A Conceptual Model for Uncertainty Demand Forecasting by Artificial Neural Network and Adaptive Neuro - Fuzzy Inference System Based on Quantitative and Qualitative Data

DOI: 10.46970/2020.26.4.3

Volume 26, Number 4

December 2020, pp. 285-302

Premporn Khemavuk

Atthiwat Leenatham

*King Mongkut's University of Technology North
Bangkok*

(ppm@kmutnb.ac.th)

(athiwat.lnt@gmail.com)

The purpose of this research is to present the new concepts for demand forecasting using artificial intelligence methods. In the first part, it demonstrates the evolution of demand forecasting from the past using traditional forecasting methods to the present using artificial intelligence forecasting methods. ANN and ANFIS were presented in this study with quantitative and qualitative data. The structure construction of the model is described to create various models in both the single forecasting method and the combined forecasting method to gain the best accuracy. There are two research questions as follows. 1. Are proposed methods with qualitative data more accurate than the one without qualitative data? 2. Is combined method forecast more accurate than single method forecast?

Keywords: Demand Forecasting, Artificial Neural Network, Adaptive-Neuro Fuzzy Inference System, Combined Forecasting Method.

1. Introduction

“Supply Chain” is used to describe the flow of goods from the first process in production to the last process delivered to consumers. Most organizations pay attention to supply chain management as it covers all activities. It also increases competitive advantages and benefits for both of consumers and stakeholders at a maximum level (Heizer, Render and Munson, 2017). Supply chain management is associated with strategic management from upstream to downstream of the flow of goods, services, finance, and information in each activity (Mentzer et al., 2001; Maestrini et al., 2017; Lima-Junior and Carpinetti, 2019). Various factors affect supply chain efficiency and one of the most important factors affecting planning of all departments in an organization is forecasting accuracy. Forecasting is one of significant leading procedures in supply chain management (Fildes, Goodwin and Lawrence, 2006) as it enables organizations to reduce risk and uncertainty from customer demands (Gamze, 2015). If forecasting is not accurate, a supply chain manager shall make a huge inventory to reduce the risk of out-of-stock inventory (Christopher and Lee, 2004). Most organizations do not know the exact demand and that is why they need to hold a large amount of safety stock by using forecasting of

demand (Asli, 2012). If the amount of inventory on hand is too much or too little, it will affect the inventory cost, an important cost to an organization. Therefore, if any organization can increase forecasting accuracy, it can help reduce the amount of safety stock and the organization will gain a higher profit and competitive advantage. In this research, different artificial intelligence and combined forecasting models are proposed with quantitative data and qualitative data.

2. Literature Review

“Demand forecasting” refers to the process of predicting customer future demand which historical sales data is most likely used to develop an estimate of expected forecast. In this regard, using only historical sales data can bring some problems when the sales data cannot represent actual customer demand, leading to products are out-of-stock (Feiler, Tong and Larrick, 2013). Thus, there are many methods using other factors to achieve more accurate forecasting (Chopra and Meindl, 2013). Forecasting can be classified into 4 basic types: qualitative methods use human judgment and opinions to predict a future, time-series methods use historical data as the basis for forecasting, causal methods find relationship of variables that affect forecasting, and simulation methods create alternative situations similar to customer behaviors.

Most previous studies forecasted customer demand using a time-series model such as moving-average technique, exponential smoothing technique and Box-Jenkins technique including causal methods like multiple linear regression and econometric models. Time-series model is a method using only historical data to forecast future sales volume and the exponential smoothing technique has been widely used in industries (Fildes, 1992) as it is a statistical forecasting method used to smooth average of data set which is different from a simple moving average method as more weights are determined to give importance to the most recent historical data. Related research studies like Chen, Drezner, et al., (2000) used a simple moving average method to assess the mean and variance of demand and to design a simple safety stock policy in accordance with a purchase order. Moreover, organization can reduce bullwhip effects by sharing these forecasting demand. Chen, Ryan, et al., (2000) indicated that an exponential smoothing technique from retailer shall trigger the bullwhip effect. The sized of increased variance depends on the nature of processes, customer demand, and forecasting techniques used by retailers, etc. Furthermore, there are ARIMA model (Auto-Regressive Integrated Moving Average) and similar ones such as ARIMAX, ARMA-GARCH, and ARFIMA models. These techniques have been popular since 1970 (Wei and Chen, 2012), these techniques were developed to cover trends, seasonality, unexpected situations, and unstable characteristics of time-series problems (Nikolopoulos et al., 2011; Box et al., 2016). Related research studies like (Kuo, Wu and Wang, 2002) used statistical methods, i.e., linear regression model and ARIMA model to make a comparison to be an option for forecasting. However, such methods are efficient to seasonal or cycle information only. In the event that information is influenced by a special case like promotion, these methods will not be efficient, similarly to the findings of Efendigil et al., (2009). They concluded that classical statistical methods are efficient to information having a seasonal pattern or similar trend. Though the aforesaid quantitative forecasting method is used for efficiently forecast, some restrictions can be seen as follows:

1. Lack of expertise can cause an error in the function connecting independent variables to dependent variables, leading to inefficient regression equation.
2. A huge amount of information is needed to achieve accurate forecasting.
3. It is not appropriate to non-linear information as it is difficult to build the model.
4. Outliers in the model can cause deviation in estimating model parameter.

