

Integrating fuzzy Delphi method with artificial neural network for demand forecasting of power engineering company

Golam Kabir^{a*} and Razia Sultana Sumi^b

^aDepartment of Civil Engineering, Faculty of Applied Science, The University of British Columbia (UBC), Okanagan Campus, Kelowna, British Columbia, Canada

^bDepartment of Business Administration, Stamford University Bangladesh, Dhaka- 1217, Bangladesh

ARTICLE INFO

Article history:

Received December 15, 2011
Received in Revised form
February, 14, 2012
Accepted 24 March 2012
Available online
April 16 2012

Keywords:

Delphi method
Demand Forecasting
Artificial Neural Network

ABSTRACT

An organization has to make the right decisions in time depending on demand information to enhance the commercial competitive advantage in a constantly fluctuating business environment. Therefore, estimating the demand quantity for the next period most likely appears to be crucial. Manufacturing companies consider forecasting a crucial process for effectively guiding several activities, and research has devoted particular attention to this issue. The objective of the paper is to propose a new forecasting mechanism which is modeled by integrating Fuzzy Delhi Method (FDM) with Artificial Neural Network (ANN) techniques to manage the demand with incomplete information. Artificial neural networks has been applied as it is capable to model complex, nonlinear processes without having to assume the form of the relationship between input and output variables. The effectiveness of the proposed approach to the demand forecasting issue is demonstrated for a 20/25 MVA Distribution Transformer from Energypac Engineering Limited, a leading power engineering company of Bangladesh.

© 2012 Growing Science Ltd. All rights reserved.

1. Introduction

Demand forecast is one of the most important inputs of production planning and supply chain (SC) planning models. In the manufacturing process, demands are predicted to perform the basic planning activities such as capacity planning, resource planning, and raw material purchasing. Demand forecasts are of great importance for marketing activities and personnel management and its accuracy affects directly the profitability of the company. There are many forecasting techniques that can be classified into four main groups: (1) Qualitative methods are primarily subjective; they rely on human judgment and opinion to make a forecast. (2) Time-series methods use historical data to make a forecast. (3) Causal methods involve assuming that the demand forecast is highly correlated with certain factors in the environment (e.g., the state of the economy, interest rate). (4) Simulation methods imitate the consumer choices that give rise to demand to arrive at a forecast (Chopra and Meindl, 2001).

* Corresponding author.

E-mail addresses: gk.raju@yahoo.com (G. Kabir)

Most prior studies have been applied to predict the customer demand primarily based on time-series models, such as moving-average, exponential smoothing, and the Box-Jenkins method, and casual models, such as regression and econometric models. More studies have been done on demand forecasting in the past 20 years. There are numerous models and many different ways to analyze and forecast demand.

Luxhoj et al. (1996) presented a hybrid econometric NN model for forecasting total monthly sales of a Danish company. Garetti and Taisch (1999) presented a selection of the most significant ANN applications to solve production planning and control (PPC) problems. Law and Au (1999) present a new approach that uses a supervised feed-forward neural network model to forecast Japanese tourist arrivals in Hong Kong. Law (2000) extends the applicability of neural networks in tourism demand forecasting by incorporating the back-propagation learning process into a non-linearly separable tourism demand data. Brännäs et al. (2002) obtained an integer-valued moving average model by cross-sectional and temporal aggregation. Smith et al. (2002) examined the theoretical foundation of nonparametric regression and answered the question of whether nonparametric regression based on heuristically improved forecast generation methods approach the single interval traffic flow prediction performance of seasonal ARIMA (Autoregressive Integrated moving average) models.

Weatherford and Kimes (2003) tested a variety of forecasting methods and to determine the most accurate method to forecast the arrivals which is one of the key inputs for a successful hotel revenue management system. Winklhofer and Diamantopoulos (2003) presented and tested a path model of export sales forecasting behavior and performance incorporating organizational and export-specific characteristics. Franses and Dijk (2005) examined the forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production series of 18 Organisation for Economic Co-operation and Development (OECD) countries. Sozen et al. (2005) develop the equations for forecasting net energy consumption (NEC) using ANN technique in order to determine the future level of energy consumption in Turkey.

Garcia-Ferrer et al. (2006) compared the empirical performance of various forecasting models in assessing the effects of policy variables, legal changes, and traffic security campaigns. Petrovic et al. (2006) proposes a demand forecasting model taking into account both statistical forecasting techniques based on historical data as well as expert judgments. Taylor (2007) constructed interval forecasts from quantile predictions generated using exponentially weighted quantile regression. Abdel-Aal (2008) used a univariate modeling of the monthly energy demand time series based only on data for 6 years to forecast the demand for the seventh year. Carbonneau et al. (2008) investigated the applicability of advanced machine learning techniques, including neural networks, recurrent neural networks, and support vector machines, to forecasting distorted demand at the end of a supply chain (bullwhip effect). Heij et al. (2008) proposed an improved method for the construction of principal components in macroeconomic forecasting with the underlying idea which was to maximize the amount of variance of the original predictor variables that is retained by the components in order to reduce the variance involved in estimating the forecast model.

Chu (2009) used univariate autoregressive moving average (ARMA) based models which are applied to tourism demand, by employing both monthly and quarterly time series generated from nine principal tourist destinations in Asian-Pacific region in the forecasting exercise to ensure the reliability of the forecasting evaluation. Sayed et al. (2009) proposes a hybrid forecasting model in which cause and effect based forecasts and time series related forecasts are combined to generate better demand forecast accuracy. Li et al. (2010) presented deterministic vector long-term forecasting (DVL) to support forecasting if there are no matching historical patterns, which is usually the case with long-term forecasting. Herrera et al. (2010) used ANN as one of the predictive models in predicting water consumption in urban areas to design efficient water supply management for the purpose of regular supply of clean water at the pressure required by consumers. Pedregal and Trapero

(2010) developed a general multi-rate methodology in order to forecast optimally load demand series sampled at an hourly rate for a mid-term horizon. Yelland (2010) described a Bayesian statistical model which was developed to forecast the parts demand for a major vendor of enterprise computer products. Andrawis et al. (2011) considered the problem of forecasting monthly tourism numbers for inbound tourism by considering the idea of diversity being accomplished by using different time aggregations. Coshall and Charlesworth (2011) used volatility, exponential smoothing, regression and naive models are considered singly and in combination by means of purely statistical criteria in terms of forecasting demand for international tourism. Chen (2011) combined the linear and nonlinear statistical models to forecast time series with possibly nonlinear characteristics by using real time series data sets of outbound tourism demand to examine the forecasting accuracy of the combination models. This research is an attempt to develop as a successful decision support tool in forecasting customer demands through integrating fuzzy Delphi method with artificial neural network technique for power engineering company.

