features of ANNs make them valuable and attractive for a forecasting task. First of all, as reported by Zhang et al (1998), White (1989), and Ripley (1993), they can be treated as multivariate nonlinear nonparametric statistical methods. Furthermore, Zhang et al (1998) opined that ANNs can generalize, and are known as universal functional approximators. In a series of studies, Irie and Miyake (1988), Hornik et al (1989), Funahashi (1989), and Hornik (1993) showed that a network can approximate any continuous function to any desired accuracy.

Finally, ANNs are nonlinear. Zhang et al (1998) noted that the traditional approaches to time series prediction, such as Box-Jenkins or ARIMA method, assume that the time series are generated from linear processes, but they may be totally inappropriate if the underlying mechanism is nonlinear.

While Wu (1995) applied multilayer feed-forward neural networks to forecast Taiwan Dollar/US Dollar exchange rate and the results are better than ARIMA models, and Sallehuddin et al (2007) proposed a GRANN-ARIMA hybrid model which combined non-linear Grey Relational Artificial Neural Network and linear ARIMA model for time series forecasting. On the other hand, Zhang (2003) presented a hybrid forecasting model that combines both ARIMA and neural network approaches, while Tenti (1996) compared the forecasting ability of three different recurrent neural network architectures and then devised a trading strategy based on the forecasting results to optimize the profitability of the system. In their own case, Chen and Abraham, (2006) used a Flexible Neural Tree(FNT) model for forecasting three major international currency exchange rates, namely US Dollar/Euro, US Dollar/British Pound and Japanese Yen/US Dollar. In another study, Lai et al (2006) proposed a multistage neural network meta-learning method for forecasting US Dollar/Euro exchange rate, while Dunis and Williams (2002) examined the use of Neural Network Regression (NNR) models in forecasting and trading the US Dollar/Euro exchange rate. In a study by Kaashoek and van Dijk (2002), the size of neural networks was reduced by using multiple correlation coefficients, principal component analysis and graphical analysis techniques.

Some other authors further introduced the use of wavelet signals into neural networks in order to produce better results. Wavelet neural network have been widely-used for forecasting. For example, Yousefi and Reinarz (2005) used it in the forecast of oil prices, while Aussem and Murtagh (1998) employed it to forecast stock index. And on the other hand, Fay and Ringwood, (2007), Zhang and Coggins (2001), as well as Zhang and Dong (2001)

used it to forecast electricity demand and other time series problems. There are not many applications of wavelet neural network for exchange rate forecasting. Tan (2009) opined that the wavelet neural network is especially suitable for forecasting exchange rates, because according to Ramsey (2002) wavelets can "decompose economic time series into their time scale components" and this is "a very successful strategy in trying to unravel the relationship between economic variables." Previous research by Kumar and Joshi (2003) and Gradojevic and Yang (2000) has shown that using economically significant variables for inputs can improve the performance of the forecasting model.

This study therefore combines wavelet transforms with statistical neural network exchange rate to achieve accurate forecasting.

2. Materials and Methods

2.1 Wavelets

Wavelets, a class of functions used to localize a given function in both position and scaling, are used in applications such as signal processing and time series analysis, and form the basis of the wavelet transforms which, as Dr I. Daubechies commented, "cuts up data of functions or operators into different frequency components, and then studies each component with a resolution matched to its scale" (Veitch 2005). In the context of signal processing, wavelet transform depends on two variables. These are scale (also known as frequency) and time.

A wavelet is a 'small wave' function, usually denoted $\psi(\cdot)$. a small wave grows and decays in a finite time period, as opposed to a 'large wave', such as the sine wave, which grows and decays repeatedly over an infinite time period.

For a function $\psi(\cdot)$, defined over the real axis $(-\infty, \infty)$, to be classed as a wavelet, it must satisfy the following three properties:

- (i) the integral of $\psi(\cdot)$ is zero

$$\int_{-\infty}^{\infty} \psi(u) du = 0$$

- (ii) the integral of the square of $\psi(\cdot)$ is unity

$$\int_{-\infty}^{\infty} \psi^2(u) du = 1$$

- (iii) the admissibility condition

$$C_{\psi} = \int_0^{\infty} \frac{|\Phi(f)|^2}{t} df$$

satisfies $0 < C_{\psi} < \infty$.

The wavelet transform is given by

$$W_x(a, \tau) = \frac{1}{\sqrt{|a|}} \int f(t) \psi\left(\frac{t - \tau}{a}\right) dt$$

where a and τ are scale and location parameters respectively.

By adjusting the scale parameter, a , a series of different frequency components in the signal can be obtained.

