Journal of Science Research (2014) Vol. 13: 133-142

On the Level of Precision of the Wavelet Neural Network in Rainfall Analysis

Christopher Godwin Udomboso*[1], Godwin Nwazu Amahia[1] and Isaac Kwame Dontwi[2]

Abstract
This research combines the efficiency of the artificial neural network and wavelet transform in modelling rainfall. The data used were decomposed into continuous wavelet signals on a scale of 48. Each of the decomposed series was subjected to correlation test with the original data. Instead of using all the series, ten series were selected on the basis of high correlation with the original data. These series included CWT 1, CWT 2, CWT 4, CWT 3, CWT 6, CWT 8, CWT 5, CWT 10, CWT 12, and CWT 7 (according to rank). The analysis showed that except in extremely rare cases, all the series performed optimally compared to the original data. The result of the study has been able to show that using the continuous wavelet transform in the ANN technique, a better performance of the network is observed.

Key words: Artificial neural network, rainfall modelling, continuous wavelet transform

Introduction

The artificial neural network has been found to model rainfall more accurately than conventional methods. This is because the artificial neural network makes use of past events of a phenomenon in predicting the future by means of learning and training. Research has shown that combining the artificial neural network with other mathematical functions and models makes the resulting model more efficient. Recently, the discrete wavelet transform has been used. However, it is known that rainfall data is rather continuous than discrete. When observed as discrete, then the interest would be on the probability of occurrence. This study therefore uses the continuous form of the wavelet transform. A wavelet network model makes use of the merits of wavelet analysis and artificial neural network, so it has excellent performance in simulation and forecast. Wavelet decomposition is a way of analyzing a signal both in time and frequency domain. The wavelet spectrum based on continuous wavelet transform (CWT), is a natural extension of the conventional Fourier spectrum analysis and short time Fourier spectrum analysis which are commonly used in climatologic time series analysis [9].

There have been considerable literatures on the modelling of rainfall data on daily basis. The majority of these models are nevertheless derived empirically. The significance of these models in meteorology is possibly attached to the fitted parameters. Most researchers make use of the gamma distribution for describing precipitation values for a variety of reasons [13, 28, 23, 26], especially because the gamma distribution is bounded on the left at zero. This is important for precipitation applications because there can never be a negative rainfall. Thus, a distribution that excludes negative values is readily applicable.

Recently, artificial neural networks (ANN) has been used in modelling rainfall. ANN constitute a useful tool to predict and forecast various hydrological variables and are used extensively in water resources research [6, 24]. The artificial neural network models are frequently employed for rainfall forecasting [22, 10]. [11] used a neural network to forecast rainfall intensity fields in space. According to [15], majority of the

Christopher Godwin Udomboso*[1], Godwin Nwazu Amahia[1] and Isaac Kwame Dontwi[2]
[1]*Department of Statistics*
University of Ibadan, Ibadan, Nigeria
[2]*Department of Mathematical Science*
Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
**Corresponding Author's E-mail:*
cg.udomboso@gmail.com,
cg.udomboso@mail.ui.edu.ng

ISSN 11179333

UNIVERSITY OF IBADAN LIBRARY

early work in this area have been mainly theoretical, concentrating on neural network performance with artificially generated rainfall-runoff data.

[18] used the combined benefits of the discrete wavelet-neural networks (DWNN) in the prediction of daily precipitation in Turkey. Wavelet transform can produce a good local representation of the signal in both the time and frequency domains, and provides considerable information about the structure of the physical process to be modeled [29, 19, 17]. [21] predicted monthly rainfall using WNN analysis. Some other areas that wavelet neural network have been used include some selected methods of thresholding for wavelet regression, SAR image segmentation, solar radiation forecasting, and forecasting IIP growth with yield spreads [2, 27, 5, 20].

Materials and Methods

Wavelets are being used in representing signals functions or images due to the fact that they allow for large compression ratios. The wavelet transform of a signal evolving in time depends on two variables – frequency (that is, scale) and time. Thus, wavelets provide a tool for time-frequency localization.

The wavelet transform is given by

$$(T^{wav}f)(a,b) = |a|^{-1/2}\int dt\, f(t)\psi\left(\frac{t-b}{a}\right) \quad (1)$$

Restricting a, b to discrete values, $a = a_0^j$, $b = kb_0a_0^j$, $j, k\epsilon\mathbb{Z}$, then

$$T_{j,k}^{wav}(f) = a_0^{-j/2}\int f(t)\psi\left(a_0^{-j}t - kb_0\right) \quad (2)$$

where Ψ is known as the mother wavelet.

