International Journal of Applied Mathematics and Statistics,
Int. J. Appl. Math. Stat.; Vol. 54; Issue No. 1; Year 2016,
ISSN 0973-1377 (Print), ISSN 0973-7545 (Online)

Applied Multiresolution B-Spline Wavelet to Neural Network Model and Its Application to Predict Some Economics Data

Syamsul Bahri[1], **Widodo**[2], **and Subanar**[3]

[1]Department of Mathematics Mataram University, Indonesia
Ph.D. Student at the Departement of Mathematics Gadjah Mada University
syamsul.mat.unram@gmail.com

[2]Department of Mathematics, Gadjah Mada University, Indonesia
widodo_mathugm@yahoo.com

[3]Department of Mathematics, Gadjah Mada University, Indonesia
subanar@yahoo.com

ABSTRACT

A Method of wavelet neural network (WNN) is a neural network method with wavelet as an activate function. The architectural design of WNN models proposed in this research is a model of WNN Plus, the WNN model coupled with stage of pre-processing data. In this model, the excellence of wavelet method in terms of denoising is used at the stage of pre-processing data, and the excellence of wavelet multi-resolution is used at the hiden layer as an activate function. For the simulation, the result of this model is applied to predict the number of tourists visiting NTB Indonesia and sales volume of PT. Semen Gresik, Tbk Indonesia.

Keywords: WNN, wavelet, neural network, pre-processing data, B-spline Wavelet.

2000 Mathematics Subject Classification: 41-04, 41A15, 41A30, 62-04, 62P20, 92B20.

1 Introduction

Time series prediction is one area of application and research that attracted many scholars. The result of time series prediction is widely used in various fields such as economy and finance, demographics, geophysics and meteorology, medicine, and various industrial processes. More specific in economic and finance, time series prediction is used in various areas for instance price prediction, stock prediction, warehouse and sales volume prediction, price and sales volume of stocks prediction, prediction of bank rate interest, prediction value of currency exchange, and so forth. Nowadays, several methods have been used to analyze time series data such as ARIMA, ARCH and GARCH. However, the level of application of all of these models has barrier. The data in these methods requires stationary and linear. On the contrary, the economic and financial data are generally nonlinear and non stationary (heteroschedastis). Besides above methods, in the decade of the 1990s also appeared new methods namely neural network method, wavelet method and wavelet transformation. Although

they were considered current methods, the excellences and capabilities of these methods attracted the attention of the researchers. Therefore, in short of time, many papers were created by using both of these methods as a tool for analyzing time series data.

Neural Network (NN) is a soft computing method that is motivated by the human brain system. NN is preferred because of its ability to handle complex and nonlinear data. This model was originally developed by Warren McCullah and Walter Pitts in 1943 (Arbib, 2003). Several studies have been conducted by using NN as tools for data analysis, for instance prediction model of rainfall in Bangkok Thailand (Hung, Babel, Weesakul and Tripathi, 2008), prediction model of load electrical power station in Nigeria (Adepoju, Ogunjuyigbe and Alawode, 2007), prediction model and data analysis of meteorology and oceanography (Hsieh and Tang, 1998), prediction of knowledge employees phychological contract (Yang, 2013), application of BP Neural network in autonomous english learning (Tai, 2013), etc.

Wavelet and NN are two methods based on soft computing, and partially have their own excellences. NN has its adaptive ability and learning algorithm, the ability of generalizing and solving nonlinear complex and cumbersome problems, while wavelet has its role as a basis function and the excellences in process of denoising, compression, and multi-resolution. The combination of these two methods is called neural network and wavelet analysis (wavelet transformation). These methods are relatively current ideas in science and technology and generally named as WNN. Based on these excellences, many research problems can be solved by using this WNN method such as modeling system, analysis system of reliability, recognition of patterns, and some applications in engineering.