2.2 Continuous wavelet transform

The Continuous Wavelet Transform (CWT) is used to transform a function or signal $x(\cdot)$ that is defined over continuous time. Hence, the parameters λ and t used for creating the wavelet family both vary continuously. A fundamental fact about CWT is that it preserves all the information from $x(\cdot)$, the original signal, and is recoverable using the inverse transform.

2.3 Wavelet neural networks

Wavelet Neural Networks (WNNs) combines the theory of wavelets and neural networks into one. The structure of WNN is very similar to that of a feed-forward neural network, taking one or more inputs, with one hidden layer and whose output layer consists of one or more linear combiners or summers. The hidden layer consists of neurons, whose activation functions are drawn from an orthonormal wavelet family. One application of WNN is that of function estimation. Given a series of observed values of a function, an WNN can be trained to learn the composition of that function, and hence calculate an expected value for a given input.

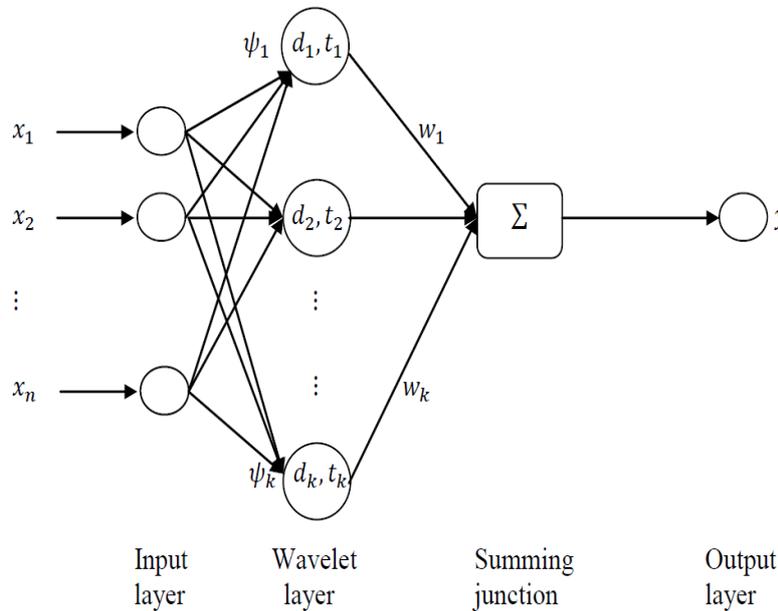


Figure 1. The structure of a typical wavelet neural network

In WNN model, the hidden neurons have wavelet activation functions of different resolutions, with weight, w_i , connecting the hidden and output layers. For an input vector, $x = [x_1, x_2, \dots, x_n]$, the output of the i th wavelet layer neuron is described as

$$\psi_k(x) = \sum_{i=1}^n \exp \left[- \left(\frac{x_i - d_k}{t_k} \right)^2 / 2 \right] \cos \left(5. \frac{x_i - d_k}{t_k} \right)$$

where x_i is the i th input vector, k is the number of wavelet node, d_k and t_k are translational and dilational parameters respectively (Wen et al 2009).

2.4 Activation function

The effectiveness of a neural network is hinged, mostly, on the activation function (also known as transfer functions). Debes et al (2005) studied the simulation of transfer functions (TFs) in ANNs by visualizing the input-output relation of an artificial neuron for a two dimensional input matrix in a three dimensional diagram. It was discovered that while different TFs and output functions can be chosen, several parameters such as weights and thresholds can be varied.

In Yonaba et al. (2010), three nonlinear transfer functions, bounded by -1 and 1 , were compared for neural network multistep ahead streamflow forecasting of the catchment of the San Juan river located on the Canadian Pacific coast. The transfer functions included the Elliot sigmoid, the bipolar (or logistic) sigmoid, and the tangent sigmoid functions. Their results endorsed the tangent sigmoid as the most pertinent transfer function for streamflow forecasting. However, the Elliot sigmoid was seen to require less computing time, which makes it a valuable option for operational hydrology.

Likewise, Dorofki et al (2012) compared the abilities of three transfer functions, namely, log sigmoid (logsig), tangent sigmoid (tansig), and purelin to simulate extreme runoff data in three catchment areas of the Johor river basin in Malaysia were investigated. The results indicated that the best transfer function (TF) was the logsig for the computation of minimum or normal runoffs. However, in the case of maximum rainfall data, the purelin proved to perform better.