Some restrictions can be improved by using an artificial neural network. A mathematical model which calculating functions can be variously selected (Garetti and Taisch, 1999; Efendigil, Öñüt and Kahraman, 2009; Asli, 2012, 2014). Nevertheless, it is important to make a choice for decision makers when unstable information and non-linear information appear (Khashei and Bijari, 2012).

In the past decade, the artificial intelligence method was used to improve forecasting efficiency. These research studies identified that an artificial neural network model was more efficient than ARIMA and classical regression equation models in terms of forecasting accuracy (Hong, 2010; Kuo, Tseng and Chen, 2014) while at the same time it was necessary to consider qualitative factors (Kuo, 2001; Kuo, Wu and Wang, 2002; Kuo, Tseng and Chen, 2014) due to higher complexity of markets and forecasting ambiguity from various factors. Thus, researchers have made attempt to seek a new method by combining all methods to gain a better outcome (Efendigil, Öñüt and Kahraman, 2009).

The combination of the neural network and fuzzy system is known as fuzzy neural network (FNN) and known today as adaptive neuro fuzzy inference system (ANFIS), one of methods being created to gain a better outcome. In literature review, such method was so little adapted to demand forecasting problems. Escoda et al., (1997) focused on language development and linguistic variables to forecast product demand by using ANN and FNN. Fuzzy Delphi and FNN were proposed to a stock trading decision support system by (Kuo, 1998). In other studies, R. J. Kuo & Xue, (1998) and R. J. Kuo et al., (2002) developed an intelligent sales forecasting system considering quantitative and qualitative factors using ANN and FNN. (Efendigil, Öñüt and Kahraman, 2009) investigated a comparison between ANN and ANFIS in demand forecasting. (Asli, 2014) proposed ANFIS method to forecast demand for fashion apparel products. R. Kuo et al., (2014) conducted ANFIS method combined with the artificial immune system adapted to the back-propagation neural network (BPNN) to forecast demand for laptop computers in Taiwan.

Based on the literature review, there was small number of studies that compared combined method forecast to single method forecast. Some studies used combination of ANN and ANFIS methods such as (Aras, Deveci Kocakoç and Polat, 2017) that made a combination among ANN, ANFIS, ARIMA, and ARFIMA. However, important qualitative data (Kuo, 2001; Kuo, Wu and Wang, 2002; Kuo, Tseng and Chen, 2014) were not included. Therefore, in order to overcome the limitations of previous studies, this research will present a new forecasting method using a combination of both ANN and ANFIS on quantitative and qualitative data. Later, forecasting accuracy will be measured with root mean squared error (RMSE), mean squared error (MSE), and mean absolute percentage error (MAPE). This research will mainly answer two research questions as follows. 1. Are proposed methods with qualitative data more accurate than the one without qualitative data? 2. Is combined method forecast more accurate than single method forecast? Both of two questions are based on the artificial intelligence methods as mentioned earlier.

3. Related Theory

3.1 Artificial Neural Network

Artificial neural network (ANN) is an artificial intelligent simulation model in accordance with structures and learning ability of human brain (Moosmayer et al., 2013; Jahangoshai Rezaee, Jozmaleki and Valipour, 2018). Though the classical artificial neural network was originated at the beginning of the decade of the 1940s by (McCulloch and Pitts, 1943), exploratory research was conducted (Rumelhart, Hinton and Williams, 1986) by presenting a back-propagation learning method. Such method created relationship between input variables and output variables efficiently and it was widely used for data forecasting (Zhou, Lai and Yen, 2012). Though there are many different types of ANN, the literature review related to adaptation of ANN in business (Smith-Miles and Gupta, 2000; Tkáč and Verner, 2015) indicated that multi-layer perception neural networks (MLP) are widely and efficiently used since MLP can forecast output variables using various input variables and can be adapted to be consistent with problems of each organization by creating different models to gain an outcome with the lowest degree of error. The structure of MLP is shown in the Figure 1 (Lima-Junior and Carpinetti, 2019), which is caused by a set of artificial neurons that are connected to each other by synaptic connection. It comprises of input layer, at least one hidden layer, and output layer. Each layer contains neuron or node to identify relationship of each variable. Nodes of each input layer show input variables of a model while nodes of output layer show output variables. Hidden layer is responsible for separating data characteristics by determining weight during the training process. Moreover, nodes of the hidden layer implement largely data processing in the MLP network (Akkoc, 2012; Tavana et al., 2016)

Figure 1 The Structure of Multi-Layer Perception Neural Networks

From Figure 2, artificial neurons consist of input signals $(x_1, x_2, \dots, x_n)$, weight values $(w_1, w_2, \dots, w_n)$, linear combination or summation $(\sum)$, threshold values (θ) , values obtained from a linear combination (Activation potential: u), activation function (g) , and output (y) . Each node has a linear combination to adjust input signals using weight values and threshold values to generate values from a linear combination. The equation (1) shows change in values from a linear combination to gain output values through activation functions such as logistic or sigmoid function, hyperbolic tangent function and linear function. Based on the Figure 1, overview of MLP is the connection of artificial neurons that each node and each layer are fully connected to the next one. The equation (1), input values of the hidden layer $(I_j^{(1)})$ is

calculated by the multiplication of input values and matrix of weights $W_{ji}^{(1)}$. The equation (2) is used to calculate output value of each node in each hidden layer ($Y_j^{(1)}$) through activation function. Similarly to the equation (3), input value of output ($I_j^{(2)}$) is calculated from the equation (4) to be output values ($Y_j^{(2)}$) using the equation (5) (Haykin, 2008; Lima-Junior and Carpinetti, 2019)