In both cases, it is assumed that $\int\psi(t) = 0$

Wavelet methods have been most studied in the non parametric regression problem of estimating a function f on the basis of observations y_i at time points t_i. This is modeled as

$$y_i = f(t_i) + e_i, \qquad i = 1, 2, \ldots, n \quad (3)$$

and $e_i \sim N(0, \sigma^2)$ is the noise [8, 1]. [16] noted some authors such as [12], [3], and [14], who did not assume any distribution for the white term but noted that they are *iid* $\sim (0, \sigma^2)$.

If we consider the regression model,

$$y = \alpha I + X\beta + e \quad (4)$$

where, $y \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ is an intercept, I is an $n-$ dimensional vector of ones, $X \in \mathbb{R}^{n\times p}$ is the matrix of the signals, $\beta \in \mathbb{R}^p$ is a vector of coefficients, and $e \in \mathbb{R}^n$ is the white, or stochastic or error term ([4] 2008).

Now, defining a wavelet transformation by an orthogonal matrix [25], say $W \in \mathbb{R}^{p\times p}$, such that $WW' = I$, I being the identity matrix.

Then, we can write $f(t_i) = \alpha_i + x_i\beta$, so that in matrix form, (3) can be written as

$$y = \alpha I + XWW'\beta + e = \alpha I + Z\beta^* + e \quad (5)$$

where, $Z = XW$ is the matrix of wavelet coefficients corresponding to the series in X and $\beta^* = W'\beta$ is the new regression vector.

The data presented to the network were decomposed into wavelet forms. We recall that wavelet transforms are mathematical functions that cut up data, functions or operators into different frequency components, and then study each component with a resolution matched to its scale. The continuous wavelet transform (CWT) is used in this study, which makes use of continuous wavelets as functions.

UNIVERSITY OF IBADAN LIBRARY

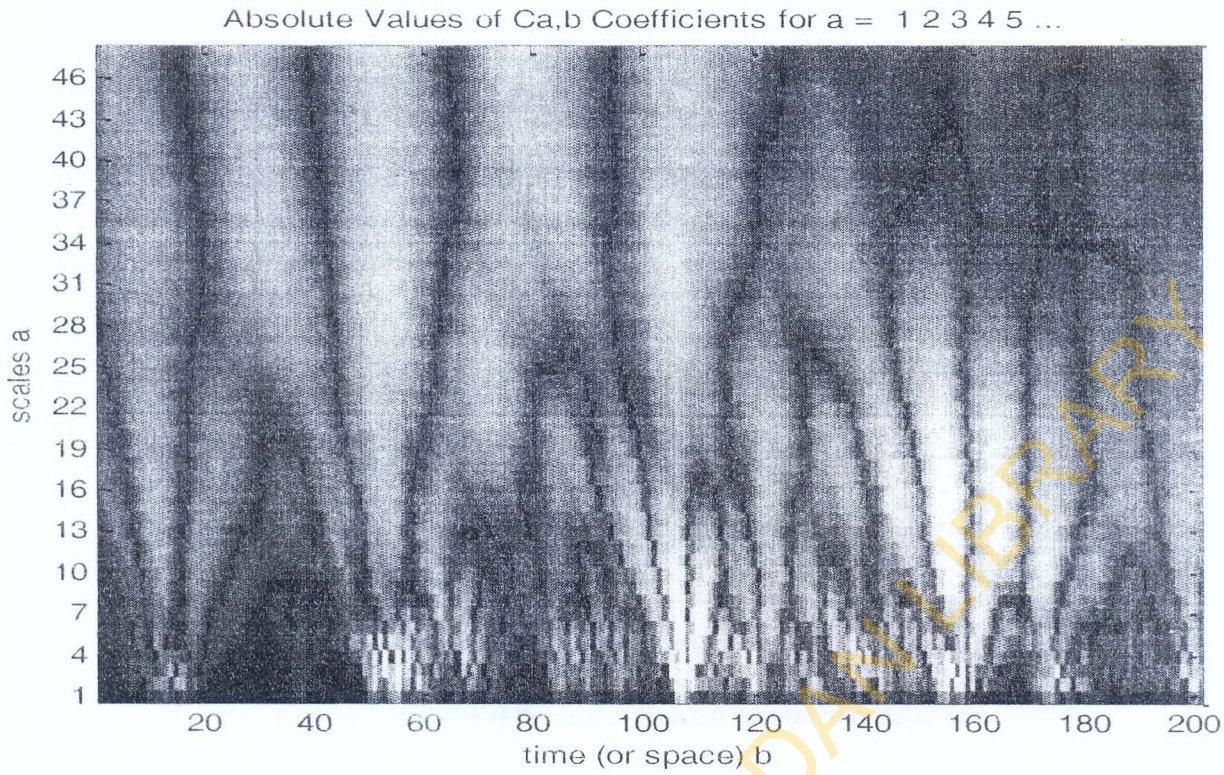

Fig. 1: Wavelet Transform of Rainfall in Ibadan for 200 days.