In this work, the activation function, also known as transfer function, used is the tangent sigmoid function (a family of the sigmoidal functions), whose form is given as

$$tansig = f_3(x) = \frac{2}{1 - e^{(-2x)}} - 1$$

Sigmoidal functions have non-local behaviour, that is, they are non-zero in infinite domain. A sigmoid function is real-valued and differentiable, having non-negative or non-positive first derivative.

Sigmoid functions are often used in neural networks to introduce nonlinearity in the model. A neural network element computes a linear combination of its input signals, and applies a sigmoid function to the results. One reason for its popularity in neural networks is because it satisfies a property between the derivative and itself such that it is computationally easy to perform, and whose derivatives are usually employed in learning algorithm.

The transfer function we used in this work was one of those based on distance measures (some other groups of transfer functions are based on dot products). In this case, the weight vectors of neurons represent the data in the input space.

2.5 Estimation of wavelet neural networks

In estimating WNN, we minimize using the usual least-squares cost function:

$$E = \frac{1}{2} \sum_{j=1}^s (y_j - o_j)^2$$

where s is the number of estimation (training) samples for each class, and o_j is the optimal output of the j th input vector.

The partial derivative of the parameters d , t , and w are as follows:

$$\frac{\partial E}{\partial d_m} = \sum_{j=1}^s 2(y_j - o_j) \left[\sum_{m=1}^k w_m \exp \left(- \left(\frac{x - d_m}{t_m} \right)^2 / 2 \right) \left(\frac{x - d_m}{t_m^2} \right) \cos \left(5 \cdot \frac{x - d_m}{t_m} \right) + \right.$$

$$\begin{aligned}
 & \left. \frac{5}{t_m} \sin \left(5 \cdot \frac{x-d_m}{t_m} \right) \right] \\
 & = \sum_{j=1}^s 2(y_j - o_j) \left[\sum_{m=1}^k w_m \exp \left(-\frac{s_m^2}{2} \right) \frac{(s_m \cos(5s_m) + 5 \sin(5s_m))}{t_m} \right] \\
 \frac{\partial E}{\partial t_m} & = \sum_{j=1}^s 2(y_j - o_j) \left[\sum_{m=1}^k w_m \exp \left(-\left(\frac{x-d_m}{t_m} \right)^2 / 2 \right) \frac{x-d_m}{t_m^2} \left(\frac{x-d_m}{t_m} \right) \cos \left(5 \cdot \frac{x-d_m}{t_m} \right) + \right. \\
 & \left. \frac{5}{t_m} \sin \left(5 \cdot \frac{x-d_m}{t_m} \right) \right] \\
 & = \sum_{j=1}^s 2(y_j - o_j) \left[\sum_{m=1}^k w_m \exp \left(-\frac{s_m^2}{2} \right) \frac{s_m}{t_m} (s_m \cos(5s_m) + 5 \sin(5s_m)) \right]
 \end{aligned}$$

and

$$\frac{\partial E}{\partial w_m} = \sum_{j=1}^s \sum_{m=1}^k \psi_m 2(y_j - o_j)$$

where $s_m = \frac{x-d_m}{t_m}$. Adjustment is made for the parameters by the following equation:

$$\Theta^n = \Theta^{n-1} - \alpha \Delta \Theta$$

where $\Theta = (d, t, w)^r$ is the vector of the parameters d, t , and w , and α is the learning rate between 0.1 and 0.9.

2.6 Description of the proposed model

The step to the present model involved decomposition of the data for the study (see section 2.7) into wavelet signals. (The details of the signals are as presented in section 2.7). The signals were then set as input into the network, where, at the summing junction, were standardized with mean 0 and variance 1. The procedure utilise in this paper is a development on the model proposed by Song et al (2017). The activation function used in mapping the signals was the hyperbolic tangent sigmoid (*transig*) given as

