Reducing the bullwhip effect from signal demand of hybrid artificial neural network models of supply chain in Indonesia

Edy Fradinata [1, 2, *], Sakesun Suthummanon [2], Wannarat Suntiamorntut [3]

[1]Industrial Engineering, Universitas Serambi Mekkah, Banda Aceh, Indonesia
[2]Industrial Engineering, Prince of Songkla University, Hatyai, Thailand
[3]Computer Engineering, Prince of Songkla University, Hatyai, Thailand

ARTICLE INFO

ABSTRACT

The bullwhip effect becomes a problem for the factory to manage the inventory policy in the warehouse. This study proposes to reduce the bullwhip effect through signal demand forecast from hybrid artificial neural network (ANN) models. The original ANN is combined with analytical hierarchy process (AHP), Monte Carlo simulation (MC), and geometric random distribution at the parts of the input weight and input bias from the ANN. The variation of forecast signal demands from the hybrid models are used to reduce the variance from the signal customer demands. The results from this study, $AHP_{iw}ANN_b$ has the smallest mean square error (MSE) from signal demands, it implys that the variance signal demands should reduce the bullwhip effect (BWE) in the supply chain. It can be concluded that the small variance signal demand should reduce the bullwhip effect in the supply chain.

1. Introduction

The accuracy of the forecast signal demand is quite an important to predict the external signal demand on the retailer (Dejonckheere et al., 2003). ANN is one method that could be combined with other methods to obtain better performance. The combination of the network could be implemented at the four spaces; input weight, input layers, bias at the hidden layer, and bias output layer. This study explores the hybrid models from original ANN to minimize the forecast error and to reduce the bullwhip effect in the supply chain member. The forecast error is reduced by the hybrid model from AHP, MC and geometric random distribution combine with the original ANN. The algorithm hybrid methods work with the six input variables and one variable demand as the target; they are separated into 80% as training process and the rest of 20% as testing and validating. The weight and bias variables from AHP, ANN and the geometric random distribution substitute at the input weight, input bias of the backpropagation ANN system. The cooperation of the weight, bias and the transfer function of *trainlm* iterate the training process

before it is tested and validating by the iteration process to obtain the small result of MSE. However, a quite few studies from researchers have reported the combination methods to obtain the best performance compare to other methods. The contribution of this paper is to estimate the signal cement demand from the determinant of cement demand factors using the hybrid ANN from AHP, MC, and the geometric random distribution. The geometric random normal distribution is selected from other random distributions to substitute at the bias in the hidden layer of ANN network. The MC and AHP are a replacement at the input weight into the ten neurons in the hidden layer of the backpropagation ANN. The additional contribution of this paper is reducing the bullwhip effect from the several algorithms of hybrid forecasting signal demand models from an original ANN. Both of these contributions become the advantage for the factory as the alternative methods to forecast and reduce the bullwhip effect from signal demand in the supply chain.

Most of the business seems to agree that the bullwhip effect is the phenomenon of the information varies from customer signal demand, which has placed an order to the retailer, and it starts to raise the variation in the downstream to the upstream to fulfil the inventory in the distribution channels (Lee et al., 2004). Lack of sharing information forced all members in a supply chain to anticipate the space by doing the stockpile. The

preparation of stockpile needs to anticipate the degree of uncertainty conditions. It leads to obtaining poor customer services, miss-production schedules, incorrect capacity plans, inefficient shipping, and be a high cost for the manufacturers (Nikdel et al., 2014). Many companies have observed the lack of information on the bullwhip effect. The suppliers and retailers are aware of the bullwhip effect phenomenon that can influence the company's inventory for each supply chain member. Safety stock utilization is taken to forecast the signal demand fluctuation. The supplement of signal demand forecast at the retailer should overreact to manufacture. The time series of signal demand dataset and the stationarity from the normal random distribution would create the bullwhip effect in the supply chain member (Lenny Koh et al., 2007).

The bullwhip effect has the classification of amplification ratio for the building industry group, retail industries has the ratio of 0.92 at the retailer level (Cachon et al., 2007). The statistical methods have implemented to minimize the forecasting error, but the results do not meet the expectations. Some methods such as data mining, heuristic processes and others work to obtain smaller forecasting errors to minimize the bullwhip effect. Lately, ANNs have become extremely popular for prediction and forecasting in many areas. The comparison of the artificial neural network (ANN) method to some traditional forecasting methods are applied for prediction, such as multiple linear regressions, like Holt's, Naïve and moving average (MA). The results indicate that the ANN is the better perform comparisons (Law, 2000). The assessments of ANN with traditional forecasting methods have been explored with the techniques of Arima and other regression methods from different researchers. The result, ANN is better than the conventional methods (Kihoro et al., 2004; Fradinata et al., 2015). The simulation of mitigating the bullwhip effect, allocating the safety stock, relates to the VMI to mitigate the bullwhip effect, and loss of sales by minimizing the safety inventory and considers the Burbidge effect and Houlihan effect in the research (Kristianto et al., 2012; Disney and Towill, 2003). The investigate of properties from the ground are examined by combining the stochastic data to be a variable input, using Gaussian random distribution to simulation MC distribution collaborate with the space of the entry of ANN network (Hattab et al., 2013). The companies observe the fluctuation to minimize the inventory for balancing to the customer demand. The suppliers and retailers are aware of the bullwhip effect phenomenon which can influence the company's inventory for each supply chain member.

Safety stock utilization is undertaken to forecast the demand of fluctuation (Zequeira et al., 2008). The supplement of demand forecast from a retailer should overreact to manufacturing (Jüttner et al., 2007). The time series of demand dataset and the stationarity from the normal random distribution would create the bullwhip effect in the supply chain

member (Chandra and Kumar, 2000). Based on the industrial level the bullwhip effect is prevalent among individual companies present from all firms and products that have been analysed by the information of aggregate data. The lack of information sharing forces all members of the supply chain to anticipate the space by doing the stockpile. The preparation of the stockpile needs to anticipate the degree of uncertainty conditions (Nyoman Pujawan et al., 2014). It leads to becoming poor customer services, misproduction of schedules, incorrect capacity plans, inefficient shipping, and high cost to the manufacturers.

The findings of the study are developing the model of ANN to minimize the bullwhip effect in the supply chain with the indicator of the small MSE from the signal demand of ANN forecasting models.

The structures parts of this paper are managed by sections are as follows. Section two reviews the work on the ANN method, AHP, MC methods, determinant the demand variables, and measure the accuracy method where it sets some methodologies are used, and test of parameters of ANN, robustness test, the signal demand, the bullwhip effect. On section three, it shows the empirical result where they found from the research. The section four shows that the empirical result from the study. The discussion is in the fifth section where it discusses all the effect and advantage from the models and reduces the bullwhip effect in the supply chain member. The final section summarizes the conclusion of the paper where it reflects the performance of hybrid methods compares to the original ANN and the relationship to the bullwhip effects in the supply chain member, retailer, wholesaler, distributor and factory.