Figure 2 The Structure of Artificial Neural Network

$$y = g\left(\sum_{i=1}^n w_i \cdot x_i - \theta\right) \quad (1)$$

$$I_j^{(1)} = \sum_{i=1}^n w_{ji}^{(1)} x_i \quad (2)$$

$$Y_j^{(1)} = g(I_j^{(1)}) \quad (3)$$

$$I_j^{(2)} = \sum_{i=0}^{n1} w_{ji}^{(2)} Y_i^{(1)} \quad (4)$$

$$Y_j^{(2)} = g(I_j^{(2)}) \quad (5)$$

Learning techniques for weight value adjustment to get better results can be divided into 2 types as unsupervised learning and supervised learning. Supervised learning is adopted in this research study through back-propagation. It starts from calculating error between actual value and calculated value. Then a minimization procedure is used to adjust the weights between two connection layers starting backwards from the output layer, hidden layer to input layer. In minimization procedures, many algorithms are adopted to get the best value differently such as Gradient descent, Conjugate gradient, Quasi-Newton, and Levenberg-Marquardt. Different kinds of algorithms depend on weight adjustment (Manevitz, Bitar and Givoli, 2005; Haykin, 2008). Further reading on detailed explanation of learning algorithms is located at (Demuth and Beale, 2000; Haykin, 2008).

3.2 Adaptive Neuro-Fuzzy Inference System

Fuzzy inference system (FIS) was proposed by (Zadeh, 1965) to provide a solution for making decision based on vague, ambiguous, imprecise or missing data. Then it was developed to be adaptive neuro-fuzzy inference system by (Jang, 1993) to create relationship between input and output on the basis of human knowledge by employing fuzzy if-then rules and artificial neural networks to determine shapes of

member functions and separation logic. The pattern of ANFIS shown in the Figure 3 uses hybrid learning process of an adaptive network. This method plays an important role in initiating rules from observation of ambiguous logic (Demuth and Beale, 2000). For being easy to understand, the fuzzy inference system under consideration has two inputs x and y and one output f of Takagi-Sugeno fuzzy model (Takagi and Sugeno, 1983), which contains two fuzzy if-then rules shown as below:

Rule 1: If x is A_1 and y is B_1
Then $f_1 = p_1x + q_1y + r_1$

Rule 2: If x is A_2 and y is B_2
Then $f_2 = p_2x + q_2y + r_2$

When A_i and B_i are fuzzy sets, and $p_1, q_1, r_1, p_2, q_2, r_2$ are given from learning process.

Figure 3 The Structure of Adaptive Neuro-Fuzzy Inference System

In the ANFIS structure as shown in the figure 3 can explain for each layer:

Layer 1 Every node in this layer is a square symbol. Each node i^{th} in this layer are function

$$O_i' = \mu_{A_i}(x), \quad i = 1, 2. \quad (6)$$

Let x is input to node i^{th} and A_i as the language value (such as small, large, etc.) of the function on this node. In other words, O_i is a member function of A_i at level x that is defined in accordance with the quantifier, A_i . Basically, $\mu_{A_i}(x)$ is usually chosen as a bell-shaped function, which has a maximum value of 1 and a low value. The maximum is 0, which the general bell function can display as

$$\mu_{A_i}(x) = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i} \right)^2 \right]^{b_i}} \quad (7)$$

where $\{a_i, b_i, c_i\}$ is a set of member function parameters. When the values of these parameters change, the bell function will change to show the various forms of the A_i member function. The parameters in this layer are called the premise parameters.

Layer 2 Every node in this layer is a circle symbol. The symbol Π is the multiplication of the incoming signal and sending results as output below

$$O_i^2 = w_i = \mu_{A_i}(x) \mu_{B_i}(y), \quad i = 1, 2. \quad (8)$$

Each output of the node shows the firing strength of the rule.

Layer 3 Every node in the layer shows the symbol N , which calculates the ratio of the firing strength of rule i^{th} to the sum of the firing strength of all rules, shown as follows

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2. \quad (9)$$

To make it easier to understand, the output of this layer is called normalized firing strengths.

Layer 4 Every node in this layer is a square symbol, explained as function here

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i), \quad i = 1, 2. \quad (10)$$

where $\bar{w}_i$ is the output of layer 3 and $\{p_i, q_i, r_i\}$ is the parameter set, the parameters in this layer are called consequent parameters.

Layer 5 In this layer, there is only one node. The $\sum$ symbol represents the calculation of output, which is the sum of all incoming signals shown as follows

$$O_i^5 = \text{overall output} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}. \quad (11)$$

In the ANFIS structure, premise parameters and consequent parameters are important values obtained from learning as both parameters are used to find output values in data forecasting. It is interesting that when the values of the premise parameters are fixed, the overall output can be expressed as a linear combination of the consequent parameters. For further clarification, the output of ANFIS model can be written (Jang, 1993) as shown below

$$f = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2. \quad (12)$$

Substitute Equation 9 into Equation 12, resulting in equation 13.

$$f = \bar{w}_1 f_1 + \bar{w}_2 f_2. \quad (13)$$

Instead of the If-then fuzzy rule in Equation 13, will get equation 14

$$f = \bar{w}_1 (p_1 x + q_1 y + r_1) + \bar{w}_2 (p_2 x + q_2 y + r_2). \quad (14)$$

After rearranging, the f output can be represented as equation 15.

$$f = (\bar{w}_1 x) p_1 + (\bar{w}_1 y) q_1 + (\bar{w}_1) r_1 + (\bar{w}_2 x) p_2 + (\bar{w}_2 y) q_2 + (\bar{w}_2) r_2 \quad (15)$$

Hybrid learning in ANFIS consists of a forward pass and a backward pass. In the forward of the hybrid learning algorithm, node outputs go forward and the consequent parameters are identified while premise parameters are fixed using the least squares method. In the backward pass, the error signals propagate backward and the premise parameters are updated by gradient descent while consequent parameters are fixed. In this regard, both parameters values enable outcomes to be consistent with learning algorithms to be obtained the best accuracy between actual and predicted value.