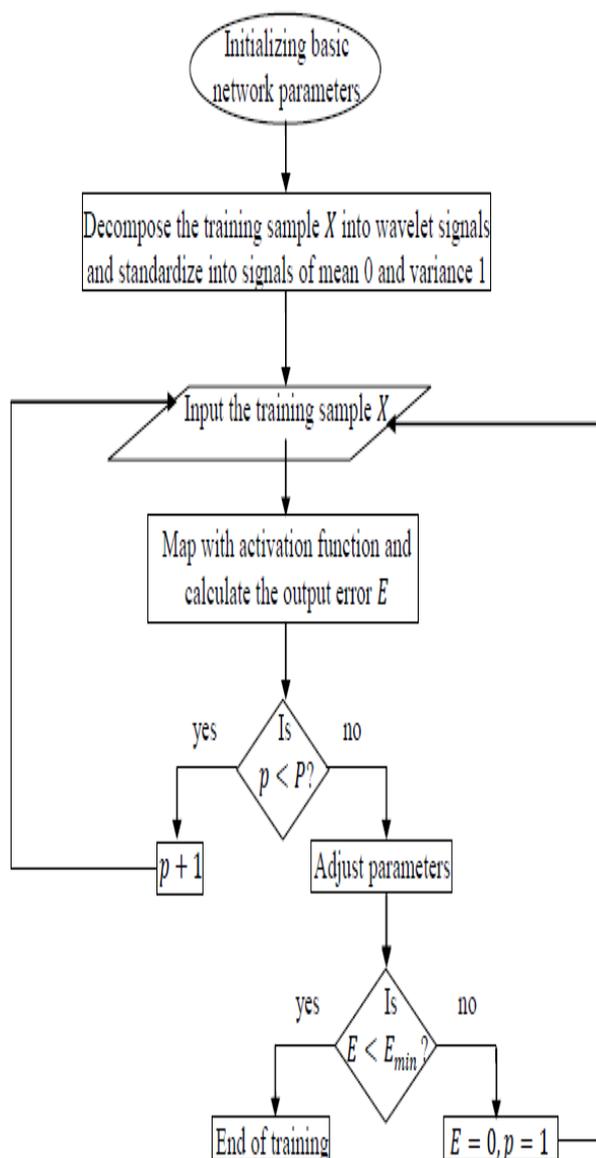
$$transig = f_3(x) = \frac{2}{1 - e^{-2x}} - 1$$

This activation function is used because it has been shown by Udomboso (2014) to be the best in mapping the input neuron to the output neuron. This verified the thoughts of several other researchers.

The flowchart in Figure 2 is a summary of the procedure in this paper.

2.7 Data for the study

The data used in this work is the annual naira US dollar exchange rate obtained from the 2012 Central Bank of Nigeria Bulletin for the last quarter of 2012. The exchange rate



Note: X is the naira-US dollar exchange rate, $p = 1, 2, \dots, P$ is the number of wavelet signals

Figure 2. Flowchart for the algorithm used in the study

variables considered are the buying rate (BR), selling rate (SR) and the central rate (CR). BR refers to the rate at which dollar is bought from a customer who wishes to exchange for dollar for naira, while SR refers to the rate at which the dollar is sold to a customer who wishes to exchange naira for dollar, and CR refers to the fixed exchange rate of the naira US dollar. The SNN model for each rate is denoted PBR, PSR, and PCR. Both the wavelet and the statistical neural network code were written in MATLAB 2009b to analyze the data. Each rate was decomposed into 10 continuous wavelet signals (W1, W2, W3, W4, W5, W6, W7, W8, W9, and W10). This gives a total of 120 points in each rate. Each signal, apart from the original data, was used as inputs into the network. The prediction from each signal was correlated with the original rate from where they were decomposed. Our interest is in showing that the models do not give the same results.

3. Results and Discussion

This section discusses the results of the analyses of the wavelet signals across the models. Results of buying rate in Table 1 shows stability in the mean prediction of the original

Table 1: Summary of results for buying rate

Model	Individual Model		Paired Differences (Actual vs Other Models)				
	Mean Prediction	Standard Error	Standard Error	Correlation Coefficient	P Value (Correl)	Test of Means	P Value (Test of)
OBR	132.49537	4.14688					
PBR_TSRM	132.49537	3.32095	2.48353	0.801	0.002	0	1
PBR	132.34085	3.95086	1.39204	0.942	0	0.111	0.914
PBRW1	134.88776	5.65221	2.7525	0.887	0	-0.869	0.403
PBRW2	133.86178	4.33388	1.58139	0.931	0	-0.864	0.406
PBRW3	134.41201	7.49693	4.49967	0.855	0	-0.426	0.678
PBRW4	147.76588	12.38872	9.02221	0.869	0	-1.693	0.119
PBRW5	126.81591	9.82977	7.68927	0.671	0.017	0.739	0.476
PBRW6	152.961	11.11307	9.62502	0.521	0.082	-2.126	0.057
PBRW7	147.98407	11.83671	1.9798	0.14	0.663	-1.293	0.223
PBRW8	165.68466	11.78165	2.77526	-0.074	0.82	-2.598	0.025
PBRW9	164.19873	12.04628	5.20976	-0.691	0.013	-2.084	0.061
PBRW10	181.51527	10.45894	3.53325	-0.652	0.022	-3.622	0.004