2. The purpose of signal demand hybrid models

The main objective of this study is to minimize the signal demand from hybrid ANN models to reduce the bullwhip effect from downstream to upstream in the supply chain. The methodology of this study can be in Fig. 1. The initial data are generated from normal random distribution to obtain the most optimal conditions for the training process in the network. In the preprocessing data, the generating random normal datasets use Durbin-Watson method to select the variables in autocorrelation, and significant collinearity relationship among variables (Fradinata et al., 2015).

Formerly, it makes the scaling the data with the mapstd function to drop the data on the certain value for input variables and works to match to the transfer function in the network system. The determinant of demand is as input variables, AHP and MC are as the weight of entry at the weight of entry, the geometric random substitute the bias in the hidden layer. Furthermore, the network is tested with the part of 10 data to generate the test the model and 10 data to validate it. The input variables cooperate with the weight and bias extremely before producing the mean square error through the algorithm.
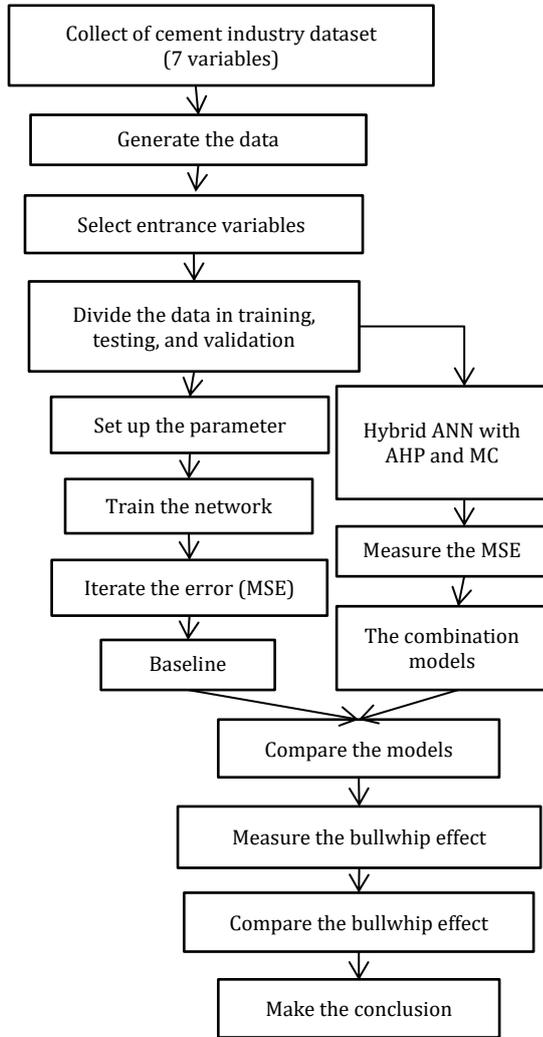
Fig. 1: The methodology

The approach of the neural network of time series forecasting or auto regression model is defined as a general formula, $Y_t = \mu + Q_1 \cdot Y_{t-1} + e$, where μ is constant, Q_1 is the weight, Y_{t-1} is the input variables to the network of ANN, and e is the bias. There are some parameters of the neural network; they are a training function, activation function, and neurons. Training function has the function to adjust or arrange the weight, and bias in the network estimates the parameter error in the model. The transfer function is used to iterate the weight and bias to form the mean square error smaller than the initial error; the input variables suggests the same range with the specific range of activation function, lease it work more smoothly. It has been pointed out that as produce the mean square error is commonly used to obtain the optimal connection among the weights; the data need to be normalized in the certain range in order just to obtain optimal results.

2.1. The AHP and Monte Carlo

The transformed data works from the input data to be the system with Analytical Hierarchy Process software. The result from raw data is converted to AHP weight model. The process of Analytical Hierarchy Process obtains from the transform data

to the Analytical Hierarchy Process system (Saaty, 2004; Ciptomulyono, 2008). The notation of *w* is the weight into each variable. The modified weight from Analytical Hierarchy Process replaces the default weight of a network in the hidden layer. Then, it works with the input from the input layer in hidden layers to generate to be better MSE than the initial case.

The output of Monte Carlo (MC) simulation results from the normal random generation to form the random draw, and some samples are taken from them. This method is chosen because the MC simulation generates a pool of probability data in the center of the data. This output should produce better data close to each other. The samples from MC are selected to be the weight of the system of ANN. The modified weight from MC simulation works together with the input weight in the area of the hidden layer. The data is iterated to obtain the smaller error than the initial error in the hidden layers in the training process. The output of AHP and MC substitute to the input weight and input layer in the ANN system. It can be seen in Fig. 2.

The algorithm feedforward backpropagation:

Phase-1. Input
1. Each input unit receives the signal and forwards it to the hidden layer. z_j, (j=1, 2, ...).

$$z\,net_i = v_{jo} + \sum_{i=1}^{n} x_i v_{ji}$$
$$z_j = f(z\,net_i) = \frac{1}{1+e^{-znet_j}}$$

2. Calculate all outputs in hidden y_k, (k=1, 2, ..., m)

$$y\,net_k = w_{ko} + \sum_{j=1}^{p} z_j w_{kj}$$
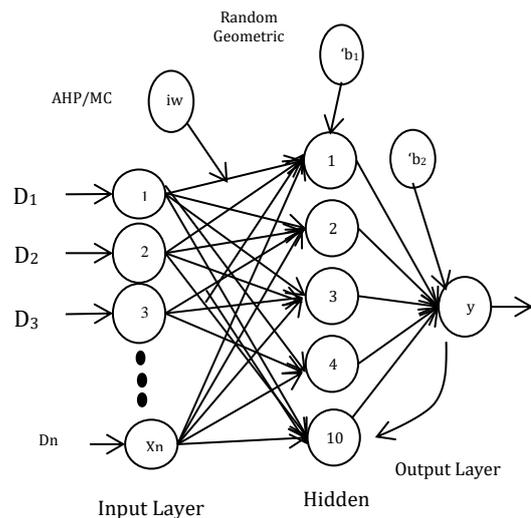$$y_k = f(y\,net_k) = \frac{1}{1+e^{-ynet_k}}$$

Fig. 2: The Purpose hybrid ANN network with the backpropagation neural network system

