World Academy of Science, Engineering and Technology
International Journal of Energy and Power Engineering
Vol:5, No:10, 2011

Development of Neural Network Prediction Model of Energy Consumption

Maryam Jamela Ismail, Rosdiazli Ibrahim, Idris Ismail

***Abstract*—**In the oil and gas industry, energy prediction can help the distributor and customer to forecast the outgoing and incoming gas through the pipeline. It will also help to eliminate any uncertainties in gas metering for billing purposes. The objective of this paper is to develop Neural Network Model for energy consumption and analyze the performance model. This paper provides a comprehensive review on published research on the energy consumption prediction which focuses on structures and the parameters used in developing Neural Network models. This paper is then focused on the parameter selection of the neural network prediction model development for energy consumption and analysis on the result. The most reliable model that gives the most accurate result is proposed for the prediction. The result shows that the proposed neural network energy prediction model is able to demonstrate an adequate performance with least Root Mean Square Error.

***Keywords*—**Energy Prediction, Multilayer Feedforward, Levenberg-Marquardt, Root Mean Square Error (RMSE)

I. Introduction

IN Artificial Neural Network (ANN), some fundamentals of neural network concepts in developing a neural network model which will determine the reliability and robustness of the system is the most vital part. A lot of research have been done in establishing appropriate guidelines in choosing the network architectures, determining the pre-processing data, learning algorithm, and other performing criteria. These parameters selections will actually affect the neural network presentation and improve the performance. It is essential to adopt a systematic approach in the development of ANN models, taking into account factors such as data pre-processing, the determination of adequate model inputs and a suitable network architecture, parameter estimation and model validation [1]. The main objective of this paper is to develop Neural Network Model for energy consumption. This paper discusses every aspect of ANN model development such as training data collection, data pre- and post-processing, different types of activation functions, training algorithms in finding the best architecture and performance. This is justified not only by the fact that it is directly associated with the model's performance but also because there is no theoretical background as to how this architecture will be found or what it should look like [2]. The selection of parameter in the Neural Network model will definitely provide some implication towards the performance of the model prediction. As an example the researchers has investigated the effect of changing the number of hidden layers of the MLPs and the number of processing elements that exist in the hidden layers of the analyzed properties of Jordan Oil Shale. The outcome of the study shows that by changing the number of hidden layers, the number of processing elements in the hidden layers will therefore be affected [3].

Rosdiazli Ibrahim is with the Universiti Teknologi PETRONAS Bandar Seri Iskandar, 31720 Tronoh, Perak, Malaysia (phone: 605-368-7821; fax: 605-365-7443; e-mail: rosdiazli@petronas.com.my).

Idris Ismail with the Universiti Teknologi PETRONAS Bandar Seri Iskandar, 31720 Tronoh, Perak, Malaysia (phone: 605 368 838; fax: 605-3688386; e-mail: idrisim@petronas.com.my).

Maryam Jamela Ismail is with the Universiti Teknologi PETRONAS Bandar Seri Iskandar, 31720 Tronoh, Perak, Malaysia (e-mail:maryamjamela@yahoo.com)

II. Neural network in energy prediction

Energy prediction has always been one of the major research fields, as that an accurate prediction and forecast is essential to produce a reliable energy distribution system. In this case, the energy can be the natural gas, electricity load, wind and solar energy and other energy systems. However, for the past decades, the research on forecasting is done more for electric load problem. A well known analogy between electricity and natural gas consumption allowed using references related to electric load forecasting problem [4].

Accurate prediction and forecasting of natural gas consumption for specific distributive area is of great importance for economical and reliable operation of distributive network [4]. Customers are billed according to the amount of energy calculated from the natural gas composition and consumption. A slight error in calculation will lead to significant monetary impact. On the view point of distributors, with accurate forecasting the number of false alarms would be significantly decreased and transship limits would be scheduled. On the view point of consumers, there will be no disconnection or breakdown. This would make the gas system more reliable and profitable [5].