Table 2: Summary of results for selling rate

Model	Individual Model		Paired Differences (Actual vs Other Models)				
	Mean Prediction	Standard Error	Standard Error	Correlation Coefficient	P Value (Correl)	Test of Means	P Value (Test of)
OSR	133.56137	4.12807					
PSR_TSRM	133.56137	3.32783	2.44266	0.806	0.002	0	1
PSR	134.04742	3.82846	1.35489	0.945	0	-0.359	0.727
PSRW1	136.42056	5.39495	2.5456	0.891	0	-1.123	0.285
PSRW2	138.08818	5.82293	3.26375	0.838	0.001	-1.387	0.193
PSRW3	136.17584	6.09573	5.76609	0.416	0.178	-0.453	0.659
PSRW4	146.20227	7.009	5.15072	0.685	0.014	-2.454	0.032
PSRW5	144.4064	5.59145	8.01036	0.567	0.055	-1.354	0.203
PSRW6	157.18843	11.61902	0.86747	0.354	0.259	-2.174	0.052
PSRW7	151.43448	10.89118	1.34569	0.077	0.812	-1.575	0.143
PSRW8	167.33313	11.02256	1.89897	-0.033	0.918	-2.838	0.016
PSRW9	164.0892	11.01882	4.21983	-0.701	0.011	-2.147	0.055
PSRW10	182.4726	10.50165	3.48644	-0.629	0.028	-3.627	0.004

data, TSRM model and the ANN (PBR) models. The wavelet signals on the other hand shows an upward trend in the mean prediction across the signals, with some short, almost unnoticed nose-dives at signals 2, 5, and 8. Signals 7 and 8 are poorly correlated. Signals 1, 2, 3, 4, 5, 9, and 10 are well correlated and significant. The test results show that except for signals 8 and 10, others predictions are more precise than the original data set.

In Table 2, results show stability in the mean prediction in the selling rate of the original data and the TSRM models. The wavelet signals on the other hand also shows an upward trend in the mean prediction across the signals, with some short, almost unnoticed nose-dives at signals 3, 5, 7 and 9. Signals 3, 6, 7 and 8 are poorly correlated. Signals 1, 2, 4, 5,

9, and 10 are well correlated and significant. The test results show that except for signals 4, 8 and 10, others predictions are more precise than the original data set.

Table 3: Summary of results for central rate

Model	Individual Model		Paired Differences (Actual vs Other Models)				
	Mean Prediction	Standard Error	Standard Error	Correlation Coefficient	P Value (Correl)	Test of Means	P Value (Test of
OCR	133.44889	4.12078					
PCR_TSRM	133.44889	3.19588	2.60139	0.776	0.003	0	1
PCR	134.05465	3.78906	1.49979	0.931	0	-0.404	0.694
PCRW1	137.02318	5.31324	2.58875	0.879	0	-1.381	0.195
PCRW2	134.50127	4.40222	1.75186	0.918	0	-0.601	0.56
PCRW3	139.96738	5.17691	2.68042	0.858	0	-2.432	0.033
PCRW4	132.05498	6.95578	5.1177	0.683	0.014	0.272	0.79
PCRW5	152.71146	9.60853	9.20723	0.31	0.327	-2.092	0.06
PCRW6	140.03233	13.80432	2.89967	0.362	0.248	-0.51	0.62
PCRW7	160.93003	12.90907	3.63372	-0.021	0.948	-2.016	0.069
PCRW8	160.72853	11.85879	2.58113	-0.007	0.983	-2.168	0.053
PCRW9	166.82519	12.23253	5.35663	-0.687	0.014	-2.173	0.052
PCRW10	182.78726	10.71259	3.75286	-0.65	0.022	-3.587	0.004

Table 3 results show stability in the mean prediction in the selling rate of the original data and the TSRM models. The wavelet signals on the other hand also shows an upward trend in the mean prediction across the signals, with some short, almost unnoticed nose-dives at signals 2, 4, and 6. Signals 5, 6, 7 and 8 are poorly correlated. Signals 1, 2, 3, 4, 9, and 10 are well correlated and significant. The test results show that except for signals 3 and 10, others predictions are more precise than the original data set.

The forecast results and the signals are given in the Appendix. The forecast results for the buying rate in Figure 1 shows the possibility of all models converging at period 24, while at period 16, the ANN, WNN signal 1 (BRW1), and TSRM would likely converge. In Figure 2, forecast shows that all models have the tendency to converge for the selling rate period 17. While the central rate in Figure 3 would likely converge at period 30 using ANN, WNN signal 1 (CRW1), and TSRM.