3. Calculate all network output in the yk=k=1, 2, ..., m)

Phase-2. Backpropagation

4. Calcuate the factor δ arithmetic unit output by an error in each unit of output yk(k=1,2, … ,m).

$$\delta_k = (t_k - y_k)f'(ynet_k) = (t_k - y_k)y_k(1 - y_k)$$

5. The δ_k Is a unit of error that will be used in a layer below the weight change.

$$\Delta w_{ij} =\propto d_k z_j \ ; k = 1,2,….,M: j = 0, 1, …. , p$$

6. Calculate the weight change rate x with the accelerated pace of z_j $(j = 1,2, …. , p)$

$$\delta net\_j = \sum\_(k = 1)^m \llbracket \delta\_k \ w\_kj \rrbracket$$

7. The factor δ in the hidden layer:

$$\delta_j = \delta net_j f'(znet_j) = \delta net_j f'(znet_j) = \delta net_j \ z_j(1 - z_j)$$

8. Calculate the change in weight, v_{ji}.

$$\Delta v_{ji} = \alpha \delta_j x_i \ ; j = 1, 2, …, p; i = 0, 1, …, n.$$

Phase 3. Change the weight
9. Calculate all weight changes.

$$w_{kj}(new) = w_{kj}(old) + \Delta w_{kj}, \ (k = 1,2, …, m \ and \ j = 0,1, …, p)$$

10. The weight change to the hidden layer is leading to output.

$$v_{ji}(new) = v_{ji}(old) + \Delta w_{kj}, \ (j = 1,2 … , \ m \ and \ i = 0,1, …, p)$$

Fig. 2 The $x_1, x_{2,…,}$Xian is scalar demand. Input, w_1 is the network weight; b_1 is the bias in the network of ANN. The scalar input transmitted through a connection that multiplies its strength by the scaled weight with the transfer function of in the hidden layer. The weight updates equation for the classical Newton algorithm (Jek Siang, 2005).

$$w_{n+1} = w_n - H^{-1\nabla_{w_n}E} \tag{1}$$

Eq. 1 shows that H^{-1} is the inverse of the Hessian matrix where the iteration process needs space to give to the free parameter number (Kwok and Yeung, 1997). The hidden layers continue the process to iterate before sending the process to the output layers. The modification of the weight can be substituted for the input weight and layer.

2.2. Signal demand accuracy

There are various measures of signal forecast errors. One of them is a *mean square error (MSE).* The MSE is the square quantity the unit term of an equal to the original value, which is the difference between the actual value and an estimated value, it can be defined in (2).

$$MSE = \frac{\sum_{i=1}^{n}(Y_t - \hat{Y}_t)^2}{n} \tag{2}$$

where:

Y_t is an actual value for period t,
$\hat{Y}_t$, is a period value of forecast t, and
n is a period number

Cross-validation is producing an independent assessment of the best signal-forecasting model to generalize the initial data. The substance in cross-validation is the normal random split data into two parts, one training set and another is a test set. The last the neural network used for training and estimating the model's performance (He et al., 2014).

2.3. Bullwhip effect

The signal demand results the forecast demand for hybrid ANN models. The supply chain assumes that the external customer demand behaviour has the normal random distribution.

The normal density has the function of f(x) which is called as the bell-shaped curve that is symmetric around μ, and attains maximum value of $1/\sigma 1/\sigma\sqrt{2\pi}$ ≈0.399/σ at x = μ. Hence the variance of the distribution is

$$Var(X) = \sigma^2 \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy = \sigma^2 \tag{3}$$

Where the density function of a normal random variable with parameters μ=0 and σ = 1, and its integral must be 1. The random normal variables that if X is normal with mean and variance, then for any constants a and b, b ≠ 0, the random variable Y=a+bX is also a normal random variable with the parameters

$$E[Y] = E[a + BX] = A + Be[X] = a + b\mu \tag{4}$$

and variance

$$Var(Y) = Var(a + bX) = b^2 Var(X) = b^2\sigma^2 \tag{5}$$

the distribution is written as

$$D_t = \frac{1}{|b|} fx\left(\frac{y-a}{b}\right) \tag{6}$$

$$D_t = f(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi\sigma|b|}} e^{\frac{-(y-a/b-\mu)^2}{2\sigma^2}} \tag{7}$$

That means showing Y = a + box is normal with mean a +bμ and variance $b^2\sigma^2$, then it follows the normal function with X∿N (μ,σ^2) then Z = (X-μ)/σ. Z is random variable called to have a standard normal distribution function when X is normal with parameters μ and σ^2.

The forecasting techniques, lead time, and the variant which has observed from the customer demand that put on the retailer from external demand, it is known as an independent, identically and normal (iidn). The independent random variables, let assume that X and Y are independent if any two sets of real numbers A and B with P{$X \le a, \ Y \le b$} = P{$X \le a$}P{$Y \le b$}, where the

joint distribution F of X and Y, then X and Y are independent if $F(a,b) = F_X(a)F_y(b)$ for all a, b. In the jointly continuing case the independent has the, equivalent of f (x, y) = FD $_{(x)}$ FY $_{(y)}$ in the condition of all x and y.

In particular, for any *n* sets of real numbers $A_1, A_2, ..., An$,

$$P\{X_1 \in A_1, A_2, \in A_2, \ldots \ldots, X_n \in A_n\} = \int_{A_{n-1}} \ldots \int_{A_1} f(x_1 \ldots \ldots, x_n)dx_1 dx_2 \ldots dx_n \qquad (8)$$

The concept of independence is defined for more than two random variables. In general, random variables $X_1, X_2, \ldots, X_n$ are said to be independent if for all set of real numbers $A_1, A_2, \ldots \ldots, A_n$,

$$P\{X_1 \in A_1, A_2, \in A_2, \ldots \ldots, X_n \leq a_n\} = \prod_{i=1}^n P\{X_1 \leq a_i\}$$
for all $a_1, a_2, \ldots, a_n$ (9)

The infinite collections of random variables are an independent if each finite sub-collection of them is independent. The mean and variance distribution is independent with (n-1) S^2/σ^2 which having the chi-square distribution with n-1 degree of freedom. To start, the number of $x_1, \ldots, x_n$, let yi = xi - μ, I = 1, ...,n. Then, as $\bar{y} = \bar{x} - \mu$, it follows from the identity of

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - n\overline{y^2} \qquad (10)$$

that

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n (x_i - \mu)^2 - n(\bar{x} - \mu)^2 \qquad (11)$$

If $X_1, \ldots, X_n$ is a sample from a normal distribution which having μ and σ^2, then it obtains from the preceding identity that

$$\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} + \left[\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}\right]^2 \qquad (12)$$

Due to $(X_i-\mu)/\sigma$, i=1,....,n are independent standard normal, where the left side of (12) is a chi-square random variable with n degree of freedom, $\sqrt{n}(\bar{X} - \mu)/\sigma$ is standard normal random variable and, its square is a chi-square random variable with 1 degree freedom. It has been established that the sum of two independent chi-square random variables is also chi-square with a degree of freedom. Thus, it would be seem that there is a reasonable possibility that the two terms on the right side (12) are independent with $\sum_{i=1}^n (X_i - \bar{x})^2/\sigma^2$ having the chi-square distribution with n-1 degree of freedom.