The decomposition here is based on a scale of 48 using the MATLAB *m-code*. This means that the original data was decomposed into 48 sub-time series. It would be unnecessary to use each sub-time series to run the MATLAB code. The first 10 highest ranked sub-time series were selected for the analyses. These included CWT 1, CWT 2, CWT 4, CWT 3, CWT 6, CWT 8, CWT 5, CWT 10, CWT 12 and CWT 7 (in that order).

A network with 100 hidden neurons was chosen for the purpose of generality, except for the convolutions where both 100 and 10 hidden neurons were used (since the network with 10 neurons had lower error variances. For the individual transfer functions, the functions that have been shown to perform best were used to investigate the continuous wavelet neural network (CWNN). These functions include:

(i) Hyperbolic Tangent transfer function
(ii) Hyperbolic Tangent Sigmoid transfer function
(iii)Symmetric Saturating Linear transfer function

Results and Discussions

The analysis begins with the construction of the time plots of the original rainfall data and the decomposed continuous wavelet transform data on a 48 scale. The original data was plotted on graph C1, while the decomposed data were plotted on the remaining graphs. That is, in Figure 2, C2 to C49 represent the entire 48 decomposed CWT data. From the plots, it could be seen that the time plots of the CWT shows an evenly distributed rainfall pattern. We see that the plots become more sparsely distributed along the scale.

ISSN 11179333

UNIVERSITY OF IBADAN LIBRARY

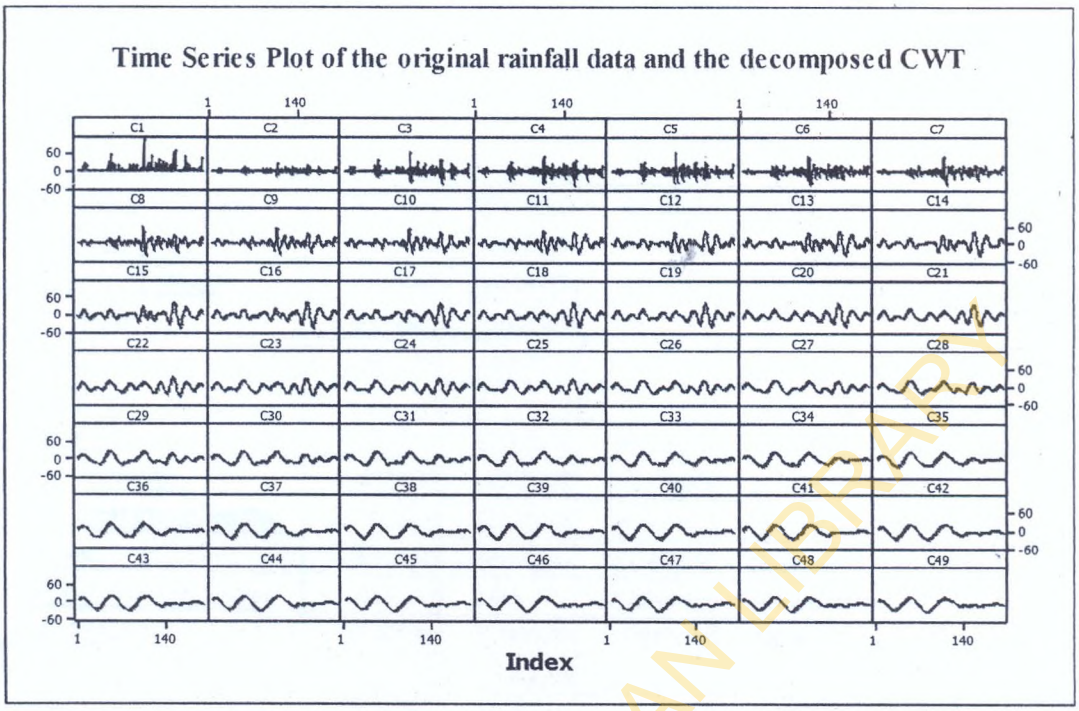

Fig. 2: Time Series Plot of the original rainfall data and the decomposed CWT.

In this study, twenty transfer functions were initially used for network training with the help of MATLAB, but fourteen emerged workable. MATLAB 2009a was used in developing the program for training the network. Five hidden layers were used. These are 2, 5, 10, 50 and 100. The number of iterations involved in this work is 1000. Daily data were used to train the network (for 200 days). Twenty transfer functions were used in the network. However, six transfer functions, namely Linear transfer function, Positive Linear transfer function, Hyperbolic Sine transfer function, Hyperbolic Cosine transfer function, exponential transfer function and gamma transfer function did not yield any output in 2, 5 and 10 hidden layers network. Nevertheless, in the higher hidden layers of 50 and 100, the *Linear transfer function* yields outputs, while others in the list of non-function transfer functions did not yield any output. The MATLAB code was trained to plot the error between the input and the output. The principle of efficiency was used in selecting the best performing transfer function.