2. Proposed model

Artificial intelligence forecasting techniques have been receiving much attention lately in order to solve problems that are hardly solved by the use of traditional methods. They have been cited to have the ability to learn like humans, by accumulating knowledge through repetitive learning activities. The proposed approach is aimed to explain a new systematic demand forecasting technique in a fluctuating environment integrating fuzzy Delphi method with artificial neural network which consists of three main phases. The detailed steps of each phase are discussed as follows:

2.1 Fuzzy logic to assign weights to the decision makers

As the Decision Makers (DM) have different experience, designation and qualification, their opinion enjoys different weights in the decision making, so the weights have been assigned to the analysts on this basis. By merging the opinions of almost everybody in the senior management, it is established that the opinion of the decision maker with more experience, higher designation and bigger qualification is more reliable. The linguistic variables for the experience, designation and qualification can be quantified using triangular fuzzy numbers as per Table 1. These linguistic variables can be expressed in positive triangular fuzzy numbers, as in Fig. 1.

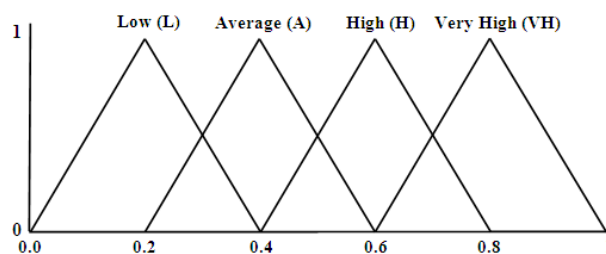


Fig. 1. Linguistic variables for experience, designation and qualification

Table 1

Linguistic variables and FTNs for the experience, designation and qualification

Designation	Qualification	Experience	Linguistic Variables	FTN
Up to manager	Under graduate	0 -<10	Low	0.0,0.2,0.4
Manager to SM	Graduate	10-<20	Average	0.2,0.4,0.6
SM to GM	Specialized graduation	20-<30	High	0.4,0.6,0.8
Sr GM and above	Post graduate	30- above	Very high	0.6,0.8,1.0

2.2 Fuzzy Delphi Method

Fuzzy Delphi Method was proposed by Ishikawa et al. (1993), and it was derived from the traditional Delphi technique and fuzzy set theory. Noorderhoben (1995) indicated that applying the Fuzzy Delphi Method to group decision can solve the fuzziness of common understanding of expert opinions. To shortlist the important criteria for demand forecasting, fuzzy Delphi approach is used in this study. In this method, the unimportant criteria can be identified and eliminated from further consideration. The detailed steps of this preliminary screening phase are described below (Gupta, 2010):

The team of experts from industry (Decision Makers) and academics should determine all possible criteria specific to the industry prior to demand forecasting which may vary dramatically from company to company. Each DM is asked through a questionnaire to specify the importance of the each evaluation criteria. As human judgments are often vague and cannot estimate his preference with an exact numerical number, each analyst must select the appropriate linguistic terms. Its goal is to integrate the opinions of all the DMs to eliminate the unimportant criteria. The seven linguistic terms which can be employed in the questionnaire are as follows: very low, low, medium low, medium, medium high, high, and very high as shown in Figure 2.

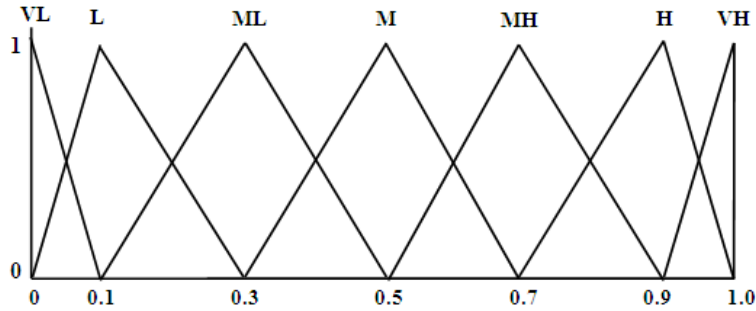


Fig. 2. Linguistic scale for relative importance

The outcome of the questionnaire is the decision matrix as follows:

$$\begin{matrix}
 & \tilde{X}_1 & \tilde{X}_2 & \dots & \tilde{X}_n \\
 & D_1 & D_2 & \dots & D_n \\
 C_1 & \left[\begin{matrix} \tilde{L}_{11} & \tilde{L}_{12} & \dots & \tilde{L}_{1n} \end{matrix} \right] \\
 C_2 & \left[\begin{matrix} \tilde{L}_{21} & \tilde{L}_{22} & \dots & \tilde{L}_{2n} \end{matrix} \right] \\
 \vdots & \left[\begin{matrix} \vdots & \vdots & & \vdots \end{matrix} \right] \\
 C_m & \left[\begin{matrix} \tilde{L}_{m1} & \tilde{L}_{m2} & \dots & \tilde{L}_{mn} \end{matrix} \right]
 \end{matrix}$$

where C_i : the i^{th} evaluation criterion, $i = 1, 2, \dots, m$. D_j : the j^{th} analyst, $j = 1, 2, \dots, n$. \tilde{X}_j : weight of the j^{th} analyst, \tilde{L}_{ij} : the linguistic evaluation of criterion i by the analyst j . Each element \tilde{L}_{ij} in the decision matrix is represented as a triangular fuzzy number $(l_{ij}^a, l_{ij}^b, l_{ij}^c)$.