In NN method, the process of determining the weight of interconnection among neurons by process of learning is one of the processes that determines the speed of convergence model. In general, the more input or more ”ugly” input will make the convergence model longer. Hence, one way to overcome this issue is through the stages of pre-processing data. Pre-processing data is an important step in data analysis which aims to improve the quality of input. Methods of pre-processing data can be divided into four categories, namely cleaning, integration, transformation, and reduction of data (Nisbet, Elder and Miner, 2009). One of the latter techniques that is widely used in data processing is wavelet. The excellence of wavelet in the process of denoising, compression and multi-resolution becomes reasons that this method is widely used in the analysis of time series data (Li, Li, Zhu and Ogihara, 2003). Numerically, (Bahri, Widodo and Subanar, 2013) have shown that the applied of Daubechies wavelet transformation at the stage of pre-processing data to the standard of neural network models gives significant results compared with no pre-processing stage of the data.

WNN models proposed in this paper is a sub model of FWNN model as a research of the author’s dissertation. WNN models proposed is a model that accommodates the excellence of wavelet as a method of denoising at the stage of pre-processing and excellence of wavelet in terms of multi-resolution as an activation function in the learning of NN. Characteristic of Daubechies wavelet is used to optimize the quality of the input at the stage of pre-processing

data and the excellence characteristics of B-spline wavelets (Unser, 1997) is used as the activate function on the model of WNN.

2 Wavelet and Data Processing

2.1 Wavelet

Wavelet is a family of function that is constructed from the process of translation and dilatation to a function which is defined as follows ((Debnath, 2002) and (Sarkar, Magdalena and Michael, 2002)):

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}), \ \ a, b \in \mathbb{R} \text{ and } a > 0 \tag{2.1}$$

a is a scaling parameter which determines the degree of compression or scale, and b is a translation parameter which determines the time location of wavelet. Notation ψ represents a wavelet function which is called mother wavelet. Function ψ in (2.1) satisfy the admissibility conditions as follows (see (Debnath, 2002) and (Misiti, Misiti, Oppenheim and Poggi, 2007)):

$$(i) \int_{-\infty}^{\infty} \psi(t)\, dt = 0, \ \ (ii) \int_{-\infty}^{\infty} \psi^2(t)\, dt = 1, \ \text{ and } \ (iii)\, C_\psi = \int_{-\infty}^{\infty} \frac{|\psi - w|^2}{w}\, dw < \infty \tag{2.2}$$

For discrete wavelet transforms, the parameters a and b of the wavelet ψ in (2.1) are valued discrete variables by choosing $a = a_0^m$, $a_0 > 1$ and m integers. Because the width of $\psi(a_0^{-m}x)$ proportional to the value of a_0^m, it can be chosen discretization value b with $b = nb_0a_0^m$ for a $b_0 > 0$ and n,m integers, so the discrete wavelet transform is given by

$$\psi_{m,n}(t) = a_0^{-m/2}\psi(a_0^{-m}t - nb_0) \tag{2.3}$$

Similar to fourier series, for each $f(t) \in L^2(\mathbb{R})$, the series wavelet representation is defined by

$$f(t) = \sum_n \sum_m c_{m,n}\psi_{m,n}(t) \tag{2.4}$$

where $\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n)$ and $c_{m,n} =< f, \psi_{m,n} >$

Daubechies wavelet is one type of wavelet developed by Ingrid Daubechies around 1990s. Daubechies wavelet family is called Daubechies Wavelet orders N (DbN) for some N natural numbers and $N \geq 2$ (Figure 1). It is one type of wavelet with the smooth curve design, by this physical characteristics of smoothness (smooth) Daubechies wavelet becomes very good in analyzing the signal that requires smoothness, and as an effective way to reconstruct the image in high resolution with various forms of wavelet (DbN family) it can handle various forms of signals (Misiti et al., 2007).