Inspecting the general performances of the network based on the errors generated by the hidden neurons, outstanding results were obtained as follows.

The *Hyperbolic Tangent transfer function* ranked best in the overall performance with error variance 1.343777949, followed by the *Hyperbolic Tangent Sigmoid transfer function* with error variance 2.132441698 and the *Symmetric Saturating Linear transfer function*, having error variance 22.04938194. As expected, the *Radial Basis transfer function* ranked last in overall performance. This can be seen in Table 1. We can note the result from Figures 3a and 3b. The graph shows that as the hidden layers increases, error variance due to Symmetric Saturating Linear transfer function reduces drastically.

UNIVERSITY OF IBADAN LIBRARY

Table 1: Error Variances of the Transfer Functions

OVERALL RANKING	Hidden Neuron - 2	Hidden Neuron - 5	Hidden Neuron - 10	Hidden Neuron - 50	Hidden Neuron - 100	Mean	Rank
Competitive transfer function	204.5773265	223.8483493	203.0092008	139.1934749	317.1809794	217.5618662	10
Hard Limit transfer function	207.8392972	191.248995	165.7895568	96.80318859	213.7932459	175.0948567	7
Symmetric Hard Limit transfer function	129.0696401	182.6077394	115.1224912	61.87730151	85.87222364	114.9098792	5
Log-sigmoid transfer function	222.299622	331.0173098	265.5221818	268.5481521	352.1146609	287.9003853	12
Inverse transfer function	14947.78244	0.018911123	171.4533406	218.8261943	202.7426787	3108.164712	14
Linear transfer function				15.16966137	162.9709612	89.07031126	4
Radial Basis transfer function	94.48467713	97.66291743	589.5135165	13487.15479	27633.56187	8380.475554	15
Saturating Linear transfer function	135.5935607	135.732966	169.2290683	350.7622309	528.9470077	264.0529667	11
Symmetric Saturating Linear transfer function	73.30123498	19.12790579	14.39814939	2.117148666	1.302470864	22.04938194	3
Softmax transfer function	206.4595437	179.9706259	202.4876644	197.6673605	236.0353431	204.5241075	9
Hyperbolic Tangent Sigmoid transfer function	10.6083217	0.003174559	0.000954226	0.041566456	0.008191549	2.132441698	2
Triangular Basis transfer function	107.3976903	91.71668251	118.874954	2386.839526	9959.251909	2532.816152	13
Hyperbolic Tangent transfer function	6.651800849	0.002590439	0.003918007	0.00048087	0.060099581	1.343777949	1
Sine transfer function	219.0672742	240.6836446	217.0671159	1.391879726	0.417669241	135.7255167	6
Cosine transfer function	274.6038618	207.0571003	161.3597498	115.0618724	169.5493846	185.5263938	8