3.3 Combined Method Forecast

Combined method forecast was firstly proposed for business forecasting by (Bates and Granger, 1969). Then it has been used successfully in empirical work such as forecasting gross domestic product, market fluctuation, currency, inflation rate, stock market, meteorological data, population number, risk (Timmermann, 2006). Researchers such as (Clemen, 1989) made empirical evidence summary in the literature review related to a combined method forecast that “outcomes can be clearly seen that a combined method forecast proposed by researchers leads to an increase in forecast accuracy. Sometimes we can amazingly improve efficiency from using methods to find an average.”

Combined method forecast is the combination of more than one forecasting method to create a new forecast while only one forecasting method has its own mechanism to create a forecasting value. The combined method forecast will integrate forecasts of the selected single methods. The combined method forecast is different from the hybrid algorithm techniques created to integrate various methods to be the same method with the intention to overcome disadvantages of each method. The combined method forecast has been presented in various techniques but in this research a simple and easy understanding combined method forecast is used so that it can be actually adopted with a tendency to lower levels of error and not to depend on any hypothesis. Based on these reasons, such method has gained popularity from forecasters (Aras, Deveci Kocakoç and Polat, 2017). The combined method forecast used in this research study includes the following:

Simple Mean is all forecasts at time t from the same method that shall determine the same weight value. Considerable studies indicated that this method is efficient (Aiolfi and Timmermann, 2006; Jose and Winkler, 2008; Aras, Deveci Kocakoç and Polat, 2017).

Least Square Weights is regression analysis technique from many single forecasting methods. Then regression coefficient is brought to be the weight value of a model (Granger and Ramanathan, 1984; Aiolfi and Timmermann, 2006; Aras, Deveci Kocakoç and Polat, 2017).

Mean Square Error Weights is a method to find weight value from MSE, which was introduced by (Stock and Watson, 1999). MSE value calculated from a period of forecasting by a single method forecast is used to calculate the weight of a certain single method forecast as seen in the equation 16.

$$w_i = (1 / MSE_i)^k / \sum_{j=1}^n (1 / MSE_j)^k \quad (16)$$

when k means a generated power number, if k is set to be 0, it shall have the same weight in a single method forecast. When k value is increased, emphasis is placed on a more efficient forecasting method. K = 1 is the most popular choice (Armstrong, 2001).

4. Proposed Method

4.1 Research Methodology

Research methodology of this study on artificial intelligence forecast comprises 3 major parts as collecting quantitative data, collecting qualitative data, and forecasting with single ANN, ANFIS models before using combined method forecast. The research methodology is shown in Figure 4. Data was divided into two parts, the first part (75%) is used for training and validating, the second part (25%) is used for testing. Data was normalized according to the equation 17 (Kuo, Tseng and Chen, 2014).

$$x'_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (17)$$

Where x'_i is the normalized value, x_i is the value before normalization, $x_{\max}$ and $x_{\min}$ is the maximum and minimum value of the data set before normalization.

Procedures shown in the Figure 5 indicate different outcomes from considerable methods that shall be compared in the study to answer the two research questions mentioned earlier.

4.2 Data Collection

(1) Qualitative data – qualitative factors affecting product demand based on the literature review of (Kuo and Xue, 1998; Kuo, 2001; Kuo, Wu and Wang, 2002; Aburto and Weber, 2007; Efendigil, Önüit and Kahraman, 2009; Kabir and Sumi, 2012; Kabir and Hasin, 2013, 2014; Kuo, Tseng and Chen, 2014) are shown in number 1 to number 6. The rest is factors that were considered in the first time in this research.

1. Product quality – this factor is related to product quality assessment in accordance with manufacturers and customers. Scores range from 1-9.
2. Levels of customer satisfaction – this factor shows sales behavior and after sales behavior of manufacturers given to customers. Scores range from 1-5.

3. Influence of seasons – this factor means an increase or decrease in percentage of sales volume associated with seasons.
4. Effect of competition – this factor shows competition (amount of companies manufacturing the same product) in the market.
5. Influence of promotion – this factor means promotional intensity released to the market at a certain time.
6. Influence of festivals and holidays – this factor shows effect of festivals and holidays taken place at a certain time.
7. Influence from data direction of previous periods – this factor indicates an effect from one-time prior data showing an increase or decrease in demand.

A questionnaire was made and submitted to experts to screen out the factors and remain 4 factors used as input variables to avoid too many variables in the models.

(2) Quantitative data – Time series data of product demand was collected to produce retrospective X -years forecast. Therefore, there will be data values to present data in monthly.