Daubechies polynomials of order N is defined as (Ruch and Van Fleet, 2009)

$$\mathbf{P}_N(t) = \sum_{k=0}^{N} \binom{2N}{k} t^k (1-t)^{N-k} \tag{2.5}$$

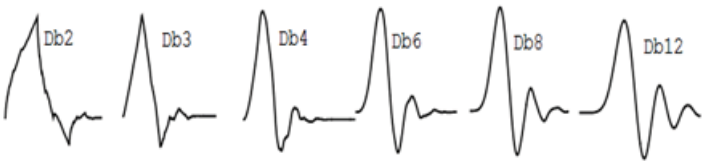

Figure 1: Graph of Daubechies Wavelet for Different Order (2, 3, 4, 6, 8, and 12)

For $m \geq 1$, B-spline order m defined as (Goswami and Andrew, 2011)

$$N_m(x) = \underbrace{(N_1 \cdot N_1 \cdot N_1 \ldots N_1)}_{n}(x) = (N_{m-1} \cdot N_1)(x) = \int_0^1 N_{m-1}(x-t)\,dt \tag{2.6}$$

with $N_1(x)$ is a characteristic at $[0,1]$ or known as Haar scale function,

$$N_1(t) = \begin{cases} 1, & 0 \leqslant t \leqslant 1 \\ 0, & otherwise \end{cases} \tag{2.7}$$

Based on the convolution, the equation (2.7) can be written as

$$N_m(x) = \frac{1}{(m-1)!}\sum_{k=0}^{m}(-1)^k\binom{m}{k}(x-k)_+^{m-1} \tag{2.8}$$

with $(x)_+^m = max\{0, x\}^m$

Furthermore, for $m \geq 2$ the first derivative of the $N_m(x)$ is called B-spline wavelets order $m-1$ and is given by the following equation ((Unser, 1997) and (Goswami and Andrew, 2011))

$$N_m^{'}(x) = \psi_{m-1}(x) = \frac{1}{(m-1)!}\sum_{k=0}^{m}(-1)^k\binom{m}{k}(x-k)_+^{m-2} \tag{2.9}$$

Numerically, (Unser, 1997) has resulted in the formulation of wavelet B-spline order m as follows,

$$\psi^m(x) = \frac{4b^{m+1}}{\sqrt{2\pi(m+1)\sigma_w^2}}\cos\left(2\pi f_0(2x-1)\right)\exp\left(\frac{-(2x-1)^2}{2\sigma_w^2(m+1)}\right) \tag{2.10}$$

with b = 0.697066, f_0 = 0.409177 and σ_w^2=0.561145.

B-spline wavelet has several excellences which can be a primary consideration to choose it as an analysis tool. Wavelet has an explicit formulas in the form of time and frequency domain, best approximation properties, symmetry and has a compact support, and simple manipulation process (Unser, 1997).

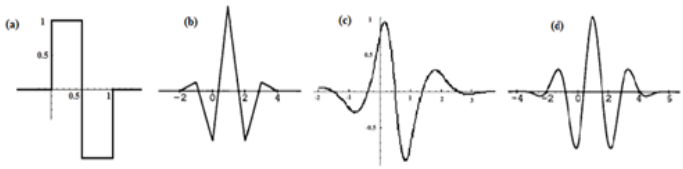

Figure 2: Graph of B-spline wavelets 1-st to 4-th order

2.2 Multi-resolution Analysis

Multi-resolution analysis (MRA) is a procedure to find a subspace nested $V_j \subseteq L^2(\mathbb{R}), j \in \mathbb{Z}$, that will be used to approximate function in $L^2(\mathbb{R})$.

Function scale $\phi(t) \in L^2(\mathbb{R})$, span nested subspace $V_j \subseteq L^2(\mathbb{R})$ so (Goswami and Andrew, 2011)

$$\{0\} \leftarrow \cdots V_{-2} \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq V_2 \cdots \rightarrow L^2 \tag{2.11}$$

and satisfy of the scaling equation

$$\phi(t) = \sum_k p_k \phi(at - k) \tag{2.12}$$

for $a > 0$ and $a \neq 1$, and coefficient sequence of $\{p_k\} \in \ell^2(\mathbb{R})$. Subspace V_n is spanned by $\{\phi(2^n t - k)\}$.