By using the appropriate fuzzy operators, weighted average of each criteria is calculated as follows:

$$\tilde{W}_i = \frac{\sum_{j=1}^n X_j \otimes L_{ij}}{n}$$

where \tilde{W}_i = weighted average of the i^{th} criteria and $i = 1, 2, \dots, m$. This value is defuzzified using average method by the equation given as:

$$W_i = \frac{W_{ai} + W_{bi} + W_{ci}}{3}$$

The large the number of criteria for the selection process, the more cumbersome and time consuming will be the selection process so only the important criteria are considered for the subsequent evaluation, while the unimportant criteria are eliminated. By integrating the opinions of the all the analysts, a minimum acceptable weight \tilde{R}_δ for all of the criteria are defined which is calculated as:

$$\tilde{R}_\delta = \frac{\sum_{j=1}^n X_j \otimes R_j}{n}$$

where R_j : the minimum acceptable weight for the criteria to be included for evaluation of the service provider defined by j^{th} analyst. This value is defuzzified using average method by the equation given as:

$$R = \frac{R_a + R_b + R_c}{3}$$

A defuzzified value of ' W_i ' is compared with the value of ' R '. The criterion C_i with ' W_i ' less than the value of ' R ' will be eliminated. The remaining criterion will be used in the final selection phase. This way Delphi assists the analysts to identify the important evaluation criteria and to obtain the weights of the criteria for the provider selection (Gupta, 2010).

2.3 Artificial Neural Network (ANN)

An artificial neural network is a mathematical model or computational model based on biological neural networks. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. A neural network is a system composed of many simple processing elements operating in parallel whose function is determined by network structure, connection strengths, and the processing performed at computing element or nodes.

Neural network architecture is inspired by the architecture of biological nervous systems, which use many simple processing elements operating in parallel to obtain high computation rates. The neural networks resemble the brain mainly in two respects-knowledge is acquired by the network from its environment and interneuron connection strengths, known as synaptic weights are used to store the acquired knowledge (Haykin, 2001). The basic element in an ANN is a neuron. The model of a neuron is depicted in Fig. 3 (Haykin, 2001).

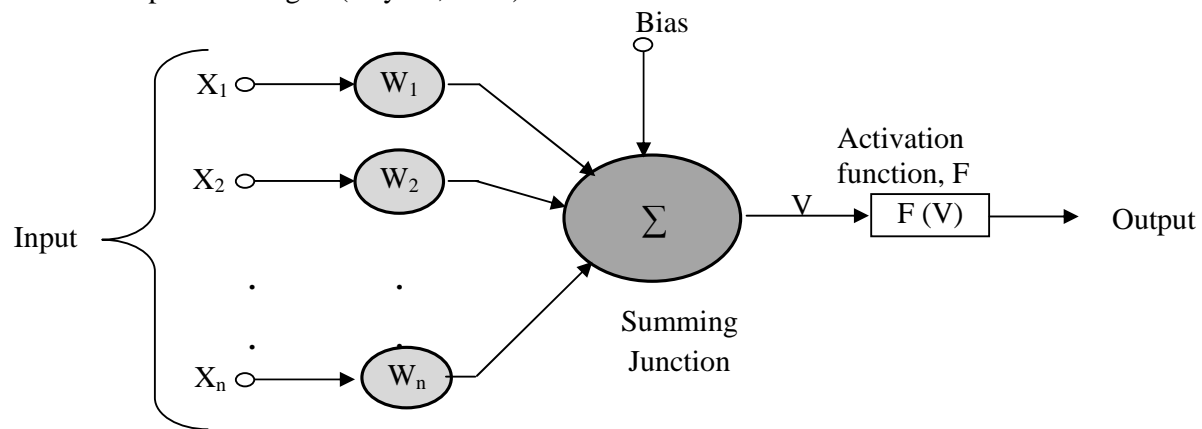


Fig. 3. A Single ANN Neuron with its Elements

In a neural network model, simple nodes (called neurons) are connected together to form a network of nodes hence the term neural network. Other names for the field include connectionism, parallel

distributed processing, neural computation or computational neural network (CNN). Neural networks are composed of nodes or units connected by directed links. Each link has a numeric weight associated with it which determines the strength and sign of the connection. Figure 4 shows ANN with input, output and one hidden layer with four neurons.

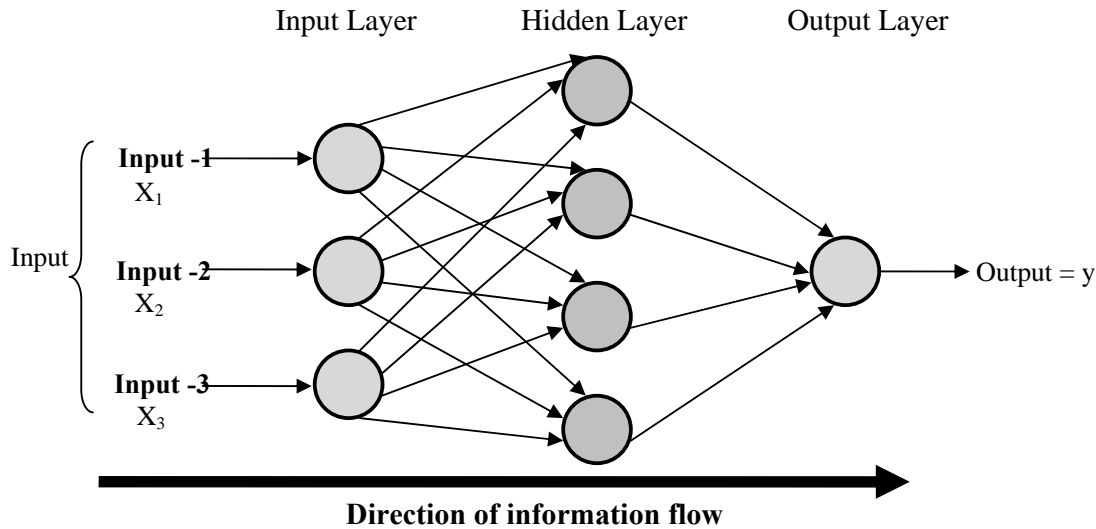


Fig.4. A Feed Forward ANN with Input, Output and One Hidden Layer

Here there are three inputs which are X_1 , X_2 , and X_3 and they produce a single output Y . The hidden layer establishes the complex relationship between the inputs and the outputs. The number of hidden neuron and hidden layer varies depending on the nature and complexity of the problems and distributions of the data set. An ANN may be seen as a *black box* which contains hierarchical sets of neurons (e.g., processing elements) producing outputs for certain inputs. Each processing element consists of data collection, processing the data and sending the results to the relevant consequent element. The whole process may be viewed in terms of the inputs, weights, the summation function, and the activation function (Figure 4) (Palau et al., 1999). The description of the Figure 4 and the working principle is given below:

- [1] The inputs (e.g. X_1, X_2, \dots, X_n) are the activity of collecting data from the relevant sources. These data are fed to the neural network.
- [2] The weights control the effects of the inputs on the neuron. In other words, an ANN saves its information over its links and each link has a weight (e.g. W_1, W_2, \dots, W_n). These weights are constantly varied while trying to optimize the relation in between the inputs and outputs. Synaptic weights characterize themselves with their strength (value) which corresponds to the importance of the information coming from each neuron. In other words, the information is encoded in these strength-weights.
- [3] Summation function is to calculate of the net input readings from the processing elements. (E.g. $\text{Sum} = W_1 X_1 + W_2 X_2 + W_3 X_3 + \dots + W_n X_n$)
- [4] Transfer (activation) function (f) determines the output of the neuron by accepting the net input ($\text{Sum} = W_1 X_1 + W_2 X_2 + W_3 X_3 + \dots + W_n X_n$) provided by the summation function.