4.3 The structure of the Proposed Neural Network

A multi-layer artificial neural network was used in the study with back propagation learning as shown in Figure 4. There were 2 parts of input variables; quantitative variables and qualitative variables. Quantitative variables comprised of retrospective product demand at p value written as D_{t-p} to D_{t-1} . In this study, retrospective 3-month data ($p=3$) was use. In terms of qualitative variables, it consisted of 4 qualitative factors at time t written as $F_{1(t)}$ to $F_{4(t)}$. All qualitative factors were converted to be numbers to be used in the models. Output is product demand at time t written as D_t . The researcher designed the internal structure of artificial neural networks to be different for the purpose of comparison as follow:

Figure 4 *The Structure of the Proposed Neural Network*

The different structure of artificial neural networks was designed for the purpose of comparison as follow:

1. Activation function – a sigmoid or logistic function was used in hidden layers since values of input variables ranged between 0 to 1 and a logistic and linear function were used in output layers.
2. Number of hidden layers – in this study 1 and 2 hidden layers were used since a large number of hidden layers can enable forecasting in learning process to get over fitting and when they are forecasted with tested values, error will be extremely high.
3. Number of nodes in hidden layers – The literature review (Patuwo, Hu and Hung, 1993; Piramuthu, Shaw and Gentry, 1994; Lenard, Alam and Madey,

1995; Ali et al., 2011) introduced a guideline to find number of nodes in hidden layer from $n/2$, $0.5(n+m)$, $0.75n$, $2n$, $2n+1$; where n is number of input variables and m is output variables. Consequently, the number of nodes in hidden layers in this study was 4, 5, 6, 16, 17.

4.4 The Structure of the Proposed Adaptive Neuro-Fuzzy Inference System

Figure 5 *The Structure of Research Methodology*

The structure of adaptive neuro-fuzzy inference system presented in this study was Sugeno-fuzzy system. ANFIS used hybrid learning with least squares method and Gradient descent method to find premise parameters and consequent parameters

using input variables similar to the structure of artificial neural networks. Models are able to learn when the structure of ANFIS is determined as follow.

1. Input membership function is the function of quantitative data and qualitative factors that determines a degree of membership of all input variables. It is used to assign membership values that are vague or ambiguous because shapes of membership function are important to output variables. Membership functions used in this study were triangular shaped membership function (triMF), trapezoidal shaped membership function (trapMF), generalized bell-shaped membership function (gbellMF), and gaussian curve membership function (gaussMF).
2. Number of membership function is number of functions used to determine a degree of membership of each input variable. For example, if the number of membership function of historical demand values are set as 2, it means that the values of retrospective are high and low. If they are set as 3, it means that the values of retrospective demand shall be high, medium, and low. In this study, the number of membership function is set as 2, 3, and 4 to make a comparison of each method.
3. Output membership function is the function of output variables which can be divided into two types as constant function or linear function according to Sugeno-fuzzy principles. The difference can be seen in linear function comprising parameters p_i , q_i , and r_i as stated in the equation 10 but the constant function contains only parameter r_i . Therefore, outcomes are different. This study conducted a comparison of the two types as the obtained results are different and most research studies are required to consider both types.

4.5 Measuring forecast efficiency

Measuring forecast efficiency requires various values such as coefficient of determination, root mean square error (RMSE), mean square error (MSE), and mean absolute percent error (MAPE) as show in the equations 18, 19, 20, and 21.

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - F_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (18)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - F_i)^2}{n}} \quad (19)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - F_i)^2, \quad (20)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \left(\frac{Y_i - F_i}{Y_i} \right) * 100 \right|. \quad (21)$$

When F_t is determined as forecast value, Y_t is actual value and n is the number of data set. R^2 ranges between 0-1 represent the variance for a dependent variable that is explained by an independent variable. R^2 indicates high or low relationship between independent variables and dependent variables in forecasting models. As it is widely known that no methods can best measure errors (Makridakis et al., 1982; Armstrong and Fildes, 1995). Error measurement should be performed based on the 3 values in the equations 19, 20, and 21. The best model is selected from the errors that are consistently in the same direction. However, if contrast happens between of error's outcomes, MAPE shall be majorly considered since it is the most often used in industries and academic studies (Fildes and Goodwin, 2007; Fildes et al., 2009; Syntetos, Kholidasari and Naim, 2015).

5. Conclusion

This research firstly presents a new forecasting method with a combination of ANN and ANFIS models on quantitative and qualitative data. Not only a single method forecast, a combined method forecast was used to prove and answer the research questions as mentioned earlier in the literature review. This research study is to present a gap of research study, artificial intelligence methods, combined method forecast, research methodology, including a design of artificial intelligence model structure in details so that at least persons who pay attention to conduct the same study shall be able to read evolution of product demand forecasting and processes in designing such models. With regard to research management and research results shall be presented in the next issue.

6. References

1. Allwood, J.M., Laursen, S.E., Malvido de Rodriguez, C.M., and Bocken, N.M.P. (2006), *Well dressed? The present and future sustainability of clothing and textiles in the United Kingdom*, Cambridge: Institute for Manufacturing, University of Cambridge. Available from: www.ifm.eng.cam.ac.uk.
2. Aburto, L. and Weber, R. (2007) 'Improved supply chain management based on hybrid demand forecasts', *Applied Soft Computing*, 7, pp. 136–144. doi: 10.1016/j.asoc.2005.06.001.
3. Aiolfi, M. and Timmermann, A. (2006) 'Persistence in forecasting performance and conditional combination strategies', *Journal of Econometrics*, 135(1), pp. 31–53. doi: <https://doi.org/10.1016/j.jeconom.2005.07.015>.
4. Akkoç, S. (2012) 'An empirical comparison of conventional techniques, neural networks and the three stage hybrid Adaptive Neuro Fuzzy Inference System (ANFIS) model for credit scoring analysis: The case of Turkish credit card data', *European Journal of Operational Research*, 222, pp. 168–178. doi: 10.1016/j.ejor.2012.04.009.
5. Ali, S. M. et al. (2011) 'Forecasting of optimum raw material inventory level using artificial neural network', *International Journal of Operations and Quantitative Management*, 17(4), pp. 333–348.