The independent and authentic has tested by the L-Jung Box method where we found the probability of α > 0.05. Then, the normality has tested the distribution with the Kolmogorov-Smirnov method with spss software.

The signal demand shows the fluctuation of variance demand from a customer in each period of lead time. It can be seen in Fig. 3.

We assume that the replenishment lead-time is constant and the external demand observes as the lead-time (*L*) in each period of *t* where it starts at the each next period of *t+L*. The continuing demand forecasters have been triggered from the period *t* to *t+1* where it has not reached the inventory due to the lead time. Then it fills the inventory follow on optimal order quantity in each period of *t*. The actual order starts from period *t* with the normal random distribution or stochastic demand (D_t). It is called the standardized normal distribution with a mean (μ) is 0 and variance (σ^2) is 1 at the demand during a lead time on period t, where **-∞**< x < ∞. It can be seen in (13).

$$D_t = f(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi\sigma|b|}} e^{\frac{-(y-a/b-\mu)^2}{2\sigma^2}} \qquad (13)$$
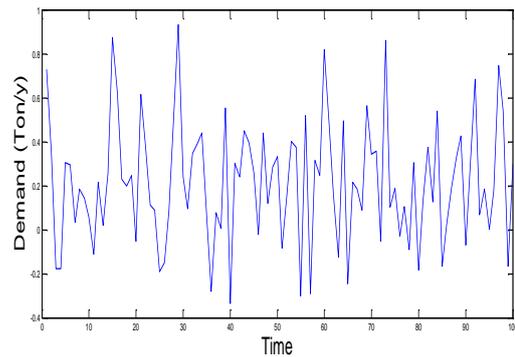

Fig. 3: Signal demand for hybrid ANN models

The external forecast demand put on the retailer in the supply chain. It fills that the inventory demand follows by each period t. In many practical situations, the lead-time follows the normal random distribution. The order quantity is calculated by (14). The assumption of lead-time, holding cost, and order cost are constants. *S* is an ordered cost, D $_{(t+1)}$ is demanded during each period, the *I* is an inventory cost, *i* is an investment cost, and C is a carrying cost. Thus, the costs are as a holding cost; it is *h* since the optimal order quantity is in (14).

$$TC = S\frac{D_{t+1}}{Q} + \frac{(I+i)(C)}{2}Q$$
$$\frac{SD}{Q^2} = \frac{(I+i)(C)}{2}$$
$$(I + i)(C)Q^2 = 2SD_{t+1}$$

where:

$$h = (I + i)(C)$$
$$Q_t = \sqrt{\frac{2 * D_{t+1} * S}{h}} \qquad (14)$$

To determine the policy in the situation where the *Q* satisfies in distribution continue to order policy 's hard situation due to the difficulty of the probability of external demand at the warehouse of the retailer. This study is assumed that *Q* = 0, this means that the *s* policy where the inventory position is kept at *s* all the time, then the distribution is normally distributed. As the order up to the policy, the order as continuing when the order drop below *s* and *Q* automatically order, the inventory system is designed in the computer program. The retailer

magnifies the external demand to anticipate the uncertainty condition of demand. The retailer magnifies it with a deviation of normal random deviation from the customer demand plus the optimal order quantity. It displays on (15).

The function of *NormInv* computes the inverse of the normal cumulative distribution function (cdf) using the corresponding μ and σ (sigma) at the corresponding probabilities where they can be a multidimensional vector arrays that all have the same size. A scalar input is expanded to a constant array with the same dimensions as the other inputs. The parameters in SD must be positive, and the values in P must lie in the interval between 0 and 1. The *pcov* is to estimate the parameters with the specifies alpha 100 (1 - alpha) % confidence bounds. The *NormInv* is computing the $\hat{\mu} + \hat{\sigma}q$. The normal inverse function is defined as $x = F^{-1}(p \mid \mu, \sigma) = \{x : F(x \mid \mu, \sigma) = p\}$. Where the p from (13) is changed to (15) with the normal distribution is as follows:

$$D_t = f(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{15}$$

where (14) is the optimal order quantity to be (16)

$$AoQ_{t+1} = Q_{t+1} + NormInv \text{ (Rand. ('), μ, SD)} \tag{16}$$

Where the AoQ_{t+1} is actual order quantity where it is a normal random deviation from a usual order from a retailer, Q_{t+1} is optimal order quantity, and the deviation of normal random probability is *NormInv* (Rand. ('), μ, SD). The order quantity (16) is calculated with the assumptions that the lead-time, holding cost, and order cost are assumed constants.

The delivery of the pipeline to fulfill the safety stock per year for each member follows in (17) where the D $_{(t+1)}$ is the observed demand for the next member, Q_{t+1} is the optimal order quantity each period.

$$Delivery = \frac{D_{t+1}}{Q_t}. \tag{17}$$

The observe demand is calculated on (18).

$$Observe\ demand = Delivery * AoQ_{t+1} \tag{18}$$

In practice, the retailers usually calculate the average and standard deviation of demand when the customer demand is observed. The retailer estimates the mean lead-time demand as averages of the previously observe demand. Thus, if D_i represents customer demand in period *i*, and *p* is the period, then the mean and standard deviation are illustrated (19) and (20).

$$\hat{\mu}_t = \frac{\sum_{i=t-p}^{t-1} D_i}{p} \tag{19}$$

and

$$S_t^2 = \frac{\sum_{i=t-p}^{t-1}(D_i - \hat{\mu}_t)^2}{p-1} \tag{20}$$

The retailer calculates the mean and standard deviation to get the variation of the order, and the variety of demand shows in (21). It shows that the Var ($O_t/\mu_{(o\_t)}$) is the variability of customer demand and the Var ($D_t/\mu_{(Dt)}$) is the variance in order at the supply chain member. The ratio of the variance (21) is called the bullwhip effect; it occurs when the ratio more than 1.

$$\frac{Var\left(\frac{O_t}{\mu_{o_t}}\right)}{Var\left(\frac{D_t}{\mu_{D_t}}\right)} = \frac{V_o}{V_D} \geq 1 \tag{21}$$

The bullwhip effects where V_o is the variance order and V_D is the variance demand in the supply chain (Kristianto et al., 2012; Lee et al., 2004) shows in (21).

The unstable conditions are caused by the production and the demand on the chain will produce the bullwhip effect; it will be making the small and the high ratio among them. If the ratio of bullwhip effect less than 1, it is considered good. Then, if the ratio more than one, it indicates that the bullwhip effect occurs.

3. Empirical result

The results of forecasting model show in Fig. 3. The illustration that the points are the prediction signal of forecasting from points 80 to 100 after the training process. The rest of 20 datasets are divided to be tested and validating cases. The testing and validating process are completed randomly from the training data to predict the 20 points from testing than it is validating to obtain the accuracy of prediction.