The learning method can be divided into two categories, namely, unsupervised learning and supervised learning. A backpropagation supervised learning model is designed in this study. The error between the expected output and the calculated output is computed. Then a minimization procedure is used to adjust the weights between two connection layers starting backwards from the output layer to input layer. There are a number of variations of minimization procedures that are based on different optimization methods, such as gradient descent, conjugate gradient, Quasi-Newton, and Levenberg-

Marquardt methods. The differences between them are based on various weight adjustments (Haykin, 2001; Manevitz et al., 2005). In this research, Levenberg-Marquardt training algorithm has been used.

A practical problem with NN is the selection of the correct complexity of the model, i.e., the correct number of hidden units or correct regularization parameters. The design of hidden layer is dependent on the selected learning algorithm. Supervised learning systems are generally more flexible in the design of hidden layers. A greater quantity of hidden layers enables a NN model to improve its closeness-of-fit, while a smaller quantity improves its smoothness or extrapolation capabilities. It was concluded that the number of hidden layers is heuristically set by determining the number of intermediate steps to translate the input variables into an output value (Choy et al., 2003). According to some literature studies, the number of hidden layer nodes can be up to (1) $2n + 1$ (where n is the number of nodes in the input layer), (2) 75% of the quantity of input nodes, or (3) 50% of the quantity of input and output nodes (Lenard et al., 1995; Piramuthu et al., 1994).

A transfer function is needed to introduce the non-linearity characteristics into the network. The non-linear function will make the hidden units of multi-layer network more powerful than just plain perception. The used transfer function is a standard function for backpropagation, that is, the sigmoid transfer function. The sigmoid transfer function is chosen due to its ability to help the generalization of learning characteristics to yield models with improved accuracy (Choy et al., 2003).

The backpropagation training paradigm uses three controllable factors that affect the algorithm's rate of learning. They are the learning rate coefficient, momentum and the exit conditions. Learning coefficient governs the speed that the weights can be changed over time, reducing the possibility of any weight oscillation during the training cycle. Momentum parameter controls over how much iteration an error adjustment persists. There is no definitive rule regarding the momentum, in general it is set to 0.5 which is half of the maximum limit for training to reduce the damping effect. NNs use a number of different stopping rules to control the termination of the training process (Choy et al., 2003).

3. Application of the model

In order to prove the applicability and validity of the proposed approach, it is demonstrated for Energypac Engineering Limited (EEL), a leading power engineering company of Bangladesh. Energypac Engineering Ltd. is the manufacturer of different types of Transformers and Switchgears. Energypac Engineering Ltd. is the manufacturer of Transformer (Power Transformer, Distribution Transformer and Instrumental Transformer) and Switchgear (Outdoor vacuum circuit breaker, Indoor vacuum circuit breaker, Control, Metering and Relay panels, Low Tension and Power Factor Improvement panel, Indoor type Load Break Switch, Outdoor Offload disconnecter and By-pass switch). This study was performed for a 20/25 MVA Distribution Transformer. In the following section, the detailed demand forecasting process for the company is described.

3.1 Determination of criteria

As the DMs have different experience, designation and qualification, their opinion enjoys different weights in the decision making. Four analysts who hold the right to make the final decision (two from marketing, one from technical and corporate departments and further to be referred as DM 1, DM 2, DM 3 and DM 4 respectively) from the related industry are chosen to form the decision team. Based on their experience, designation and academic qualification, using Table 1 the weights of the decision makers are DM 1 (X_1) = (0.47,0.67,0.87), DM 2 (X_2) = (0.27,0.47,0.67), DM 3 (X_3) = (0.47,0.67,0.87) and DM 4 (X_4) = (0.07,0.27,0.47).

The decision team agreed to adopt the 9 criteria for demand forecasting as the initial evaluation criteria used for the fuzzy Delphi process. Those are Unit sales price (USP), Product quality (PQ), Customer satisfaction level (CSL), Versatility and lead time (VLT), Payment flexibility (PF), Effect of seasonality and promotions (ESP), Bureaucracy advantage (BA), Effect of reputation (ER) and Effect of competitors (EC). Each DM is asked through a questionnaire to specify the importance of the each evaluation criteria. Table 2 shows the judgments of the DM's and the aggregated values of the selected 9 criteria for initial evaluation.

Table 2
Weighted aggregate of criteria

Sr. No	Criteria	DM 1	DM 2	DM 3	DM 4	Weighted Aggregated of each Criteria	Defuz-ified Value	Select or Reject
		Weightage of the decision makers (X_n)						
		0.47,0.67,0.87	0.27,0.47,0.67	0.47,0.67,0.87	0.07,0.27,0.47			
1	USP	H	VH	VH	VH	0.264,0.503, 0.72	0.496	S
		0.7,0.9,1	0.9,1,1	0.9,1,1	0.9,1,1			
2	PQ	H	VH	VH	H	0.261,0.496, 0.72	0.492	S
		0.7,0.9,1	0.9,1,1	0.9,1,1	0.7,0.9,1			
3	CSL	H	H	MH	H	0.200, 0.434, 0.698	0.444	S
		0.7,0.9,1	0.7,0.9,1	0.5,0.7,0.9	0.7,0.9,1			
4	VLT	MH	MH	M	MH	0.136,0.330, 0.605,	0.357	R
		0.5,0.7,0.9	0.5,0.7,0.9	0.3,0.5,0.7	0.5,0.7,0.9			
5	PF	M	MH	MH	M	0.133,0.317, 0.581	0.344	R
		0.3,0.5,0.7	0.5,0.7,0.9	0.5,0.7,0.9	0.3,0.5,0.7			
6	ESP	H	H	H	H	0.224,0.468, 0.72	0.470	S
		0.7,0.9,1	0.7,0.9,1	0.7,0.9,1	0.7,0.9,1			
7	ER	MH	ML	M	ML	0.102, 0.256, 0.490	0.283	R
		0.5,0.7,0.9	0.1,0.3,0.5	0.3,0.5,0.7	0.1,0.3,0.5			
8	BA	M	M	MH	H	0.126, 0.320, 0.582	0.343	R
		0.3,0.5,0.7	0.3,0.5,0.7	0.5,0.7,0.9	0.7,0.9,1			
9	EC	H	MH	H	MH	0.207, 0.431, 0.691	0.443	S
		0.7,0.9,1	0.5,0.7,0.9	0.7,0.9,1	0.5,0.7,0.9			