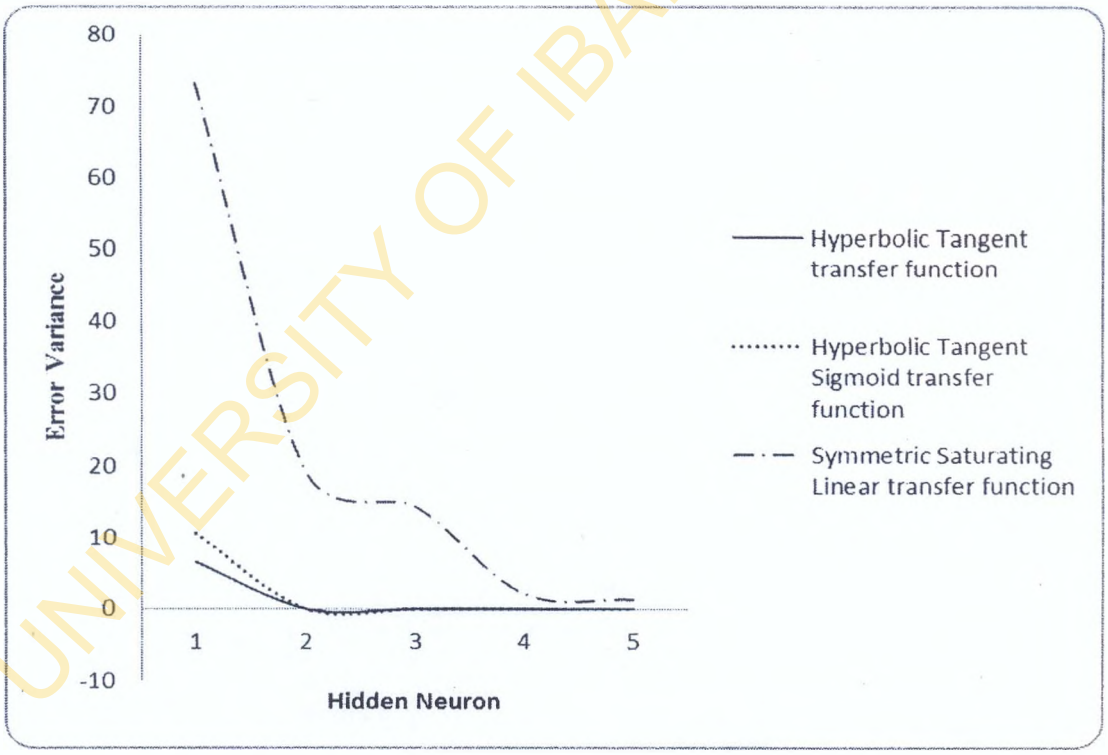

Fig. 3a: Graph of error variance of the best three transfer functions.

UNIVERSITY OF IBADAN LIBRARY

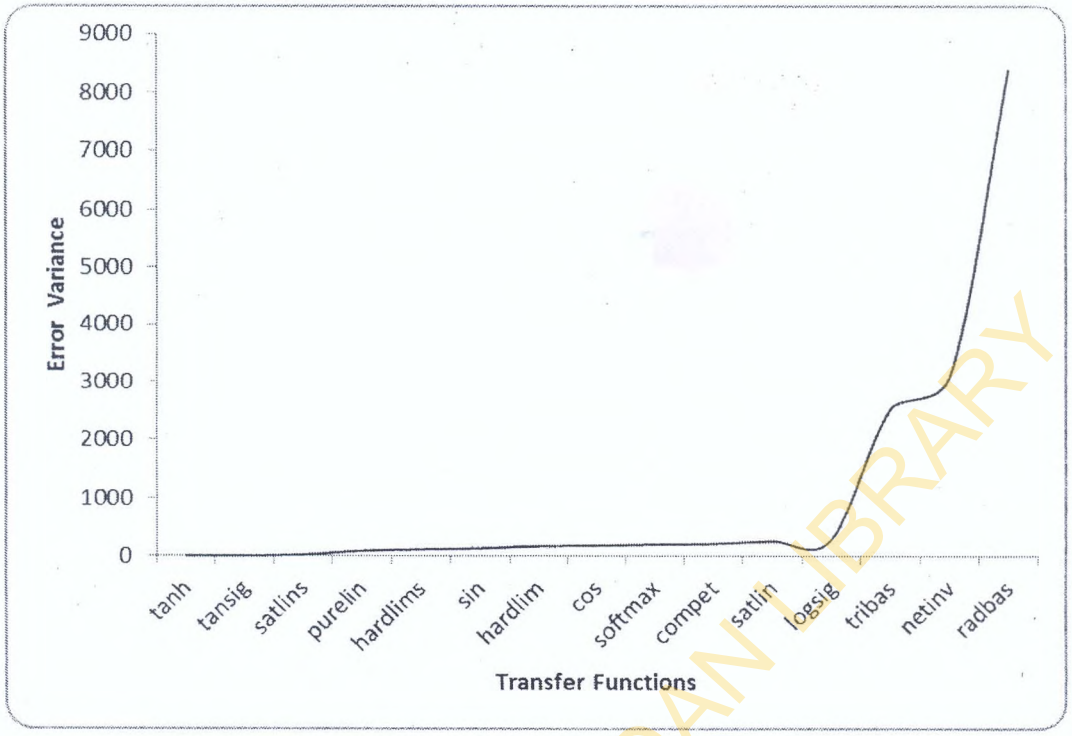

Fig. 3b: Overall Performances of Error Variance of the Transfer Functions.

In the results involving wavelets, it is found in every analysis that the continuous wavelet neural network gives better results. Tables 3a – 3c consist of the CWNN results of the best performing transfer functions earlier discovered. These are the Hyperbolic Tangent transfer function, Hyperbolic Tangent Sigmoid transfer function and the Symmetric Saturating Linear transfer function. For the Hyperbolic Tangent transfer function, CWNN resulted in a range of error variances, 0.000729592 - 0.009715933, with CWT 1 recording the least, and CWT 6 recording the highest. This is in contrast to the result of the function using the original data, having error variance 0.060099581, which is higher.