This condition is illustrated that represents the testing and validation data have the R-correlation values in good condition with the value of 1.000. That value represents the overlap line at the validation has a perfect performance of the signal demand models due to the process of the best-fit regression line between output and target relationship. The prediction error works on the global minima. The optimal conditions of validations are chosen for the value of 1.000 validating processes. The combined methods are called AHP$_{iw,}$ $_{il}$ANN$_{bn}$ and MC$_{iw, il}$ANN$_{bn}$ combinations.

The subscript notations show that the *b* (input bias), *iw* (input weight), and *il* (input layer). The abbreviations of hybrid symbols for ANN are as follows:

AHP_{iw}	*AHP weight substitutes at part of input weight of ANN. (Net.IW{I, J}).*
MC_{iw}	*MC weight substitutes at part of the input weight of ANN. (Net.IW{I, J}).*
ANN_{b1}	*Geometric random distribution works in the bias hidden layers of ANN. (Net.b{I}).*
ANN_{b2}	*Use 1 as a bias at output layer of ANN. (Net.b{J}).*

This study uses eight models where seven combination methods and one the original method.

The original ANN creates the new approaches from ANN with the components of AHP, MC, and random geometric distribution. The original ANN is as a baseline for comparison to the other combination methods. These new various methods are as follows: The symbol of i_W is the input weight, and b is the bias. These notations show the several of the combined methods within the neural network.

The selected parameters are applied to entrance the variables to the neural network feedforward-backpropagation. This study is used three layers; input layer, hidden layer, and an output layer with ten neurons, *trainlm* training function, and *tansig* as the transfer function.

The performances of MSE from forecasting models are indicating the difference perform from each model. The sample of AHP, MC, and geometric random are combined with original ANN in the network to iterate the performance of MSE, Table 1. The AHP is better than the MC combination ANN because the transform data of AHP produce the smaller weight than the MC. AHP weight has the ratio consistency index quite less than 0.1. This condition contributes to the process of random work in the network by iterating the better process to produce the mean square error than the MC. While the MC works in the normal distribution, then it is selected from the random work around the number of 0.2. Furthermore, the bias has the quite small value with the geometric random distribution; it means that the results of iteration process of these two factors work with the backpropagation algorithm have smaller values in the iteration process. The initial iteration on the random values works with the training process. The testing process from this hybrid method is tested after the training process finish, and then the validating works to compare the real data to the prediction. The cross-validation data must be quite a difference from the training process. It comes from the generating dataset or the real dataset of the determinant of demand. Then, the error is predicted based on the training data and testing dataset. The results of this process iterate the error to obtain the smaller than the initial iteration. The minimum mean square error (MSE) for the hybrid method it can be seen in Table 1.

Table 1: The mean square error from post processing signal demand of hybrid models in cement industry

Models	$AHP_{iw} ANN_{b1}$	$MC_{iw}ANN_{b1}$	$AHP_{iw}ANN$	$MC_{iw}ANN$
MSE	$8.08e^{+7}$	$7.54e^{+8}$	$8.53e^{+9}$	$9.97e^{+9}$
Models	$AHP_{b1}ANN$	$MC_{b1}ANN$	ANN_{b1}	ANN
MSE	$1.05e^{+10}$	$1.32e^{+10}$	$2.61e^{+10}$	$5.05e^{+10}$

The statistical information of variance demand ($\sigma_d = \frac{V_d}{\mu_d}$) with their SD, μ, σ, where they are in vary of hybrid models of ANN provided in the Table 2.

The output of standard deviation and mean are used to measure the order variance in the supply chain. Then the calculation obtains the bullwhip effect formula with the rule of it in Table 3.

At the first time, in the retailer, there is no bullwhip effect occurring due to the order nearly with the same frequency and nearly the same amount as the retailer did, but it did not reflect the retailer must hold a high level of safety stock which increases its inventory holding and carrying costs, Table 3. It is possible to say that the bullwhip effect is hidden in inventory level or the safety stocks. In the case of the external demand and the review is a continuous review, an order starts exactly the reorder point when the inventory position reaches the reorder point of *s*. The order arrives from the pipeline at the L time later when the average stock on hand; it is called the safety stock. The safety stock can be an additional stock when on the pipeline when arriving to process against demand variation from external demand.

Table 2: Statistical information on various combination methods for cement industry

Models	Demand			
	SD	μ	σ	σ_d *
$AHP_{iw}ANN_{b1}$	$1.16e^7$	$3.62e^8$	$1.36e^{14}$	$3.754e^5$
$MC_{iw}ANN_{b1}$	$1.04e^7$	$3.62e^8$	$1.09e^{14}$	$3.002e^5$
$AHP_{iw}ANN$	$1.30e^7$	$3.61e^8$	$1.68e^{14}$	$4.644e^5$
$MC_{iw}ANN$	$3.64e^7$	$3.58e^8$	$1.32e^{15}$	$3.692e^6$
$AHP_{b1} ANN$	$1.69e^7$	$3.6e^8$	$2.84e^{14}$	$7.89e^5$
$MC_{b1} ANN$	$1.18e^7$	$3.62e^8$	$1.38e^{14}$	$3.81e^5$
ANN_{b1}	$1.40e^7$	$3.62e^8$	$1.95e^{14}$	$5.38e^5$
ANN	$2.71e^7$	$3.63e^8$	$7.34e^{14}$	$2.02e^6$
* Note: $\sigma_d = \frac{V_d}{\mu_d}$				

Table 3: Statistical information about variance order in supply chain and BWE for $AHP_{iw}ANN_{b1}$ for cement industry

Order	Supply Chain			
	Retailer	Wholesaler	Distributor	Factory
SD	$1.16e^4$	$4.45e^7$	$8.81e^7$	$1.76e^8$
μ	$3.62e^5$	$1.38e^9$	$2.71e^9$	$5.40e^9$
σ	$1.36e^8$	$1.98e^{15}$	$7.77e^{15}$	$3.08e^{16}$
$\sigma_o = \sigma/\mu$	$3.75e^2$	$1.44e^6$	$2.86e^6$	$5.71e^6$
BWE, (σ_o/σ_d)	0.001	3.8	7.6	15.2

Moreover, It means that the safety stock to guaranty the satisfaction to reorder point condition. In practice, it must be understood to determine the service level is based on the stock on hand or safety stock. The safety stock can be chosen by the manager follow the condition of the filtrate. The level of filtrate can determine the service level where it is based on the probability *z* distribution to make sure the probability of a stock out during the lead time (1-α). The expected demand is in the interval between the range of t+L and t_{n+1}+L cannot meet the stock on hand, its demand is back ordered. The continuing review policy (s, Q) where the inventory position can never become negative; then we assume that the *s* ≥ 0 and it can be simplified assumption that Q ≥ *s*. It comes consequence that if one order is outstanding, the inventory position is strictly above the reorder point, therefore, that the order must be controlled automatically well by the program on the computer.