Eliminate unimportant criteria. It was decided to select all the criteria whose weight are more than 0.40 and eliminate the rest. The selected main criteria are Unit sales price (USP), Product quality (PQ), Customer satisfaction level (CSL), Effect of seasonality & promotions (ESP) and Effect of Competitors (EC) (Table 2). It can be mentioned that more criteria can be selected for final evaluation by reducing the minimum acceptable weight \tilde{R}_ϕ for all of the criteria. The demand forecasting system for this study is built by the factors in the following:

Unit sales price (input): Unit sales price is a competitive factor affecting the customer behaviors. It is processed as quantitative information.

Product quality (input): This factor includes the evaluation about product quality according to manufacturer and customers via a 1-9 scale. It is processed as qualitative information.

Customer satisfaction level (input): This factor shows the sales and post-sales behaviors of manufacturer to the customers. It is processed as quantitative information.

Effect of seasonality and promotions (input): This factor means the percent increase of sales related to seasonality and promotions. It is processed as quantitative information. Effect of Competitors (input): This factor shows the competition (no. of company manufacturing the same product) in the market. It is processed as qualitative information.

3.2 Data Collection

Unit sales price, product quality, customer satisfaction level, effect of seasonality & promotions and effect of competitors for the 42 monthly periods were used for the proposed neural network forecasting mechanism. Demand curve for the last 42 monthly periods is given in Figure 5 which shows that there is some periodically pattern such as trend and seasonality.

Although the time-series methods perform well in this situation, they suffer from some limitations. First, lack of expertise might cause a mis-specification of the functional form linking the independent and dependent variables together, resulting in a poor regression. Secondly, a large amount of data is often required to guarantee an accurate prediction. Thirdly, non-linear patterns are difficult to capture. Finally, outliers can bias the estimation of the model parameters. Some of these limitations can be overcome by the use of artificial intelligence approaches, which have been mathematically demonstrated to be universal approximates of functions (Garetti & Taisch, 1999). The main characteristic of a neural network is the ability to learn from environment, and to improve its performance through learning (Haykin, 2001; Skapura, 1996). ANN have three great advantages over traditional methods such as; universal approximation capabilities, recognize “on their own” implicit dependencies and relationship in data, learn” to adapt their behavior (Bodyanskiy and Popov, 2006). That is why ANN has been used for this research to forecast the demand of the transformer.

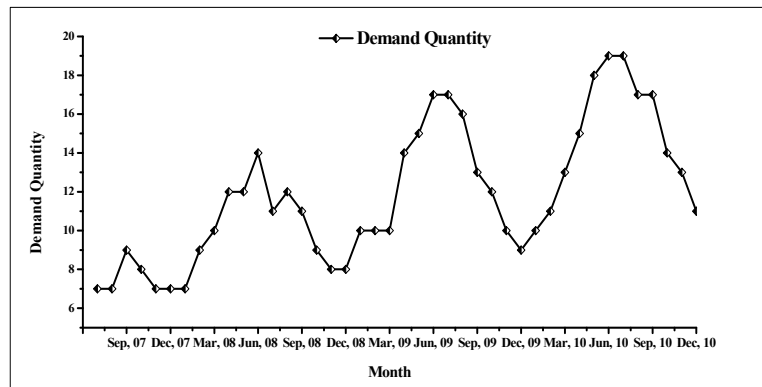


Fig. 5. Demand for 42 Monthly Periods

3.3 Demand forecasting using artificial neural network

The input/output dataset of the ANN model that is going to be formulated to forecast demand is illustrated schematically in Figure 6. The five basic steps used in general application of neural network have been adopted in the development of the model: assembly or collection of data; analysis and pre-processing of the data; design of the network object; training and testing of the network; and performing simulation with the trained network and post-processing of results.

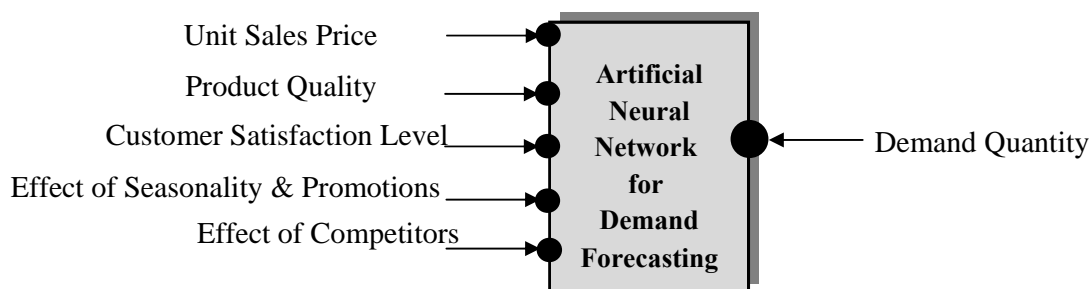


Fig. 6. Schematic Diagram of ANN for Demand Forecasting

The neural network has been designed with MATLAB 7.6 software. The feed forward backpropagation network consists of input layer with five neurons corresponding to each of the five input variables and one neuron in the output layer. The algorithm used for the neural network learning is ‘the backward propagation algorithm’ with Levenberg-Marquardt (LM) version. The proposed methodology was trained with 32 samples corresponding to 75% of the data set and tested with 10 samples (approximately 25% of the data set) that were selected in random from the data set. The momentum constant and learning rate used in this model is 0.5 and 0.5 respectively. For the optimal network architecture, logarithmic transfer function ‘logsig’, tangent sigmoid transfer function ‘tansig’, linear transfer function ‘purelin’ has been used in the output layer with different number of hidden neurons. The maximum number of training epochs set was 10,000 and the training error goal was 0.0001. The performance of the network was evaluated by mean absolute percentage of error (MAPE) between the measured and the predicted values for every output nodes in respect of training the network. The different combinations of this experimental study of demand forecasting are summarized in Table 3. To find out the optimal model, 56 different neural network architecture models have been constructed (Table 4).