Table 3a: CWNN Result of Hyperbolic Tangent Transfer Function

Hidden Layer - 100	Mean Error (act-pred)	Mean Absolute Error	Variance
Sub-Time Series 1	0.00389005	0.013951741	0.000729592
Sub-Time Series 2	0.004260697	0.024978109	0.003239021
Sub-Time Series 4	-0.008116418	0.030443781	0.002145085
Sub-Time Series 3	-0.033570647	0.05209204	0.004102457
Sub-Time Series 6	-0.005403483	0.071986567	0.009715933
Sub-Time Series 8	0.001421891	0.023469652	0.002512283
Sub-Time Series 5	0.019393035	0.032890547	0.002339264
Sub-Time Series 10	-0.003360199	0.046734328	0.005444832
Sub-Time Series 12	-0.01788209	0.022665174	0.00073663
Sub-Time Series 7	0.023131841	0.039681095	0.002958337

UNIVERSITY OF IBADAN LIBRARY

Table 3b: CWNN Result of Hyperbolic Tangent Sigmoid Transfer Function

Hidden Layer - 100	Mean Error (act-pred)	Mean Absolute Error	Variance
Sub-Time Series 1	0.001477114	0.014383582	0.000739871
Sub-Time Series 2	0.035310448	0.131156219	0.039282011
Sub-Time Series 4	0.004956716	0.029195522	0.002329514
Sub-Time Series 3	0.028981592	0.048633333	0.004778149
Sub-Time Series 6	-0.005168657	0.028030348	0.002541222
Sub-Time Series 8	0.004595522	0.040067164	0.006018716
Sub-Time Series 5	0.014223881	0.035047761	0.002771785
Sub-Time Series 10	-0.001948756	0.016926866	0.000814949
Sub-Time Series 12	-0.007950249	0.025757214	0.001418593
Sub-Time Series 7	-0.038218905	0.043672637	0.002743456

Table 3c: CWNN Result of Symmetric Saturated Linear Transfer Function

Hidden Layer - 100	Mean Error (act-pred)	Mean Absolute Error	Variance
Sub-Time Series 1	0.012547761	0.083100995	0.099012345
Sub-Time Series 2	0.015978607	0.216885075	0.493311007
Sub-Time Series 4	-0.028518408	0.217135323	0.42369717
Sub-Time Series 3	-0.054432338	0.286941791	0.709184674
Sub-Time Series 6	0.031933333	0.247026866	0.427719785
Sub-Time Series 8	-0.035263184	0.352124876	0.951654567
Sub-Time Series 5	-0.005807463	0.285989552	0.587782167
Sub-Time Series 10	0.053563682	0.381137811	0.791104316
Sub-Time Series 12	-0.004406468	0.407753731	0.763670662
Sub-Time Series 7	-0.045742289	0.270354229	0.62697559

In the case of Hyperbolic Tangent Sigmoid transfer function, it is only CWT 2 with error variance 0.039282011 that is higher than the result obtained using the original data. While in the case of the *Symmetric Saturating Linear transfer function*, all the sub-time series data show smaller variances compared to the result from the original result.

Conducting the test of hypotheses on the results obtained, it was found out that there are significant differences between each decomposed data and the original data as shown on Tables 4 and 5. The null hypothesis in the test for difference in the means stated that there is no difference. Three alternative hypotheses were constructed. These are:

1. $\mu_i - \mu_0 < 0$,
2. $\mu_i - \mu_0 = 0$,
3. $\mu_i - \mu_0 > 0$,

where μ_i is the mean of the i^{th} decomposed data.

It is noted that only the first alternative hypothesis has the $p - value > 0.05$ (that is 1.0000). The other two has $p - value < 0.05$ (that is, 0.0000).

The variance ratio test was used to test the validity of the model based on the error generated by the network from each decomposed data. Similar to the previous test, three alternative hypotheses were also constructed, which are as follows:

UNIVERSITY OF IBADAN LIBRARY

1. $\sigma_i - \sigma_0 < 0$,
2. $\sigma_i - \sigma_0 \neq 0$,
3. $\sigma_i - \sigma_0 > 0$,

where, σ_i is the standard deviation of the i^{th} decomposed data

The test shows that at $H_1: \sigma_i - \sigma_0 \neq 0$, sub-series 1, 2, 4 and 6 are significant, while at $H_1: \sigma_i - \sigma_0 > 0$, sub series 1, 2, 4, 6 and 5 are significant. These can be seen on Table 10.