Table 4 shows that the retailer has ordered on the initial order quantity Q; then the retailer starts to magnify the information to the wholesaler to upstream in the supply chain. The bullwhip effect

increases from the retailer, wholesaler, distributor, and factory in the supply chain member. The increases of bullwhip effect due to the variance and standard deviation increase from the downstream to the upstream of the supply chain member. The variance follows the normal distribution as the deviation of magnifying the information from the retailer. The movement of data follows the characteristic of normal distribution where it spread from the center to the mean data point. In the bullwhip effect, the amplification of information and the order activities in the supply chain influences the preparation of stock inventory in every member of the supply chain. In this situation, the member of the supply chain must be able to anticipate the uncertain demand from external demand. The variance between the retailer to the factory level and for each model can be illustrated in Figs. 4 and 5.
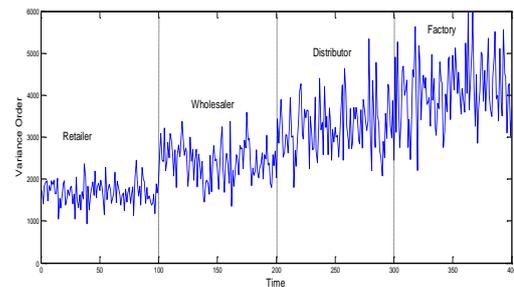
Table 4: The bullwhip effect from retailer to factory and the total bullwhip effect from varying hybrid methods in the cement industry

| Hybrid method	Retailer	Wholesaler	Distributor	Factory	Total BWE
$AHP_{iw}ANN_{b1}$	0.001	3.8	7.6	15.2	26.6
$MC_{iw}ANN_{b1}$	0.001	4.0	8.0	15.9	27.9
$AHP_{iw}ANN$	0.001	5.4	10.8	21.6	37.8
$MC_{iw}ANN$	0.001	5.7	11.2	22.4	39.3
$AHP_{b1}ANN$	0.001	5.3	10.5	20.9	36.7
$MC_{b1}ANN$	0.001	5.5	11.0	22.0	38.5
ANN_{b1}	0.001	5.8	11.5	22.9	40.2
ANN	0.001	5.9	11.7	23.4	41.3

The bullwhip effect increases from the retailer to the upstream level. The external demand is changing the standard deviation automatically and mean of the external demand is changed and plus the quantity order factor as well. Both those factors will magnify to increase with the normal random distribution. Let say the increase is 20% on periodically from the customer demand by the retailer.

The order quantity and the increase of a deviation factor from normal random distribution magnify the information to the wholesaler. This means that the reorder point is automatically increased from the retailer and the other member increase their order to the factory. The magnification is caused by the variance of the customer demand. The variations of variances are illustrated in Fig. 4 from retailer to factory level in each hybrid methods. Fig. 5 shows that one hybrid method of ANN for variance order or reorder point from downstream to upstream in the supply chain.
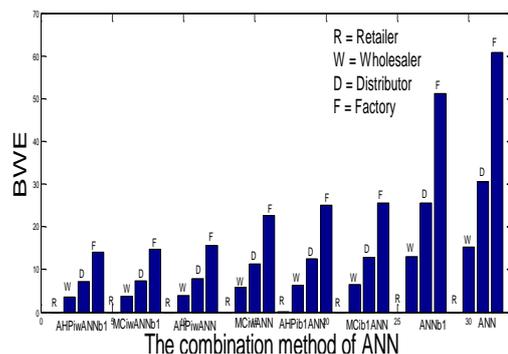
their warehouse to expect opportunity and advantage for the unpredicted situation from the customers' demand, it is illustrated Fig. 6. Many retailers do not want to lose their prospect and loss their customer on the situation of shortage on the stock inventory.

Fig. 5: The single variance order from one hybrid ANN model from downstream to upstream of the supply chain.

Fig. 6: The bullwhip effect from retailer to factory for each hybrid
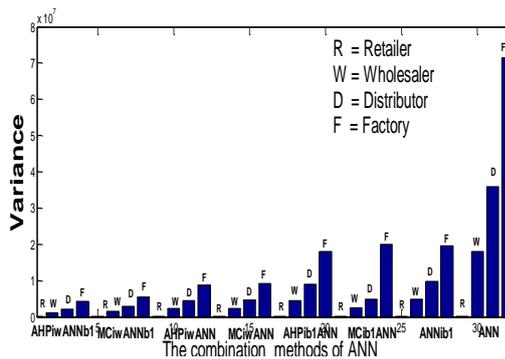
Fig. 4: The variance order from retailer to factory for each hybrid method

Based on the uncertainty condition of customer demand the lead-time also influences by the member order. If the lead-time short it is possible the demand from each member change should be short of small among the members of supply chain. The entire member of the chain will anticipate the probability of uncertain condition to keep the buffer stock in

The comparison of total bullwhip effect performance between the hybrid methods to original ANN in the cement industry is in Fig. 7, where the $AHP_{iw}ANN_{b1}$ and $MC_{iw}ANN_{b1}$ show that the total bullwhip effect performance increase 35.59% and 32.69%, respectively compare to the original ANN. The $AHP_{iw}ANN$ and $MC_{iw}ANN$ increase 8.47%, and 4.84%, respectively, compared to original ANN, the $AHP_{b1}ANN$ and $MC_{b1}ANN$ the performance increase 11.91% and 7.02% compared to original ANN respectively, and at last, the comparison of ANN_{b1}

increase the performance 3.15% compare to the original ANN. The average total bullwhip effect