Table 3
Summary of the ANN Model for Demand Forecasting

	Multi layer feed-forward back-propagation neural networks	
	Input Neurons: Unit sales price, Product quality, Customer satisfaction level, Effect of seasonality & promotions, Effect of Competitors	
	Output Neurons: Demand Quantity	
	Hidden Layer: 1 or 2	
	Hidden Neurons: 5, 7, 8, 10, 12, 15, 20	
Architecture	Learning Algorithm: Levenberg- Marquardt Optimization	
	Transfer Function:	Input layer logarithmic sigmoid (logsig), tangent sigmoid (tansig) Output layer logarithmic sigmoid (logsig), tangent sigmoid (tansig), linear transfer function (purelin)
	Sample pattern vector: 32 (for training), 10 (for testing)	
	Learning rate/coefficient: 0.5	
	Momentum rate: 0.5	
	All input data normalized between 0 and 1	
Computation/ Termination	Training: Feed-forward back-propagation Training Termination: Stop training when reaches 10,000 epochs or 0.001 error tolerance	

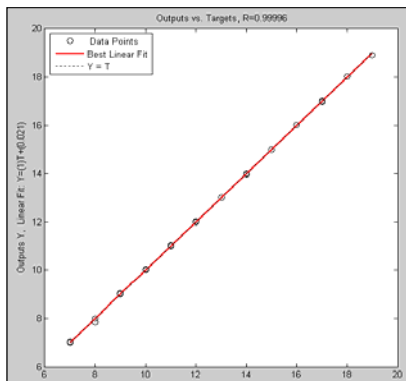


Fig. 7 (a) Training dataset

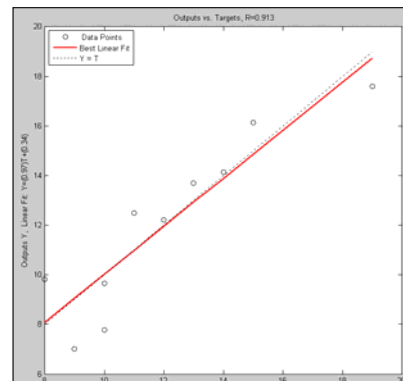


Fig. 7 (b) Testing dataset

Fig. 7. Correlation between ANN Prediction and Actual Values of Demand Forecasting (5-20-8-1 ANN Architecture)

Table 4
Neural Network Architecture for Multicriteria Demand Forecasting

Learning Rate	No. of Hidden Layer (HF)	No. of Hidden Neurons		Activation Function (AF)			Train/Test Error			
		HL 1	HL 2	HL 1	HL 2	Output Layer	Test R ²	Test MAPE	Train R ²	Train MAPE
0.5	1	5		tansig		purelin	0.0137	10.858	0.9998	0.0504
0.5	1	7		tansig		purelin	0.8745	1.9475	1.0000	0.0043
0.5	1	8		tansig		purelin	0.4794	5.5757	1.0000	0.0071
0.5	1	10		tansig		purelin	0.5300	15.824	0.9999	0.2953
0.5	1	12		tansig		purelin	0.7836	3.2526	1.0000	0.0055
0.5	1	15		tansig		purelin	0.7314	3.5502	0.9998	0.0263
0.5	1	20		tansig		purelin	0.9193	0.7332	0.9998	0.1268
0.5	1	5		logsig		purelin	0.7755	0.3681	0.9981	0.0156
0.5	1	7		logsig		purelin	0.9015	5.6087	1.0000	0.0008
0.5	1	8		logsig		purelin	0.9047	7.3209	0.9999	0.0700
0.5	1	10		logsig		purelin	0.6040	8.9115	0.9999	0.0444
0.5	1	12		logsig		purelin	0.5486	8.2826	0.9998	0.2227
0.5	1	15		logsig		purelin	0.8455	2.0449	1.0000	0.0066
0.5	1	20		logsig		purelin	0.8100	5.5384	0.9998	0.1153
0.5	1	5		logsig		tansig	0.0879	2.9701	0.9998	0.0344
0.5	1	7		logsig		tansig	0.5358	6.7603	0.9998	0.1402
0.5	1	8		logsig		tansig	0.7151	0.9664	0.9997	0.0100
0.5	1	10		logsig		tansig	0.7369	0.0130	0.9998	0.1155
0.5	1	12		logsig		tansig	0.8355	4.2032	0.9999	0.0196
0.5	1	15		logsig		tansig	0.6400	8.9895	0.9999	0.0089
0.5	1	20		logsig		tansig	0.8151	8.0000	0.2298	34.025
0.5	1	5		tansig		logsig	0.8272	13.281	0.6983	11.741
0.5	1	7		tansig		logsig	0.6857	16.309	0.6983	11.740
0.5	1	8		tansig		logsig	0.6551	10.977	0.6983	11.740
0.5	1	10		tansig		logsig	0.6552	10.973	0.7216	12.002
0.5	1	12		tansig		logsig	0.7774	12.385	0.7216	12.002
0.5	1	15		tansig		logsig	0.6558	10.984	0.6983	11.741
0.5	1	20		tansig		logsig	0.6551	9.1153	0.2898	6.1694
0.5	2	20	20	tansig	tansig	purelin	0.9251	1.8821	0.9999	0.1306
0.5	2	20	5	tansig	logsig	purelin	0.8754	4.9648	1.0000	0.0050
0.5	2	5	20	logsig	tansig	purelin	0.5677	9.3519	0.9999	0.1927
0.5	2	5	5	logsig	logsig	purelin	0.7766	0.3682	0.3682	0.0004
0.5	2	20	20	tansig	tansig	tansig	0.8155	5.4365	0.9999	0.0963
0.5	2	20	5	tansig	logsig	tansig	0.7854	1.3095	0.9998	0.1969
0.5	2	5	20	logsig	tansig	tansig	0.6952	3.0109	0.9999	0.0661
0.5	2	5	5	logsig	logsig	tansig	0.6956	0.3762	0.9997	0.0280
0.5	2	20	20	tansig	tansig	logsig	0.0471	0.3721	0.0847	0.8600
0.5	2	20	5	tansig	logsig	logsig	0.6764	11.203	0.7216	12.002
0.5	2	5	20	logsig	tansig	logsig	0.0443	0.3721	0.0339	0.8600
0.5	2	5	5	logsig	logsig	logsig	0.1138	0.3721	0.0637	0.8600
0.5	2	15	15	tansig	tansig	purelin	0.6893	1.2055	1.0000	0.0462
0.5	2	20	10	tansig	logsig	purelin	0.8040	0.1798	1.0000	0.0669
0.5	2	7	15	logsig	tansig	purelin	0.8450	0.8123	1.0000	0.0184
0.5	2	10	10	logsig	logsig	purelin	0.7571	3.2000	1.0000	0.0204
0.5	2	5	5	tansig	tansig	tansig	0.7686	2.4464	0.9999	0.2405
0.5	2	20	10	tansig	logsig	tansig	0.8664	4.3409	0.9998	0.0779
0.5	2	10	20	logsig	tansig	tansig	0.4160	12.458	0.9997	0.0075
0.5	2	20	20	logsig	logsig	tansig	0.8175	2.9268	0.9999	0.1650
0.5	2	20	12	tansig	tansig	logsig	0.8796	17.188	0.7216	12.002
0.5	2	15	20	tansig	logsig	logsig	0	0.3721	0	0.8600
0.5	2	7	15	logsig	tansig	logsig	0.8164	0.3721	0.1599	0.8600
0.5	2	10	10	logsig	logsig	logsig	0.0471	0.3721	0.0847	0.8600
0.5	2	20	8	tansig	logsig	tansig	0.8336	0.5557	0.9999	0.0182
0.5	2	20	8	logsig	logsig	purelin	0.7900	0.5749	1.000	0.0598
0.5	2	7	8	logsig	tansig	purelin	0.8372	1.5770	1.000	0.0061
0.5	2	8	15	tansig	logsig	purelin	0.7111	0.3734	1.000	0.0838