Using each of the network models separately, signal 1 (Figure 4) shows the possibility of both the selling and central rate converging at period 23. However, the selling rate would remain on a spot from period 19 to 30. Signal 9 (Figure 5) shows that all rates might converge at period 13. The ANN (Figure 6) results reveal that all rates might converge at period 22.

4. Concluding Remarks

In this study, we set out to forecast time series data with the statistical neural networks using the continuous wavelet decomposition. The general outcome of the models show stability in the mean predictions across the buying, selling, and the central rates. The study shows the points at which the different rates would likely converge. Especially, the selling rate has the tendency to converge faster than the other two rates. These results could be useful given that the Nigeria exchange rate remains in a state of equilibrium as it is presently. We hope that future work would consider the present state of the Nigeria economy. In conclusion, efforts are ongoing in the development of a detailed statistical analysis of the wavelet neural networks which could lead to more informed decisions.

References

- Adamowski, J. and Chan, H.F. (2011) A Wavelet Neural Network Conjunction Model for Groundwater Level Forecasting. *Journal of Hydrology*, 407: 28–40.
- Armstrong, S. and Adyaa, M. (2000). An application of rule based forecasting to a situation lacking domain knowledge. *International Journal of Forecasting*, 16: 477–484.
- Aussem, J.C.A and Murtagh, F. (1998). Wavelet-based feature extraction and decomposition strategies for financial forecasting. *Journal of Computational Intelligence in Finance*, 6: 5–12.
- Chen, B.Y.Y. and Abraham, A. (2006). Optimal design of hierarchical wavelet networks for time series forecasting. *Proceedings of ESANN 2006*: 155–160.
- Debes, K., Koenig, A. and Gross, H. (2005). Transfer Functions in Artificial Neural Networks - A Simulation-Based Tutorial. A Supplementary Material for urn:nbn:de:0009-3-1515, Department of Neuroinformatics and Cognitive Robotics, Technical University Ilmenau, Germany. Retrieved from <http://www.brains-minds-media.org>.
- Dorofki, M., Elshafie, A.H., Jaafar, O., Karim, O.A. and Mastura, S. (2012). Comparison of Artificial Neural Network Transfer Functions abilities to simulate Extreme Runoff Data. *Proceedings of the 2012 International Conference on Environment, Energy and Biotechnology*, Singapore 2012, IPCBEE, 33: 39-44.
- Dunis, C.L. and Williams, M. (2002). Modelling and trading the EUR/USD exchange rate: Do neural network models perform better? *Derivatives Use, Trading & Regulation*, 8: 211-239.
- Fay, D. and Ringwood, J.V. (2007). A wavelet transfer model for time series forecasting. *International Journal of Bifurcation and Chaos*, 17: 3691-3696.
- Funahashi, K. (1989). On the approximate realization of continuous mappings by neural networks. *Neural Networks*, 2: 183-192.
- Gradojevic, N. and Yang, J. (2000). The application of artificial neural networks to exchange rate forecasting: The role of market microstructure variables. *Bank of Canada Working Paper*, 23.
- Hornik, K. (1993). Some new results on neural network approximation. *Neural Networks*, 6: 1069-1072.
- Hornik, K., Stinchcombe, M. and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2: 359-366.
- Irie, B. and Miyake, S. (1988). Capabilities of three-layered perceptrons. *Proceedings of the First Forum on Application of Neural Networks to Power Systems*, 41-45.
- Kaashoek, J.F. and van Dijk, H.K. (2002). Neural network pruning applied to real exchange rate analysis. *Journal of Forecasting*, 21: 559-577.
- Kilby, J. and Prasad, K. (2013). Continuous wavelet analysis and classification of surface electromyography signals. *International Journal of Computer and Electrical*

Engineering, 5(1): 30-35

Krishna, B., Satyaji Rao, Y.R. and Nayak, P.C. (2011). Time series modeling of river flow using wavelet neural networks. *Journal of Water Resource and Protection*, 3: 50-59.

Lai, S.W.K.K., Yu, L. and Zhou, C. (2006). Neural-network-based meta-modeling for financial time series forecasting. *Proceedings of the 2006 Joint Conference on Information Sciences*.

Mishra, S., Saravanan, C. and Dwivedi, D.K. (2015). Study of time series data mining for the real time hydrological forecasting - a review. *International Journal of Computer Applications*, 117(23): 6-17.

Mitra, A. and Mitra, S. (2006). Modeling exchange rates using wavelet decomposed genetic neural networks. *Statistical Methodology*, 3(2): 103-124.