6. Aras, S., Deveci Kocakoç, I. and Polat, Ç. (2017) 'Comparative study on retail sales forecasting between single and combination methods', *Journal of Business Economics and Management*, 18, p. 2017. doi: 10.3846/16111699.2017.1367324.
7. Armstrong, J. (2001) *Principles of Forecasting: A Handbook for Researchers and Practitioners*. doi: 10.1007/978-0-306-47630-3.
8. Armstrong, J. S. and Fildes, R. (1995) 'Correspondence on the selection of error measures for comparisons among forecasting methods', *Journal of Forecasting*. John Wiley & Sons, Ltd, 14(1), pp. 67–71. doi: 10.1002/for.3980140106.
9. Asli, A. (2012) 'A decision support system for demand forecasting in the clothing industry', *International Journal of Clothing Science and Technology*. Edited by O. Nursel. Emerald Group Publishing Limited, 24(4), pp. 221–236. doi: 10.1108/0955562211232829.
10. Asli, A. (2014) 'Demand forecasting for apparel manufacturers by using neuro-fuzzy techniques', *Journal of Modelling in Management*. Edited by Ö. Nursel. Emerald Group Publishing Limited, 9(1), pp. 18–35. doi: 10.1108/JM2-10-2011-0045.
11. Bates, J. M. and Granger, C. W. J. (1969) 'The Combination of Forecasts', *Journal of the Operational Research Society*, 20(4), pp. 451–468. doi: 10.1057/jors.1969.103.
12. Box, B. et al. (2016) *Time Series Analysis: Forecasting and Control*, *Journal of the American Statistical Association*. doi: 10.2307/2284112.
13. Chen, F. et al. (2000) 'Quantifying the bullwhip effect in a simple supply chain', the impact of forecasting, lead times, and information, 46(3), pp. 436–443. Available at: <http://www.scopus.com/inward/record.url?scp=0033872695&partnerID=8YFLogxK>.
14. Chen, F., Ryan, J. and Simchi-levi, D. (2000) 'The Impact of Exponential Smoothing Forecasts on the Bullwhip Effect', *Naval Research Logistics - NAV RES LOG*, 47, pp. 269–286. doi: 10.1002/(SICI) 1520-6750(200006)47:43.0.CO; 2-Q.
15. Chopra, S. and Meindl, P. (2013) *Supply chain management Strategy, planning and operation*. 5th edn. New Jersey: Prentice Hall.
16. Christopher, M. and Lee, H. (2004) 'Mitigating supply chain risk through improved confidence', *International Journal of Physical Distribution & Logistics Management*, 34, pp. 388–396. doi: 10.1108/09600030410545436.
17. Clemen, R. T. (1989) 'Combining forecasts: A review and annotated bibliography', *International Journal of Forecasting*, 5(4), pp. 559–583. doi: [https://doi.org/10.1016/0169-2070\(89\)90012-5](https://doi.org/10.1016/0169-2070(89)90012-5).
18. Demuth, H. and Beale, M. (2000) 'Neural Network Toolbox for use with MATLAB', 9.
19. Efendigil, T., Önüt, S. and Kahraman, C. (2009) 'A Decision Support System for Demand Forecasting with Artificial Neural Networks and Neuro-Fuzzy Models: A Comparative Analysis', *Expert Syst. Appl. USA: Pergamon Press, Inc.*, 36(3), pp. 6697–6707. doi: 10.1016/j.eswa.2008.08.058.
20. Escoda, I. et al. (1997) 'Demand forecast by neuro-fuzzy techniques', in *Proceedings of 6th International Fuzzy Systems Conference*, pp. 1381–1386 vol.3. doi: 10.1109/FUZZY.1997.619745.