performance to reduce the bullwhip effect is 14.81% for the cement industry.
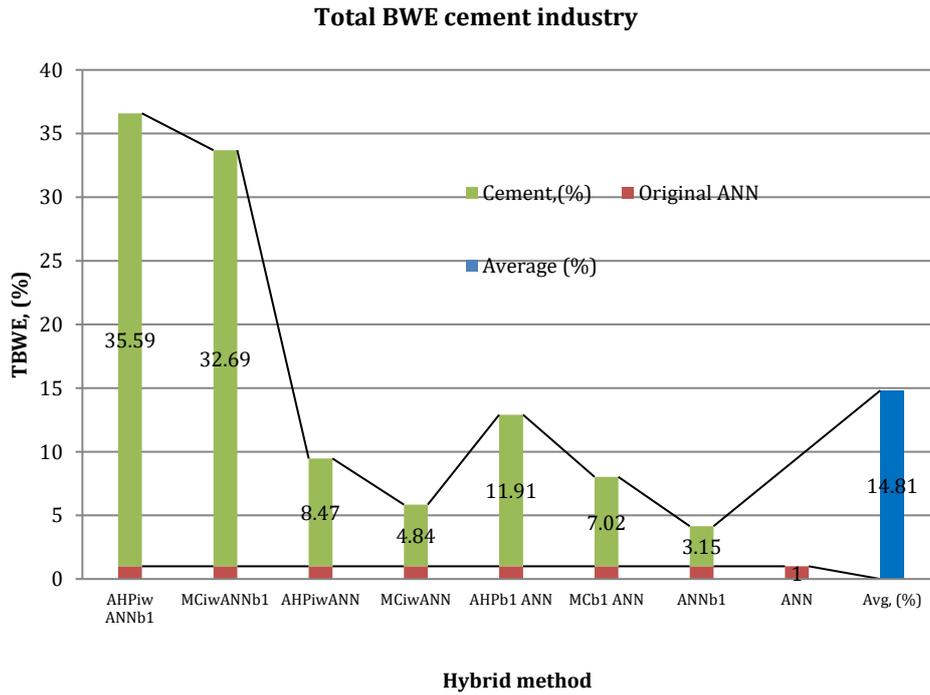
Total BWE cement industry

Fig. 7: Total bullwhip effect performance increase for cement industry

The advantage of ANN is the ability to adapt the data to the fitting process condition through the training process. The most combination methods have better performance than the original ANN method. The reasons are the biases run with the smallest means square error of geometric probability distribution than the other random distribution methods. The input weight and input layer work with the MC and AHP methods where they have the advantage to make the smallest result with close to zero than the other methods when the training process before it is validated with the actual data. The algorithm of backpropagation neural network is quite good when the robustness test it, means that the parameters work in the optimal conditions. Moreover, the measurement among the variance performance of hybrid models measured by the non-parametric test, and due to the sample has smaller than 30 observations of data. These study results are tested by the Friedman test method to measure the consistency of variance between the models due to the Friedman test can be used in a linear and nonlinear condition of data; it can be applied in a small amount of data and stationer in the mean and variance of the dataset. This test exam the differences in treatments across multiple test attempts of the data from the mean square error from the output of the models. The result can be seen in Table 5.

The results, the calculated of p-value 0.157 is greater than the significance level of α ($\alpha < 0.05$), mean that the performance of the MSE from different combination models did not make different variance occurred between the models from the cement industry.

4. Discussion

Neural networks are commonly popular is used to forecast demand variables. This study obtains the determinant of demand cement variables in Indonesia. It is utilized as input variables to the network, AHP, MC is entranced as input weights, and the geometric random substitutes the bias in the hidden layer. We divide into three parts of the process: *first* is the pre-processing phase. This process refers to supervise the neural network that is normalized the input variables from the determinant of demand dataset. It is normalized with the mapstd where this function drops the data in the range between 0 to1. The Durbin-Watson measures the autocorrelation, and the correlation amongs the correlation relation table measures the variables independent with the significant correlation. The result that we got the correlation higher number than the significant (sig<0.01) means that the correlation among the independent variables are weak and independent. *Second*, the chosen variables are transferred to the transfer function, which has relevance range to able to process in the system of the neural network. The system works with the optimal parameters for the detection of trends in the learning process. The

Table 5: Friedman test

N	8
Chi-square	2
'df	1
Asymp. Sig.	0.157

parameters are determined to identify the distribution to make the relevant patterns. The training process learns the pattern of data characteristic of data to iterate the error that works in the range of the transfer function capability at the upper bound and lower bound to produce the mean square error. At the end of the process, the data to reverse back to the real data. The advantage of ANN is the ability to adapt the condition through the retraining process when the input data entrance to the network. This study is set the training for 80 data. The 80 data elaborates together with the weight of AHP or MC, and with the bias of geometric random distribution. Due to input variables, the weight, and the geometric random variable are small enough to become smaller iteration than the initial in each iteration process. Hybrid models are obtained of input variables from AHP, MC and geometric random in the network. The AHP combination ANN is better than the MC combination ANN, due to the transform data to AHP produces the smaller weight due to the AHP weight to validate the ratio consistency index smaller than 0.1. These conditions contribute to the process of random work small in the network to iterate the mean square error compared to the MC. The MC normal distribution is selected and has the random number up to 0.2. The modifications of ANN substitute on parts of an input weight, an input layer, and a bias. The AHP and MC entrance in the input weight and the input layer while the geometric random distribution entrance in the bias part. The various signal demand of cement industries are $AHP_{iw}ANN_{b1}$, $MC_{iw}ANN_{b1}$, $AHP_{iw}ANN$, $MC_{iw}ANN$, $AHP_{b1}ANN$, $MC_{b1}ANN$, and ANN_{b1}. The $AHP_{iw}ANN_{b1}$ is better performance MSE and BWE than the $MC_{iw}ANN_{b1}$. The $AHP_{iw}ANN$ is better than the $MC_{iw}ANN$. The $AHP_{b1}ANN$ is better than $MC_{b1}ANN$. Furthermore, the ANN_{b1} is better than original ANN. The $AHP_{iw}ANN_{b1}$ is better than $MC_{iw}ANN_{b1}$ because the value of AHP as input weight has smaller than the MC input weight. The performances of bullwhip effect increase from modifying the method to original ANN. The average performance of bullwhip effect for all combination methods, percentage increase 14.81 for the cement industry and 22.30 for cement industry compares to original signal demand ANN. This because the initial iteration on the random values of AHP and geometric random as backpropagation algorithm trained a bias with the AHP value is quite close to zero. It means that the result of an iteration process of these two factors collaborate should have the small result of prediction error. Then, the testing process from this hybrid method is tested after the training process, and then the validating process works to compare the real data to the predict dataset. Then, the error function from this hybrid minimizes the mean square error (MSE). *Third,* the post processing data of forecasting demand dataset. It has the transferred to the original dataset when the process got the smallest mean square error. The combination ANN with AHP is a better signal than ANN with MC.