Define W_n subspace in $L^2(\mathbb{R})$ which is an orthogonal subspace to the V_n thus $V_{n+1} = V_n \oplus W_n$ (see Figure 3). Subspace W_n is called wavelet subspace which is constructed by wavelet function $\psi(t)$.

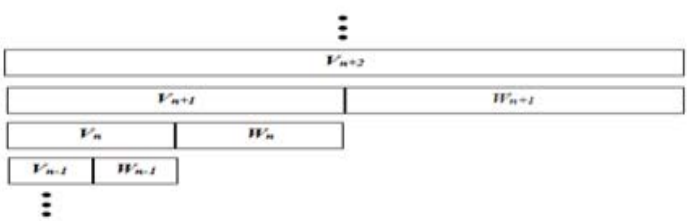

Figure 3: Decomposition Illustration of space V_{n+1} becomes subspace V_n and W_n Source : Goswami and Andrew (2011)

Because $V_{n+1} = V_n \oplus W_n$ then for each $f_j(t) \in L^2(\mathbb{R})$ and $\mathrm{g}_j(t) \in W_j$ can be written as

$$f_j(t) = \sum_{k=-\infty}^{\infty} c_{j,k}\, \phi(2^j t - k) = \sum_{k=-\infty}^{\infty} c_{j,k}\, \phi_{j,k} \tag{2.13}$$

$$\mathrm{g}_j(t) = \sum_{k-\infty}^{\infty} d_{j,k}\, \psi(2^j t - k) = \sum_{k-\infty}^{\infty} d_{j,k}\, \psi_{j,k} \tag{2.14}$$

for a sequence $\{c_{j,k}\}$ and $\{d_{j,k}\}$ in the ℓ^2. Hence, a function $f(t) \in L^2(\mathbb{R})$ can be approximated by a function $(f)_M(t) \in V_M$ by the equation

$$f(t) \simeq f_M(t) = \sum_{m=1}^{N} \mathrm{g}_{M-m}(t) + f_{M-m}(t) \tag{2.15}$$

2.3 Wavelet and Pre-Processing Data

Pre-processing data is an important stage in data analysis. The aims of this stage is to improve the quality of input. Methods of pre-processing can be divided into four categories, namely cleaning, integration, transformation, and reduction of data (Nisbet et al., 2009). In general, the data obtained from an observation contains noise, there is a missing value, and it might be not consistent. In addition, the data obtained is not common to have large amounts and high dimension ((Li et al., 2003) and (Nisbet et al., 2009)). Therefore, we need a process to reduce or eliminate the noise, interpolation of missing data, reduce the number and dimensions of data, and transforming the data. This process is a part of pre-processing data. One of the latter techniques that is widely used in processing the data is wavelet. The Excellence of wavelet in process of de-noising, compression and multi-resolution is a reason why this method is widely used in the analysis of time series data (Li et al., 2003), especially at the stage of pre-processing data.

3 Proposed Architecture Model of Wavelet Neural Networks

Architecture model of WNN developed in this paper can be seen at Figure 4. Design of proposed WNN feed-forward architecture consists of six layers. The first layer is the input of time series data and the second layer is the result of the discrete wavelet transform Daubechies (Db8 3-rd level) of the data that has been normalized. In the third layer, each data on second layer is weighted by a matrix W with size $n \times c$ by using the formula

$$Xw_k = \sum_{j=1}^{n} w_{j,k}\, x_j,\ \ k = 1, 2, \cdots, c \tag{3.1}$$

In the fourth layer, each data on the third layer is activated by using B-spline wavelet transformation of the 1-st to c-th order, and c variation of translation and scaling parameter, by notation

$\psi_j^i(t)$. $\psi_j^i(t)$ indicates a variation j-th and order i-th of the B-spline wavelet, $i = 1, 2, \cdots, n$ and $j = 1, 2, \cdots, c$. By Unser formula, $\psi_j^i(t)$ is defined as

$$\psi_j^i(z) = \frac{4b^{i+1}}{\sqrt{2\pi(i+1)\sigma_w^2}} \cos\left(2\pi f_0(2z_j - 1)\right) \exp\left(\frac{-(2z_j - 1)^2}{2\sigma_w^2(i+1)}\right) \tag{3.2}$$