21. Feiler, D., Tong, J. and Larrick, R. (2013) 'Biased Judgment in Censored Environments', *Management Science*, 59, pp. 573–591. doi: 10.1287/mnsc.1120.1612.
22. Fildes, R. (1992) 'The evaluation of extrapolative forecasting methods', *International Journal of Forecasting*, 8, pp. 81–98. doi: 10.1016/0169-2070(92)90009-X.
23. Fildes, R. et al. (2009) 'Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning', *International Journal of Forecasting*, 25(1), pp. 3–23. doi: <https://doi.org/10.1016/j.ijforecast.2008.11.010>.
24. Fildes, R. and Goodwin, P. (2007) 'Against Your Better Judgment? How Organizations Can Improve Their Use of Management Judgment in Forecasting', *Interfaces*, 37, pp. 570–576. doi: 10.1287/inte.1070.0309.
25. Fildes, R., Goodwin, P. and Lawrence, M. (2006) 'The design features of forecasting support systems and their effectiveness', *Decision Support Systems*. North-Holland, 42(1), pp. 351–361. doi: 10.1016/J.DSS.2005.01.003.
26. Gamze, O. K. (2015) 'Parameter optimization of intermittent demand forecasting by using spreadsheet', *Kybernetes*. Edited by D. O. Fahrettin. Emerald Group Publishing Limited, 44(4), pp. 576–587. doi: 10.1108/K-03-2015-0062.
27. Garetti, M. and Taisch, M. (1999) 'Neural networks in production planning and control', *Production Planning & Control*. Taylor & Francis, 10(4), pp. 324–339. doi: 10.1080/095372899233082.
28. Granger, C. W. J. and Ramanathan, R. (1984) 'Improved methods of combining forecasts', *Journal of Forecasting*. John Wiley & Sons, Ltd, 3(2), pp. 197–204. doi: 10.1002/for.3980030207.
29. Haykin, S. (2008) *Neural Networks and Learning Machines*, Pearson Prentice Hall New Jersey USA 936 pLinks. doi: 978-0131471399.
30. Heizer, J., Render, B. and Munson, C. (2017) *Operations Management: Sustainability and Supply Chain Management*. 12th edn. Pearson.
31. Hong, W.-C. (2010) 'Application of chaotic ant swarm optimization in electric load forecasting', *Energy Policy*, 38, pp. 5830–5839. doi: 10.1016/j.enpol.2010.05.033.
32. Jahangoshai Rezaee, M., Jozmaleki, M. and Valipour, M. (2018) 'Integrating dynamic fuzzy C-means, data envelopment analysis and artificial neural network to online prediction performance of companies in stock exchange', *Physica A: Statistical Mechanics and its Applications*, 489, pp. 78–93. doi: <https://doi.org/10.1016/j.physa.2017.07.017>.
33. Jang, J.-S. (1993) 'ANFIS Adaptive-Network-based Fuzzy Inference System', *Systems, Man and Cybernetics*, IEEE Transactions on, 23, pp. 665–685. doi: 10.1109/21.256541.
34. Jose, V. R. R. and Winkler, R. L. (2008) 'Simple robust averages of forecasts: Some empirical results', *International Journal of Forecasting*, 24(1), pp. 163–169. doi: <https://doi.org/10.1016/j.ijforecast.2007.06.001>.

35. Kabir, G. and Hasin, A. (2014) 'Application of adaptive neuro fuzzy inference system in demand forecasting for power engineering company', *International Journal of Industrial and Systems Engineering*, 18, pp. 237–255. doi: 10.1504/IJISE.2014.064708.
36. Kabir, G. and Hasin, M. (2013) 'Comparative Analysis of Artificial Neural Networks and Neuro-Fuzzy Models for Multicriteria Demand Forecasting', *International Journal of Fuzzy System Applications (IJFSA)*, 3, pp. 1–24. doi: 10.4018/ijfsa.2013010101.
37. Kabir, G. and Sumi, R. (2012) 'Integrating fuzzy Delphi method with artificial neural network for demand forecasting of power engineering company', *Management Science Letters*, 2, pp. 1491–1504. doi: 10.5267/j.msl.2012.04.010.
38. Khashei, M. and Bijari, M. (2012) 'Hybridization of the probabilistic neural networks with feed-forward neural networks for forecasting', *Engineering Applications of Artificial Intelligence*, 25, pp. 1277–1288. doi: 10.1016/j.engappai.2012.01.019.
39. Kuo, R. J. (1998) 'A decision support system for the stock market through integration of fuzzy neural networks and fuzzy delphi', *Applied Artificial Intelligence*. Taylor & Francis, 12(6), pp. 501–520. doi: 10.1080/088395198117640.
40. Kuo, R. J. (2001) 'Sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm', *European Journal of Operational Research*, 129, pp. 496–517. doi: 10.1016/S0377-2217(99)00463-4.
41. Kuo, R. J., Wu, P. and Wang, C. P. (2002) 'An intelligent sales forecasting system through integration of artificial neural networks and fuzzy neural networks with fuzzy weight elimination', *Neural networks: the official journal of the International Neural Network Society*, 15, pp. 909–925. doi: 10.1016/S0893-6080(02)00064-3.
42. Kuo, R. J. and Xue, K. C. (1998) 'An intelligent sales forecasting system through integration of artificial neural network and fuzzy neural network', *Computers in Industry*, 37(1), pp. 1–15. doi: [https://doi.org/10.1016/S0166-3615\(98\)00066-9](https://doi.org/10.1016/S0166-3615(98)00066-9).
43. Kuo, R., Tseng, Y. and Chen, Z.-Y. (2014) 'Integration of fuzzy neural network and artificial immune system-based back-propagation neural network for sales forecasting using qualitative and quantitative data', *Journal of Intelligent Manufacturing*, 27. doi: 10.1007/s10845-014-0944-1.
44. Lenard, M. J., Alam, P. and Madey, G. R. (1995) 'The Application of Neural Networks and a Qualitative Response Model to the Auditor's Going Concern Uncertainty Decision*', *Decision Sciences*. John Wiley & Sons, Ltd, 26(2), pp. 209–227. doi: 10.1111/j.1540-5915.1995.tb01426.x.
45. Lima-Junior, F. R. and Carpinetti, L. C. R. (2019) 'Predicting supply chain performance based on SCOR® metrics and multilayer perceptron neural networks', *International Journal of Production Economics*. Elsevier, 212, pp. 19–38. doi: 10.1016/J.IJPE.2019.02.001.
46. Maestrini, V. et al. (2017) 'Supply chain performance measurement systems: A systematic review and research agenda', *International Journal of Production Economics*. Elsevier, 183, pp. 299–315. doi: 10.1016/J.IJPE.2016.11.005.