At last, there are two methods to examine the results. F*irst,* the test of the algorithm uses to check the robustness of prediction accuracy Fig. 8. The variability of the neurons in the hidden layers of ANN network is measured to obtain the MSEs.
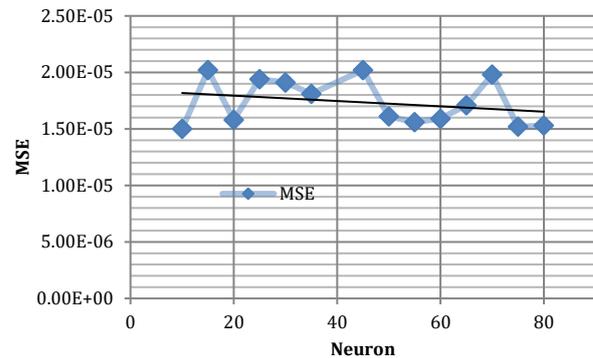

Fig. 8: The robustness of algorithm test

5. Conclusion

The neural network is a popular and widely uses a method for estimation the certain condition in the future. This study proposes the supervise back-propagation hybrid of ANN from the AHP, MC and the geometric random distribution to obtain the better performance than the original ANN. The modify methods between original ANN and additional methods are called the combination (hybrid) method of $AHP_{iw}ANN_{b1}$, and $MC_{iwl}ANN_{b1}$. The MSE is used to measure the performance of the model. The results of hybrid methods are produced the better performance than the original one. The smallest of MSE occurs at the $AHP_{iw}ANN_{b1}$ with the value of $3.67e^{+6}$. The comparison model of AHP is better than MC. The other hybrid models are also performed the excellent promising to be forecasting models compare to original ANN. The better forecasting method should be reducing the bullwhip effect in the supply chain member. The variance of signal demand from the customer is magnified by the retailer to the base stock level to protect the uncertainty customer demand in the next period. The modification methods decrease the variance demand to the effect of variance order. The best hybrid model is the $AHP_{iw}ANN_{b1}$ compared to other models. The AHP combination is better than MC combination to reduce the bullwhip effect in the supply chain. This study has the robust of the algorithm and no variance between the models.

6. Future work

The suggestion for the future work is trying to develop more models to hybrid with other original methods with data mining and artificial intelligence methods.

of this research and to the Prince of Songkla University, Hatyai, Thailand as well. Thomas Dryden, a lecturer of PSUIC from UK England who proofread this manuscript.

List of symbols

w_1	Network weight
b_1	Bias
z_j	Hidden layer
y_k	Outputs in hidden
δ_k	Unit of error used in the layer
Δw_{ij}	Weight change
μ	Mean
σ^2	Variance
π	Constanta
D_t	Signal demand normal random
D_{t+1}	Demand during period
S	Order cost
h	Holding cost
AoQ_t	Actual order quantity
Q_t	Optimal order quantity
NormInv(rand. (')m, sd)	Normal Random deviation (the amplify)
μ	Mean
sd	Standard deviation
V_o	Variance of order
V_d	Variance of demand
$\dfrac{V_o}{V_d}$	Bullwhip effect
e_n	Error forecast
n + p	Future period
Y_t	Time observation
$\hat{Y}_{n+1}$	Forecast value
n + 1	Period time
Z_t	Observation point
μ	A constant
θ	Moving average parameter
ⴲ	Autoregression parameter
Y_t	Actual value for period
n	Period number
y_0	Output signal
w_i	Initial Weight
x_i	Initial Variable
w_o	Weight
$f(x)$	Function of x
w^T	Transform weight
i_w	Input weight
i_l	Input layer
'$b_{1,2}$	Bias 1 and 2.
dL	Lower limit
dU	Upper limit
$x_1, x_{2, \ldots} x_n$	Demand scalar
i_w	Input weight
i_l	Input layer
'$b_{1,2}$	Bias 1 and 2.
dL	Lower limit
dU	Upper limit
i_w	Input weight
i_l	Input layer
'$b_{1,2}$	Bias 1 and 2.
dL	Lower limit
dU	Upper limit

References

Cachon GP, Randall T, and Schmidt GM (2007). In search of the bullwhip effect. Manufacturing and Service Operations Management, 9(4): 457-479.

Chandra C and Kumar S (2000). Supply chain management in theory and practice: A passing fad or a fundamental change?. Industrial Management and Data Systems, 100(3): 100-114.

Ciptomulyono U (2008). Fuzzy goal programming approach for deriving priority weights in the analytical hierarchy process (AHP) method. Journal of Applied Sciences Research, 4(2): 171-177.

Dejonckheere J, Disney SM, Lambrecht MR, and Towill DR (2003). Measuring and avoiding the bullwhip effect: A control theoretic approach. European Journal of Operational Research, 147(3): 567-590.

Disney SM and Towill DR (2003). Vendor-managed inventory and bullwhip reduction in a two-level supply chain. International Journal of Operations and Production Management, 23(6): 625-651.

Fradinata E, Suthummanon S, and Suntiamorntut W (2015). Forecasting determinant of cement demand in Indonesia with artificial neural network. Journal of Asian Scientific Research, 5(7): 373-384.

Hattab N, Hambli R, Motelica-Heino M, and Mench M (2013). Neural network and Monte Carlo simulation approach to investigate variability of copper concentration in phytoremediated contaminated soils. Journal of Environmental Management, 129: 134-142.

He Z, Wen X, Liu H, and Du J (2014). A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. Journal of Hydrology, 509: 379-386.

Jek Siang J (2005). Jaringan syaraf tiruan pemrograman menggunakan matlab. Penerbit Andi, Yogyakarta, Indonesia.

Jüttner U, Christopher M, and Baker S (2007). Demand chain management-integrating marketing and supply chain management. Industrial Marketing Management, 36(3): 377-392.

Kihoro J, Otieno R, and Wafula C (2004). Seasonal time series forecasting: A comparative study of ARIMA and ANN models. Asian Journal of Science and Technology, 5(2): 41-49.

Kristianto Y, Helo P, Jiao JR, and Sandhu M (2012). Adaptive fuzzy vendor managed inventory control for mitigating the Bullwhip effect in supply chains. European Journal of Operational Research, 216(2): 346-355.

Kwok TY and Yeung DY (1997). Constructive algorithms for structure learning in feedforward neural networks for regression problems. IEEE Transactions on Neural Networks, 8(3): 630-645.

Law R (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. Tourism Management, 21(4): 331-340.

Lee HL, Padmanabhan V, and Whang S (2004). Information distortion in a supply chain: The bullwhip effect. Management Science, 50(12_supplement): 1875-1886.

Lenny Koh SC, Demirbag M, Bayraktar E, Tatoglu E, and Zaim S (2007). The impact of supply chain management practices on performance of SMEs. Industrial Management and Data Systems, 107(1): 103-124.

Nikdel N, Nikdel P, Badamchizadeh MA, and Hassanzadeh I (2014). Using neural network model predictive control for controlling shape memory alloy-based manipulator. IEEE Transactions on Industrial Electronics, 61(3): 1394-1401.

Nyoman Pujawan I, Er M, Kritchanchai D, and Somboonwiwat T (2014). Uncertainty and schedule instability in supply chain: Insights from case studies. International Journal of Services and Operations Management, 19(4): 468-490.

Saaty TL (2004). Decision making—the analytic hierarchy and network processes (AHP/ANP). Journal of Systems Science and Systems Engineering, 13(1): 1-35.

Zequeira RI, Valdes JE, and Berenguer C (2008). Optimal buffer inventory and opportunistic preventive maintenance under random production capacity availability. International Journal of Production Economics, 111(2): 686-696.