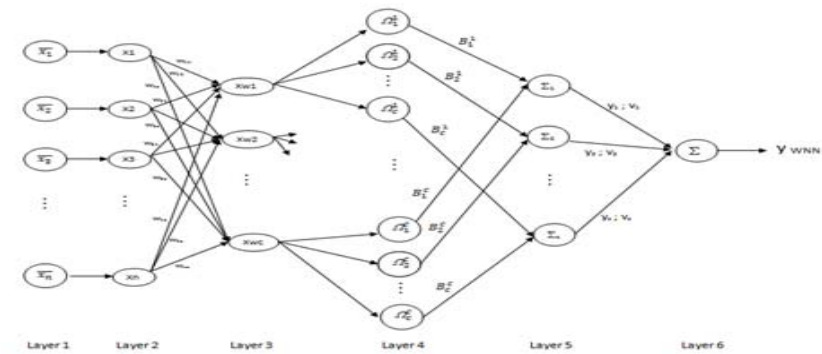

Figure 4: Design of proposed WNN architecture

The fifth layer, the aggregation of each order of B-spline wavelets uses equation

$$y_j = \alpha_j \sum_{k=1}^{c} \psi_k^j(Xw_{jk}) \tag{3.3}$$

For a constant $\alpha_j \in \mathbb{R}$, $Xw_jk = \frac{Xw_j - b_k}{a_k}$ as variation k-th $(k = 1, 2, \cdots, c)$ of the Xw_{jk}, $j = 1, 2, \cdots, c$ and the parameters a_k and b_k are indicate variations k-th of the scaling and translation parameter of the wavelet B-spline. The number of variation of the translation and scaling parameters is equal to the amount of class data classification which is determined using FCM clustering algorithm.

Then, the sixth layer is a weighted aggregation of the results of the fifth layer as follows

$$y_{WNN} = \frac{\sum_{j=1}^{c} v_j\, y_j}{\sum_{j=1}^{c} v_j} + \beta \tag{3.4}$$

for a constant $\beta \in \mathbb{R}$.

4 Learning Parameters of WNN

The model of WNN proposed is trained by using the type of supervised training. Training WNN is done to minimize the cost function

$$E = \frac{1}{2} \sum_{i=1}^{P} (y_i - o_i^d)^2 \tag{4.1}$$

with P denotes the number of WNN model output (in this research is P = 1), y_i indicates output of the WNN method and o_i^d indicates output targets. By using the gradient descent algorithm with momentum, w_{ij} and v_j weight parameters, scaling parameter a, and translation parameter b will be updated using equation (4.2) - (4.5).

$$w_{jk}(t+1) = w_{ij}(t) + \eta_w \frac{\partial E}{\partial w_{jk}} \tag{4.2}$$

$$v_j(t+1) = v_j(t) + \eta_v \frac{\partial E}{\partial v_j} \tag{4.3}$$

$$a_j(t+1) = {}_j(t) + \eta_a \frac{\partial E}{\partial_j} \tag{4.4}$$

$$b_j(t+1) = b_j(t) + \eta_b \frac{\partial E}{\partial a_j} \tag{4.5}$$

with η_w, η_v, η_a, and η_b indicate learning rate parameter for the weight parameters w_{jk} and v_j, scaling, and translation parameters of the B-spline wavelets.