Pandhiani, S.M and Shabri, A.B. (2013). Time Series Forecasting using Wavelet-Least Squares Support Vector Machines and Walet Regression Models for Monthly Stream Flow Data. *Open Journal of Statistics*, 3: 193-194.

Qiu, C.R. and Rong, Z. (2008). A case based reasoning system for individual demand forecasting. *4th International Conference on Wireless Communications, Networking and Mobile Computing*, WiCOM '08: 1-6.

Ramsey, J.B. (2002). Wavelets in economics and finance: Past and future. *Working Paper No.S-MF-02-02*, C.V. Starr Center for Applied Economics, New York University.

Ripley, B.D. (1993). Statistical aspects of neural networks. In: Barndoff-Nielsen, O. E., Jensen, J. L. (Eds), *Networks and Chaos-Statistical and Probabilistic Aspects*, Chapman & Hall, London, 40-123.

Sallehuddin, S.Z.M.H.R, Shamsuddin, S.M.H. and Abraham, A. (2007). Forecasting time series data using hybrid grey relational artificial neural network and auto-regressive integrated moving average model. *Neural Network World*, 6: 573-605.

Satyaji Rao, Y.R., Krishna, B. and Venkatesh, B. (2014). Wavelet based neural networks for daily stream flow forecasting. *International Journal of Emerging Technology and Advanced Engineering*, 4(1): 307-317.

Song, D., Zhang, Y., Shan, X., Cui, J. and Wu, H. (2017). Over-learning phenomenon of wavelet neural networks in remote sensing image classification with different entropy error functions. *Entropy*, 19(101): 1-19.

Tan, C. (2009). Financial time series forecasting using improved wavelet neural network. Unpublished Master's Thesis.

Tenti, P. (1996). Forecasting foreign exchange rates using recurrent neural networks. *Applied Artificial Intelligence*, 10: 567-581.

Udomboso, C.G. (2014). On the level of precision of an alternative heterogeneous model of the statistical neural network. Unpublished *Ph.D. Thesis*, University of Ibadan, Ibadan.

Veitch, D. (2005). Wavelet neural networks and their applications in the study of dynamical systems. *MSc Dissertation in Data Analysis, Networks and Nonlinear Dynamics*, Department of Mathematics, University of York, UK.

Wen, X., Zhang, H. and Wang, F. (2009). A wavelet neural network for SAR image segmentation. *Sensors*, 9: 7509-7515; doi:10.3390/s90907509.

White, H. (1989). Learning in artificial neural networks: a statistical perspective. *Neural Computation*, 1: 425-464.

Wu, B. (1995). Model-free forecasting for nonlinear time series with application to exchange rates. *Computational Statistics & Data Analysis*, 19: 433-459.

Yonaba, H., Anctil, F. and Fortin, V. (2010). Comparing sigmoid transfer functions for neural network multistep ahead stream flow forecasting. *Journal of Hydrological Engineering*, ASCE/APRIL 2010: 275-283. Retrieved from <http://www.ascelibrary.org>.

Yousefi, I.W.S. and Reinarz, D. (2005). Wavelet-based prediction of oil prices. *Chaos, Solitons and Fractals*, 25: 265-275.

Zhang B.L. and Dong Z.Y. (2001). An adaptive neural-wavelet model for short term load forecasting. *Electric Power Systems Research*, 59: 121-129.

Zhang B.L. and Coggins R. (2001). Multi-resolution forecasting for futures trading using wavelet decompositions. *IEEE Transactions on Neural Networks*, 12(4): 765-775.

Zhang G.P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neuro computing*, 50: 159-175.

Zhang G., Patuwo B.E., and Hu M.Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14: 35-62.