It is shown from the Table 4 that network with 2 hidden layer and 20 neurons in the first hidden layer with ‘tansigmoid’ and 8 neurons in the second hidden layer with ‘logsigmoid’ and ‘purelin’ transfer function in the output layer respectively provides the best result. So, 5-20-8-1 network architecture was selected as the optimum ANN model. Figure 7 shows that the correlation coefficient of 0.99996 and 0.913 was obtained for training dataset and the model predictions respectively. It is shown that the predicted error for the training data is 0.0182 and for the test data it is 0.5557. Figure 8 shows the ANN prediction values and actual values for 5-20-8-1 ANN architecture. From the graphs, it is clear that the proposed model can predict values which are nearly very close to experimental observations for each of the output parameters.

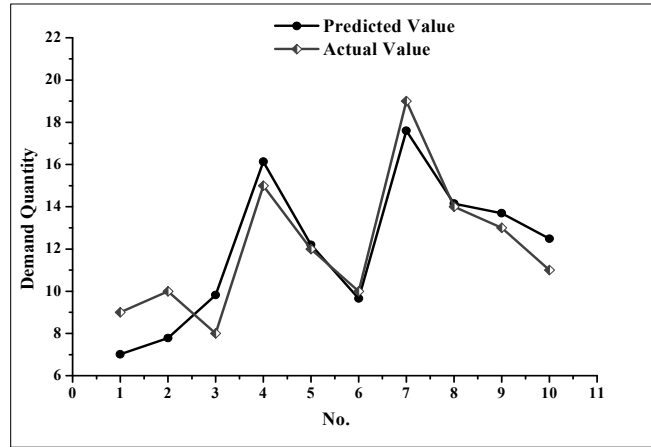


Fig. 8. Actual vs. ANN Prediction of Demand Forecasting for 10 Testing Cases (5-20-8-1 ANN Architecture)

4. Discussions

In this research, an artificial neural network (ANN) with feed-forward back-propagation algorithm was trained with Levenberg-Marquardt algorithm (LM) and the training epoch (cycles) set for each network is 10,000. The proposed methodology was trained with 32 samples corresponding to 75% of the data set and tested with 10 samples (approximately 25% of the data set) that were selected in random from the data set. The purpose of the training is to minimize the mean absolute percentage of error (MAPE). Finally, network with 2 hidden layer and 20 neurons in the first hidden layer with ‘tansigmoid’ and 8 neurons in the second hidden layer with ‘logsigmoid’ and ‘purelin’ transfer function in the output layer respectively was found out as the optimal network. The performance of the 5-20-8-1 ANN model has been highlighted in Figure 8. As shown in the figures, it is clear that the values predicted by ANN are very close to experimental values. The co-efficient of determination (R^2) was obtained 0.99996 and 0.913 for training and testing dataset respectively. The mean absolute percentage of error (MAPE) between the actual and the predicted values were 0.0182 and 0.5557 for the test data.

5. Conclusions

Demand forecast is one of the most important inputs of production planning and supply chain (SC) planning models. Demand forecasts are of great importance for marketing activities and personnel management and its accuracy affects directly the profitability of the company. This study has developed a forecasting mechanism based on fuzzy Delphi method with ANN techniques to manage the demand forecasting issue under incomplete information. Multilayer feed forward backpropagation network consisting of five inputs, 2 hidden layer and 20 neurons in the first hidden layer with ‘tansigmoid’ and 8 neurons in the second hidden layer found to be the optimum network (5-20-8-1) for the model developed in this study. This model can also be utilized to identify the effectual and influencing variables for demand forecasting. To demonstrate the effectiveness of the proposed

methodology, demand forecasting issue was investigated for Energypac Engineering Limited (EEL), a Bangladeshi power engineering company as an application of the models.

Forty two months data were used for demand forecasting using artificial intelligence approaches. If the number of input/output dataset is increased, the accuracy of the model will increase as well as the error of the model will decrease. Therefore, Monte Carlo simulation can be used to generate data considering the characteristics of samples. The work can also be analyzed by recurrent neural network or radial basis functional neural network and then compared the prediction accuracy or generalization capability of different types of neural network.