Table 4: Sample Statistics of the Original Data and the CWNN

	Mean	Standard Deviation	Standard Error of the Mean
Original Data	6.081095	14.33177	1.010885
Sub-Series 1	-0. 0105473	5. 268888	0. 3716387
Sub-Series 2	0.0268657	11.987	0.8454978
Sub-Series 4	0.1049254	12.90514	0.9102586
Sub-Series 3	0.0257711	13.10204	0.9241467
Sub-Series 6	0.1443781	11.59102	0.8175676
Sub-Series 8	0.1827363	11.85019	0.8358478
Sub-Series 5	0.2015423	12.08643	0.852511
Sub-Series 10	0.2196518	12.79639	0.9025877
Sub-Series 12	0.2327861	13.73769	0.9689816
Sub-Series 7	0.1508955	11.57113	0.8161646

Table 5: Paired Sample Statistics of the Original Data and the CWNN

	95% Confidence Interval of the Difference		t	p – value at $H_1: \mu_i - \mu_0 \neq 0$ and $H_1: \mu_i - \mu_0 > 0$
	Lower	Upper		
Original Data - Sub-Series 1	4.605974	7.577309	8.0853	0.000
Original Data - Sub-Series 2	4.788468	7.31999	9.4317	0.000
Original Data - Sub-Series 4	4.412868	7.539471	7.5381	0.000
Original Data - Sub-Series 3	4.331777	7.77887	6.9278	0.000
Original Data - Sub-Series 6	4.078572	7.794861	6.3002	0.000
Original Data - Sub-Series 8	3.893714	7.903002	5.8020	0.000
Original Data - Sub-Series 5	3.841665	7.917439	5.6892	0.000
Original Data - Sub-Series 10	3.731568	7.991318	5.4267	0.000
Original Data - Sub-Series 12	3.608266	8.088351	5.1482	0.000
Original Data -Sub-Series 7	3.809	8.051398	5.5128	0.000

UNIVERSITY OF IBADAN LIBRARY

Table 6: Variance Ratio Test of the Original Data and the CWNN

	Upper Tail (FU)	Lower Tail $(FL) = \frac{1}{F}$	p – value at $H_1: \sigma_i - \sigma_0 \neq 0$	p – value at $H_1: \sigma_i - \sigma_0 > 0$
Original Data - Sub-Series 1	171.375	0.006	0.0000	0.0000
Original Data - Sub-Series 2	7.016	0.143	0.0319	0.016
Original Data - Sub-Series 4	9.256	0.108	0.0159	0.008
Original Data - Sub-Series 3	3.314	0.302	0.1705	0.0853
Original Data - Sub-Series 6	9.142	0.109	0.0164	0.0082
Original Data - Sub-Series 8	1.830	0.546	0.4806	0.2403
Original Data - Sub-Series 5	4.807	0.208	0.0776	0.0388
Original Data - Sub-Series 10	2.653	0.377	0.2602	0.1301
Original Data - Sub-Series 12	2.834	0.353	0.2305	0.1153
Original Data -Sub-Series 7	4.232	0.236	0.1027	0.0513

Conclusion

The time plots of the actual data was scattered and all points fall on the positive side of the vertical axis, whereas the decomposed data were evenly distributed on both sides of the vertical axis. However, it was noticed that the clustering of the data became sparsely distributed as the correlation of the decomposed data with the original data became weaker. Optimal performances were noticed with Hyperbolic Tangent (tanh), Hyperbolic Tangent Sigmoid (*tansig*) and Symmetric Saturating Linear (*satlins*). Generally, as the hidden neurons increased, the error variances reduced, except in some cases where a *V*-shape is formed in the behaviour of the error variances. The data was decomposed to continuous wavelet analysis on a scale of 48 (that is, forty-eight series). Ten series were selected on the basis of high correlation with the original data. These series included CWT 1, CWT 2, CWT 4, CWT 3, CWT 6, CWT 8, CWT 5, CWT 10, CWT 12, and CWT 7 (according to rank). The analysis showed that except in extremely rare cases, all the series performed optimally compared to the original data. The result of the study has been able to show that using the continuous wavelet transform in the ANN technique, a better performance of the network is observed.

References

[1] Abramovich, F. and Silverman, B.W. 1998. Wavelet decomposition approaches to statistical inverse problems. *Biometrika* 85: 115-129.

[2] Altaher, A.M. and Ismail, M.T. 2010. A Comparison of Some Thresholding Selection Methods for Wavelet Regression. World Academy of Science, Engineering and Technology, 62.

[3] Antoniadis, A., Gregoire, G. and McKeague, I.W. 1994. Wavelet methods for curve estimation. *Journal of the American Statistical Association* 89: 1340-1352.

[4] Avarez, G. and Sanso, B. 2008. Bayesian Wavelet Regression for Spatial Estimation. *Journal of Data Science* 6: 219-229.

[5] Capizzi G., Napoli C. and Bonanno F. 2012. Innovative Second-Generation Wavelets Construction With Recurrent Neural Networks for Solar Radiation Forecasting. *IEEE Transactions on Neural Networks and Learning Systems* 23(11): 1805-1815.

[6] Cigizoglu, H.K. and Kisi, O. 2006. Methods to improve the Neural Network Performance in Suspended Sediment Estimation. *Journal of Hydrology* 317 (3/4): 221-238.