Appendix

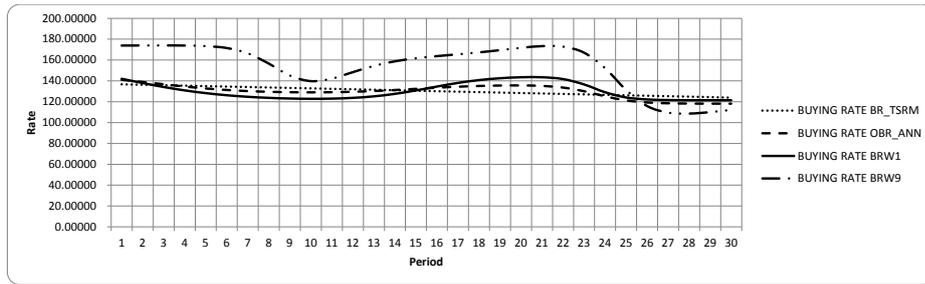


Figure 2: Forecast of buying rate across some of the models

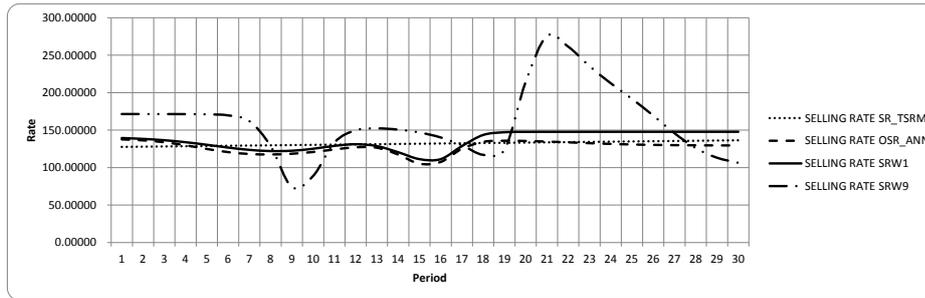


Figure 3: Forecast of selling rate across some of the models

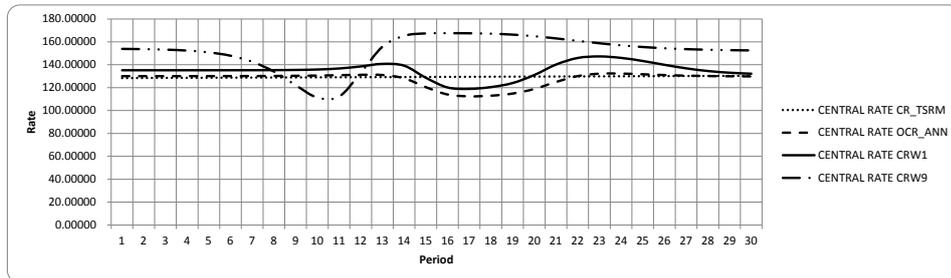


Figure 4: Forecast of central rate across some of the models

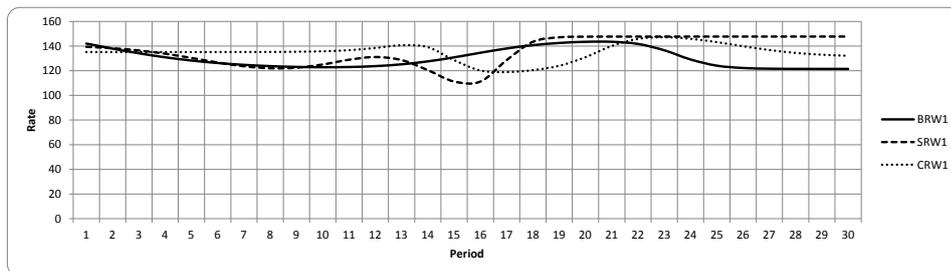


Figure 5: Forecast across Signal 1 of the WNN Models

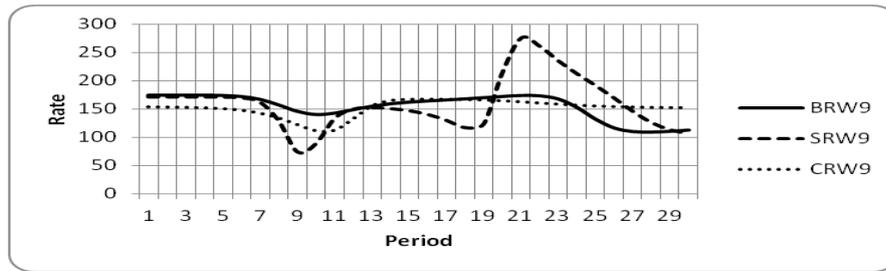


Figure 6: Forecast across Signal 9 of the WNN Models

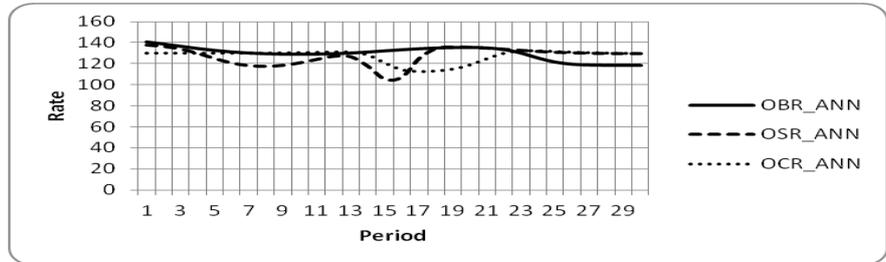


Figure 7: Forecast across the ANN Models