The value of partial derivatives in equation (4.2) - (4.5) is given by the following equation

$$\frac{\partial E}{\partial w_{jk}} = \frac{\partial E}{\partial y_{WNN}} \frac{\partial y_{WNN}}{\partial y_j} \frac{\partial y_j}{\partial \psi_j^k} \frac{\partial \psi_j^k}{\partial X w_j^k} \frac{\partial X w_j^k}{\partial X w_j} \frac{\partial X w_j}{\partial w_{jk}} \tag{4.6}$$

$$\frac{\partial E}{\partial v_j} = \frac{\partial E}{\partial y_{WNN}} \frac{\partial y_{WNN}}{\partial v_j} \tag{4.7}$$

$$\frac{\partial E}{\partial a_j} = \frac{\partial E}{\partial y_{WNN}} \frac{\partial y_{WNN}}{\partial y_j} \frac{\partial y_j}{\partial \psi_j^k} \frac{\partial \psi_j^k}{\partial X w_j^k} \frac{\partial X w_j^k}{\partial a_j} \tag{4.8}$$

$$\frac{\partial E}{\partial b_j} = \frac{\partial E}{\partial y_{WNN}} \frac{\partial y_{WNN}}{\partial y_j} \frac{\partial y_j}{\partial \psi_j^k} \frac{\partial \psi_j^k}{\partial X w_j^k} \frac{\partial X w_j^k}{\partial b_j} \tag{4.9}$$

Selection value of learning rate parameters to ensure convergence of WNN model which is determined by (Banakar and Azeem, 2008) is the value of learning rate parameter in equation (4.2) - (4.5) which ensures convergence WNN model

$$0 \leq \eta_\rho \leq \frac{2}{max_t(\frac{\partial y(t)}{\partial \rho(t)})^2}, \quad \rho = w, v, a, \text{ and } b \tag{4.10}$$

5 Simulation Model

Model of WNN which is developed in this research is applied to predict some economic data, that is the volume of tourist who visited Nusa Tenggara Barat (NTB), Indonesia and sales volume of PT. Semen Gresik, Tbk.

5.1 Volume of Tourists Visiting NTB

The data used is the number of tourists who visited NTB period of January 2007 to December 2013 (Figure 5). Data $x(t)$ is predicted by using sequences data $x(t-18)$, $x(t-16)$, $x(t-14)$, $x(t-12)$, $x(t-9)$, $x(t-6)$, $x(t-4)$, $x(t-3)$, $x(t-2)$, and $x(t-1)$. From this structure, there are 66 pairs of data that is organized into 50 data for training and 16 data for testing.

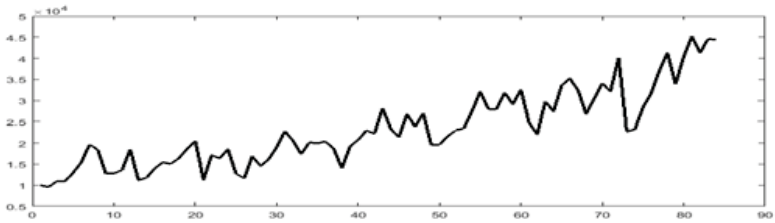

Figure 5: Number of tourists who visited NTB, Period of January 2007 to December 2013
Source: www.bappeda-ntb.go.id

By using FCM it is obtained that the number classes of data classification is 4. Thus, associated with WNN architecture in Figure 4, there are 10 data inputs and 4 class data classification. This condition has an implications on the number of parameters involved in the model, that is 58 parameters consisting of six fixed parameters (number of data classification class $\sim c$, momentum parameter $\sim mc$, and four parameter learning rate $\sim \eta_w, \eta_v, \eta_a, \eta_b$ and 52 parameters which are updated in the model (40 weight parameters for matrix $\mathbf{W}$, and each of the four parameters for the weight matrix $\mathbf{V}$, scaling parameter a and translation parameter b.

Comparison of actual data and output generated by the model of WNN results 300 epoch as shown in Figure 6 (a) and the performance of model (MSE) is 0.024496 (Figure 6(b)).