[7] Daubechies, I. 1992. Ten Lectures on Wavelets. Society for Industrial and Applied Mathematics, Philadelphia, PA.

[8] Donoho, D., Johnstone, I., Kerkyacharian, G. and Picard, D. 1995. Wavelet Shrinkage: Asymptopia? (with discussion), *Journal of the Royal Statistical Society, Ser. B*, 57: 301-369.

UNIVERSITY OF IBADAN LIBRARY

[9] Drago, A.F. and Boxall, S.R. 2002. Use of the wavelet transform on hydro-meteorological data. *Physics and Chemistry of the Earth* 27: 1387-1399.

[10] Freiwan, M. and Cigizoglu, H.K. 2005. Prediction of Total Monthly Rainfall in Jordan using Feed Forward Back propagation Method. *Fresenious Environ. Bull.* 14(2): 142-151.

[11] French, M.N., Krajewsky, W.F. and Cuykendall, R.R. 1992. Rainfall Forecasting in Space and Time using a Neural Network. *Journal of Hydrology* 137: 1-31.

[12] Hart, J.D. and Wehrly, T.E. 1986. Kernel regression estimation using repeated measurements data. *Journal of the American Statistical Association* 81: 1080-1088.

[13] Ison, N.T., Feyerherm, A.M. and Dean, B.L. 1971. Wet period precipitation and the gamma distribution. *Journal of Applied Meteorology* 10: 658-665.

[14] Kovac, A., and Silverman, B.W. 2000. Extending the scope of wavelet regression methods by coeJcient dependent thresholding. *Journal of the American Statistical Association* 95: 172-182.

[15] Minns, A.W. and Hall, M.J. 1996. Artificial Neural Networks as Rainfall-Runoff Models. *Hydrological Science Journal* 41(3): 399-417.

[16] Oyet, A.J. and Sutradhar, B. 2003. Testing variances in wavelet regression models Statistics & Probability Letters 61: 97-109.

[17] Partal, T. and Cigizoglu, H.K. 2008. Estimation and Forecasting of the Daily Suspended Sediment Data using Wavelt-Neura Networks. *Journal of Hydrology* 358(3-4): 317-331.

[18] Partal, T. and Cigizoglu, H.K. 2009. Prediction of Daily Precipitation using Wavelet-Neural Networks. *Hydrological Science Journal* 54(2): 234-246.

[19] Partal, T. and Kisi, O. 2007. Wavelet and Neuro-Fuzzy Conjunction Model for Precipitation Forecasting. *Journal of Hydrology* 342(1-2): 199-212.

[20] Pir, M.Y., Shah, F.A. and Asger, M. 2014. Using Wavelet Neural Networks to Forecast IIP Growth with Yield Spreads. *IPASJ International Journal of Computer Science (IIJCS)* 2(Issue 5): 2321-5992. ISSN.

[21] Ramana, R.V., Krishna, B., Kumar, S.R., and Pandey, N.G. 2013. Monthly Rainfall Prediction using Wavelet Neural Network Analysis. Water Resour Manage 27: 3697-3711. DOI 10.1007/s11269-013-0374-4. Published by Springer.

[22] Ramirez, M.C.V., Velho, H.F.C. and Ferreira, N.J. 2005. Artificial Neural Network Technique for Rainfall Forecasting applied to the Sao Paulo Region. *Journal of Hydrology* 301: 146-160.

[23] Thom, H.C. 1958. A note on the gamma distribution. Monthly Weather Review 86: 117-122.

[24] Toprak, F. and Cigizoglu, H.K. 2008. Predicting longitudinal Dispersion Coefficient in Natural Streams by Artificial Neural Networks. *Hydrological Processes*. 22(20): 4106-4129.

[25] Vidakovic, B. 1999. Statistical Modeling by Wavelets. Wiley Series in Probability and Statistics. New York: John Wiley & Sons.

[26] Wilks, D.S. 1995. Statistical Methods in the Atmospheric Sciences - An Introduction. Academic Press: San Diego, CA.

[27] Wen, X., Zhang, H. and Wang, F. 2009. A Wavelet Neural Network for SAR Image Segmentation. Sensors 2009, 9: 7509-7515; ISSN 1424-8220, doi:10.3390/s90907509.

[28] Woolhiser, D.A. 1992. Modeling daily precipitation – progress and problems. In Statistics in the Environmental & Earth Sciences, Walden AT, Guttorp P. Halsted Press: London; 71-89.

[29] Zhang, B.L. and Dong, Z.Y. 2001. An Adaptive Neural-Wavelet Model for Short Term Load Forecasting. Electric Power System Researches 59: 121-129.

UNIVERSITY OF IBADAN LIBRARY