References

- Abdel-Aal, R.E. (2008). Univariate modeling and forecasting of monthly energy demand time series using abductive and neural networks. *Computers & Industrial Engineering*, 54(4), 903-917.
- Aburto, L., & Weber, R. (2007). Improved supply chain management based on hybrid demand forecasts. *Applied Soft Computing*, 7(1), 136-144.
- Andrawis, R.R., Atiya, A.F., & El-Shishiny, H. (2011) Combination of long term and short term forecasts, with application to tourism demand forecasting. *International Journal of Forecasting*, 27, 870–886.
- Bodyanskiy, T., & Popov, S. (2006). Neural Network approach to forecasting of quasiperiodic financial time series. *European Journal of operation Research*, 175(3), 1375-1366.
- Brännäs, K., Hellström, J., & Nordström, J. (2002). A new approach to modeling and forecasting monthly guest nights in hotels. *International Journal of Forecasting*, 18(1), 19-30.
- Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research*, 184(3), 1140-1154.
- Chen, K.Y. (2011). Combining linear and nonlinear model in forecasting tourism demand. *Expert Systems with Applications*, 38(8), 10368-10376.
- Chopra, S., & Meindl, P. (2010). *Supply chain management: Strategy, planning and operation*. 4th Ed., NJ: Prentice-Hall.
- Choy, K. L., Lee, W. B., & Lo, V. (2003). Design of an intelligent supplier relationship management system: A hybrid case based neural network approach. *Expert Systems with Applications*, 24(2), 225–237.
- Chu, F.L. (2009). Forecasting tourism demand with ARMA-based methods. *Tourism Management*, 30(5), 740-751.
- Coshall, J.T., & Charlesworth, R. (2011). A management orientated approach to combination forecasting of tourism demand. *Tourism Management*, 32(4), 759-769.
- Franses, P.H., & Dijk, D.V. (2005). The forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production. *International Journal of Forecasting*, 21(1), 87-102.
- Garcia-Ferrer, A., Juan, A.D., & Poncela, P. (2006). Forecasting traffic accidents using disaggregated data. *International Journal of Forecasting*, 22(2), 203-222.
- Garetti, M., & Taisch, M. (1999). Neural networks in production planning and control. *Production Planning & Control: The Management of Operations*, 10(4), 324-339.
- Gupta, R., Sachdeva, A., & Bhardwaj, A. (2010). Selection of 3pl Service Provider using Integrated Fuzzy Delphi and Fuzzy TOPSIS. *Proceedings of the World Congress on Engineering and Computer Science*, Vol II, October 20-22, 2010, San Francisco, USA.
- Haykin, S. (2001). *Neural Networks- A comprehensive Foundation*, 2nd edition, Pearson Education, Inc. Singapore.
- Heij, C., Dijk, D.V., & Groenen, P.J.F. (2008). Macroeconomic forecasting with matched principal components. *International Journal of Forecasting*, 24(1), 87-100.

- Herrera, M., Torgo, L., Izquierdo, J., & García, R.P. (2010). Predictive models for forecasting hourly urban water demand. *Journal of Hydrology*, 387(1-2), 141-150.
- Ishikawa, A., Amagasa, M., Shiga, T., Tomizawa, G., Tatsuta, R., & Mieno, H. (1993). The max-min Delphi method and fuzzy Delphi method via fuzzy integration. *Fuzzy Sets and Systems*, 55(3), 241-253.
- Law, R. (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. *Tourism Management*, 21(4), 331-340.
- Law, R., & Au, N. (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. *Tourism Management*, 20(1), 89-97.
- Lenard, M. J., Alam, P., & Madey, G.R. (1995). The application of neural networks and a qualitative response model to the auditor's going concern uncertainty decision. *Decision Sciences*, 26(2), 209-227.
- Li, S.T., Kuo, S.C., Cheng, Y.C., & Chen, C.C. (2010). Deterministic vector long-term forecasting for fuzzy time series. *Fuzzy Sets and Systems*, 161(13), 1852-1870.
- Luxhoj, J.T., Riis, J.O., & Stensballe, B. (1996). A hybrid econometric-neural network modeling approach for sales forecasting. *International Journal of Production Economics*, 43(2-3), 175-192.
- Manevitz, L., Bitar, A., & Givoli, D. (2005). Neural network time series forecasting of finite-element mesh adaptation. *Neurocomputing*, 63, 447-463.
- Noorderhaben, N. (1995). *Strategic decision making*. UK: Addison-Wesley.
- Palau, A., Velo, E., & Puigjaner, L. (1999). Use of Neural Networks and Expert Systems to Control A Gas/Solid Sorption Chilling Machine. *International Journal of Refrigeration*, 22(1), 59-66.
- Pedregal, D.J., & Trapero, J.R. (2010). Mid-term hourly electricity forecasting based on a multi-rate approach. *Energy Conversion and Management*, 51(1), 105-111.
- Petrovic, D., Xie, Y., & Burnham, K. (2006). Fuzzy Decision Support System for Demand Forecasting with a Learning Mechanism. *Fuzzy Sets and Systems*, 157(12), 1713-1725.
- Piramuthu, S., Shaw, M., & Gentry, J. (1994). A classification approach using multilayered neural networks. *Decision Support Systems*, 11(5), 509-525.
- Sayed, H.E., Gabbar, H.A., & Miyazaki, S. (2009). A Hybrid Statistical Genetic-based Demand Forecasting Expert System. *Expert Systems with Applications*, 36(9), 11662- 11670.
- Skapura, D.M. (1996). *Building Neural Networks*. ACM Press, New York, 143-148.
- Smith, B.L., Williams, B.M., & Oswald, R.K. (2002). Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research Part C: Emerging Technologies*, 10(4), 303-321.
- Sozen, A., Arcaklioglu, E., & Ozkaymak, M. (2005) Turkey's net energy consumption. *Applied Energy*, 81(2), 209-221.
- Taylor, J.W. (2007). Forecasting daily supermarket sales using exponentially weighted quantile regression. *European Journal of Operational Research*, 178(1), 154-167.
- Weatherford, L.R., & Kimes, S.E. (2003). A comparison of forecasting methods for hotel revenue Management. *International Journal of Forecasting*, 19(3), 401-415.
- Winklhofer, H., & Diamantopoulos, A. (2003). A model of export sales forecasting behavior and performance: development and testing. *International Journal of Forecasting*, 19(2), 271-285.
- Yelland, Y.M. (2010). Bayesian forecasting of parts demand. *International Journal of Forecasting*, 26(2), 374-396.