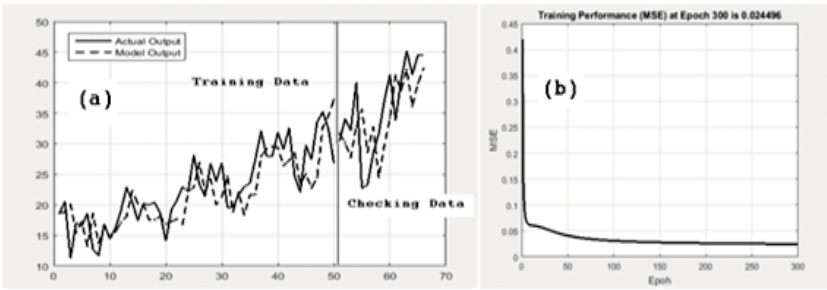

Figure 6: (a) Comparison between the actual data with the output model generated by the WNN model proposed, (b) Performance model at the 300 epoch is 0.024496.

5.2 Sales Volume of PT. Semen Gresik, Tbk.

The data used in this study is monthly sales volume data in PT. Semen Gresik Tbk. during the period of January 2007 to March 2015 (Figure 7). Data $x(t)$ is predicted using a sequences data $x(t-13)$, $x(t-12)$, $x(t-9)$, $x(t-6)$, $x(t-4)$, $x(t-3)$, $x(t-2)$, and $x(t-1)$. From this structure, there are 98 pairs of data that is organized into 75 data for training and 23 data for testing.

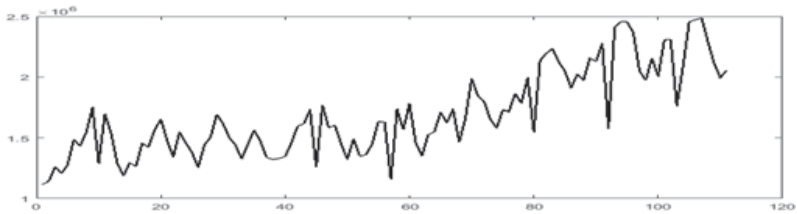

Figure 7: Sales volume of cement at PT. Semen Gresik, period of Jan 2007 to Mar 2015
Source: www.semengresik.com/ina.

By using FCM, it is obtained that the number classes of data classification is 5. Thus, it can be associated with WNN architecture in Figure 4, there are 8 data input and 5 class data classification. This condition has an implication on the number of parameters involved in the model. They are 61 parameters consisting of six fixed parameters (number of data classification class $\sim c$, momentum parameter $\sim mc$, and four parameters learning rate $\sim \eta_w, \eta_v, \eta_a, \eta_b$ and 55 parameters which are updated in the model (40 parameter weight matrix $\mathbf{W}$, and each of the five parameters for the weight matrix $\mathbf{V}$, scaling parameter a and translation parameter b.

Comparison of actual data cement sales of PT. Semen Gresik, Tbk. and the output values generated by the model WNN results 300 epoch as shown in Figure 8 (a) and the performance of the models (MSE) is 0.027282 (Figure 8 (b)).

Figure 6 (a) and 8 (a) show that the result prediction of the models is not so accurate but is quite significant, it is measured by the relatively small number of MSE. This condition is understandable, due to cement sales volume or the number of tourists who visit tourism area is not always determined by the amount/volume of previous period. The number of tourist resorts visiting tourism area is affected by many factors such as the existance of tourism, facilities and infrastructure that support the tourism area and the level of security both in tourism area or its surroundings. Likewise in cement sales volume, it is influenced by various factors such as the level of prosperity of the community (purchasing power), the rate of economic growth of the region or country, and even the condition of the season.

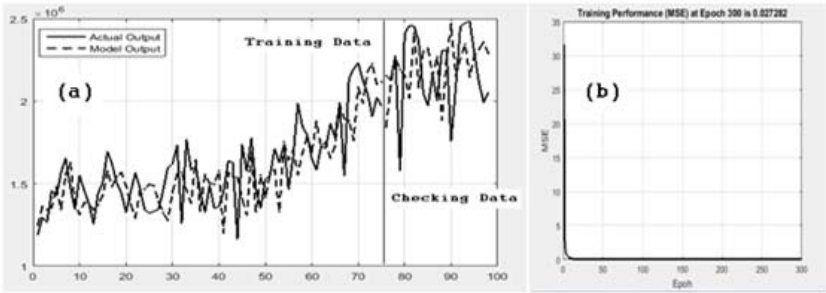

Figure 8: (a) Comparison between actual data volume of sales of PT. Semen Gresik, Tbk and the output generated by the model WNN, (b) Performance model at the 300 epoch is 0.027282.