47. Makridakis, S. et al. (1982) 'The accuracy of extrapolation (time series) methods: Results of a forecasting competition', *Journal of Forecasting*. John Wiley & Sons, Ltd, 1(2), pp. 111–153. doi: 10.1002/for.3980010202.
48. Manevitz, L., Bitar, A. and Givoli, D. (2005) 'Neural network time series forecasting of finite-element mesh adaptation', *Neurocomputing*, 63, pp. 447–463. doi: 10.1016/j.neucom.2004.06.009.
49. McCulloch, W. S. and Pitts, W. (1943) 'A logical calculus of the ideas immanent in nervous activity', *The bulletin of mathematical biophysics*, 5(4), pp. 115–133. doi: 10.1007/BF02478259.
50. Mentzer, J. et al. (2001) 'Defining Supply Chain Management', *Journal of Business Logistics*, 22. doi: 10.1002/j.2158-1592.2001.tb00001.x.
51. Moosmayer, D. C. et al. (2013) 'A neural network approach to predicting price negotiation outcomes in business-to-business contexts', *Expert Systems with Applications*, 40(8), pp. 3028–3035. doi: <https://doi.org/10.1016/j.eswa.2012.12.018>.
52. Nikolopoulos, K. et al. (2011) 'An Aggregate-Disaggregate Intermittent Demand Approach (ADIDA) to Forecasting: An Empirical Proposition and Analysis', *JORS*, 62, pp. 544–554. doi: 10.1057/jors.2010.32.
53. Patuwo, E., Hu, M. Y. and Hung, M. S. (1993) 'Two-Group Classification Using Neural Networks*', *Decision Sciences*. John Wiley & Sons, Ltd, 24(4), pp. 825–845. doi: 10.1111/j.1540-5915.1993.tb00491.x.
54. Piramuthu, S., Shaw, M. J. and Gentry, J. A. (1994) 'A classification approach using multi-layered neural networks', *Decision Support Systems*, 11(5), pp. 509–525. doi: [https://doi.org/10.1016/0167-9236\(94\)90022-1](https://doi.org/10.1016/0167-9236(94)90022-1).
55. Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986) 'Learning representations by back-propagating errors', *Nature*, 323(6088), pp. 533–536. doi: 10.1038/323533a0.
56. Smith-Miles, K. and Gupta, J. (2000) 'Neural Networks in Business: Techniques and Applications for the Operations Researcher', *Computers and Operations Research*, 27. doi: 10.1016/S0305-0548(99)00141-0.
57. Stock, J. and Watson, M. (1999) 'A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series', in Engle, R. and White, H. (eds) *Cointegration, Causality and Forecasting: A Festschrift for Clive W.J. Granger*. Oxford: Oxford University Press, pp. 1–44. Available at: http://www.economics.harvard.edu/faculty/stock/files/Granger_Festschrift_Stock_andWatson.pdf.
58. Syntetos, A., Kholidasari, I. and Naim, M. (2015) 'The effects of integrating management judgement into OUT levels: In or out of context?', *European Journal of Operational Research*, 249. doi: 10.1016/j.ejor.2015.07.021.
59. Takagi, T. and Sugeno, M. (1983) 'Derivation of Fuzzy Control Rules from Human Operator's Control Actions', *IFAC Proceedings Volumes*, 16(13), pp. 55–60. doi: [https://doi.org/10.1016/S1474-6670\(17\)62005-6](https://doi.org/10.1016/S1474-6670(17)62005-6).
60. Tavana, M. et al. (2016) 'A Hybrid Intelligent Fuzzy Predictive Model with Simulation for Supplier Evaluation and Selection', *Expert Systems with Applications*, 61, pp. 129–144. doi: 10.1016/j.eswa.2016.05.027.

61. Timmermann, A. (2006) 'Chapter 4 Forecast Combinations', Handbook of Economic Forecasting, 1(05), pp. 135–196. doi: 10.1016/S1574-0706(05)01004-9.
62. Tkáč, M. and Verner, R. (2015) 'Artificial neural networks in business: Two decades of research', Applied Soft Computing, 38. doi: 10.1016/j.asoc.2015.09.040.
63. Wei, Y. and Chen, M.-C. (2012) 'Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks', Transportation Research Part C: Emerging Technologies, 21(1), pp. 148–162. doi: <https://doi.org/10.1016/j.trc.2011.06.009>.
64. Zadeh, L. A. (1965) 'Fuzzy sets', Information and Control, 8(3), pp. 338–353. doi: [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X).
65. Zhou, S., Lai, K. K. and Yen, J. (2012) 'A dynamic meta-learning rate-based model for gold market forecasting', Expert Systems with Applications, 39(6), pp. 6168–6173. doi: <https://doi.org/10.1016/j.eswa.2011.11.115>.

About Our Authors

Premporn Khemavuk

She is an Assistant Professor at Industrial Engineering, King Mongkut's University of Technology North Bangkok in Thailand. Her research interests include forecasting, risk management and logistics performance measurement. During her PhD study in Australia, she gained great experiences working for DHL and Linfox. She ensures her researchs always based on both industry and academic point of view.

Athwart Leenatham

He is a PhD candidate at Industrial Engineering, King Mongkut's University of Technology North Bangkok in Thailand. Currently, he is Lecturer at Management and Logistics Engineering at Dhurakij Pundit University in Thailand. His research interests include artificial intelligence, statistical process improvement, operation research, supply-chain management, operation management.