6 Conclusion

This research has produced an architectural model of WNN with neural network models plus stages of pre-processing data. The excellence of wavelet method in terms of denoising at the stage of pre-processing of data is used to improve the quality of input and the excellence of wavelet multi-resolution is used as an activation function of neural network architectures. The presented WNN model has the advantages of approximation accuracy and good generalization performance. This prediction models provide significant results to predict the number of tourists who visit the region especially NTB and predict sales volume of cement (monthly) of PT. Semen Gresik, Tbk. as an example simulation in this research.

Acknowledgment

I would like to thank all parties, in particular I wish to thank the Indonesian government has provided scholarships BPPS. Our gratitude also goes to the anonymous reviewer who has given to the improvement of this paper. We also thank the Gadjah Mada University who has provided the opportunity to learn Ph.D. in mathematics.

References

Adepoju, G., Ogunjuyigbe, S. and Alawode, K. 2007. Application of neural network to load forecasting in nigerian electrical power system, *the Pacific Journal of Science and Technology* **Vol. 8**(1).

Arbib, M. A. 2003. *The Handbook of Brain Theory and Neural Networks*, A Bradford Book THE MIT PRESS, London, England.

Bahri, S., Widodo and Subanar 2013. Optimalization of applied neural network methods with pre-processing data using daubechies wavelet transformations, *Prociding of National Seminar on Statistics: Statistics on Disaster Management,* pp. 425–439.

Banakar, A. and Azeem, M. 2008. Artificial wavelet neural network and its application in neuro fuzzy models, *Applied Soft Computing* **Vol. 8**: 1463–1485.

Debnath, L. 2002. *Wavelet Transforms and Their Applications (With 16 Figures)*, Biekhauser, Boston.

Goswami, J. and Andrew, K. 2011. *Fundamental of Wavelets : Theory, Algorithms, and Applications Second Edition*, John Wiley and Sons, Inc., Publication, USA.

Hsieh, W. and Tang, B. 1998. Applying neural network models to prediction and data analysis in meteorology and oceanography, *Bulletin of the American Meteorological Society* .

Hung, N. Q., Babel, M. S., Weesakul, S. and Tripathi, N. K. 2008. An artificial neural network model for rainfall forecasting in bangkok,thailand, *Hydrology and Earth System Sciences Discussions are under open-access review for the journal Hydrology and Earth System Sciences* .

Li, T., Li, Q., Zhu, S. and Ogihara, M. 2003. A survey on wavelet applications in data mining, *SIGKDD Explorations* **Vol. 4**(2): 49.

Misiti, M., Misiti, Y., Oppenheim, G. and Poggi, J. 2007. *Wavelets and Their Applications*, ISTE Ltd, USA.

Nisbet, R., Elder, J. and Miner, G. 2009. *Handbook of Statistical Analysis and Data Mining Applications*, Elsevier Inc, Amsterdam.

Ruch, D. K. and Van Fleet, P. J. 2009. *Wavelet Theory : An Elementary Approach with Applications*, John Wiley and Sons, Inc., New Jersey.

Sarkar, T., Magdalena, S. and Michael, C. 2002. *Wavelet Applications in Engineering Electromagnetics*, Artech House electromagnetic analysis series, London.

Tai, Y. 2013. Application of bp neural network in autonomous english learning, *International Journal of Applied Mathematics and Statistics* **Vol. 51**(22).

Unser, M. 1997. Ten good reasons for using spline wavelets, *Proc. SPIE, Wavelets Applications in Signal and Image Processing* **Vol. 3169**: 422–431.

Yang, L. 2013. Application of neural network in prediction of knowledge employees' phycological contract, *International Journal of Applied Mathematics and Statistics* **Vol. 43**(13).