Artificial Neural Network (ANN) recently is one of many computational methods that has great attention in research and vastly used in prediction application. It is a well-known fact that ANN can model any nonlinear relationship to an arbitrary degree of accuracy by adjusting the network parameters [6]. It can also handle nonlinearities among variables as the expected nature of the energy consumption data is nonlinear. 16 papers on gas consumption prediction using neural network approach is studied and compared. The analysis on the model developed is shown in Table I.

TABLE I
PAPERS REVIEWED ON THE PARAMETER SELECTION IN CONSTRUCTING THE NEURAL NETWORK MODEL

No	Architecture	Time Step	Variable	Normalized Data	No. of Training Data	No. of Validation Data	Structure	Transfer Function	Learning Algorithm	Ref
1	MLP	H	Gas	N/A	N/A	N/A	10:20:1	Sig/Lin	BP	[4]
2	MLP	H	Gas	[0.05-0.95]	N/A	N/A	43:10:24	Sig/Sig	BP	[7]
3	Adaptive –MLP Functional NN	D	Gas	N/A	400	200	N/A	Sig Sig	BP LM	[8]
4	MLP	D	Gas	N/A	N/A	N/A	N/A	N/A	Kalman Filter	[9]
5	Recurrent	D	Gas	N/A	1096	365	10:6:1	N/A	Jordan	[10]
6	MLP Fuzzy-NN	M W	Gas	[-1,1] [0,1]	N/A	N/A	N/A	N/A	N/A	[11]
7	MLP	D	Gas	N/A	N/A	N/A	11:8:4:2 20:12:8:2	N/A	Kalman Filter	[12]
8	1.Adaptive-MLP 2.Functional NN	H	Gas	N/A	N/A	N/A	6:1 9:1	Sig	BP	[13]
9	MLP	D W	Gas	[-1,1]	60%	20%	19:48:1 8:38:1	Log	QP CGD	[5]
10	MLP		Gas	N/A	80%	20%	18:20:1	Log	BP	[14]
11	MLP	D&W	Load	N/A	N/A	N/A	75:10:24	N/A	BP	[15]
12	Fuzzy NN	H&D	Load	N/A	52	N/A	N/A	N/A	N/A	[16]
13	MLP	Y	Load	N/A	N/A	N/A	4:1:1	TanSig/Lin	BP	[6]
14	RBF	Y	Load	[0,1]	15	5	5:10:5	Radbas	N/A	[17]
15	MLP	D	Load	N/A	100	N/A	11:5:5:1	N/A	ALBP	[18]
16	1.MLP 2.Recurrent	Y	Load	N/A	20	3	10:9:1	N/A	BP	[19]

*Note: N/A – not available

A. Prediction Horizon

In load prediction, the supply industry requires forecasts with lead times that range from the short term (minutes, hours, or days ahead) to the long term (up to 20 years ahead) [20]. Most of the prediction addressed more on the short-term load forecasting (STLF) because of its importance to the economic and secure operation of power systems [16]. Daily gas consumption referred to as daily gas send out, is influenced by many factors that affect the amount if natural gas consumed [13]. In four season countries, past investigations show that the majority of gas consumed for most dwellings in a year is for heating purposes, which is mainly related to temperature, wind speed and many other weather factors. However in low humidity countries like Malaysia, the demand for natural gas is not significantly affected by those factors. Unlike short-term load forecasting, long-term load forecasting (LTLF) is mainly affected by economical factors rather than weather conditions. The economic factors and their contributions on long-term loads are the main focus of the long-term load forecasting study [19].

B. Network Architecture

In Table I, most of the authors chose MLP as the network architecture except for Musilek [10], which uses the recurrent architecture, Bakirtzis et al. [16], used Fuzzy NN and Zhi-Sheng Li et al. [17], used Radial Basis Function, (BF)as the network architecture. Some chose to use combination of two or more network architectures to create an adaptive NN. Khotanzad et al. [8],[13], build a combination of three networks; MLP trained with back-propagation, MLP trained with Levenberg-Marquardt algorithm (LM) and Functional Link network. These three separate forecasts are non-linearly combined in the second stage using a functional link ANN combiner [8].

All networks built by the authors have only one hidden layer except for [18], which has two hidden layer for the network structure. There is no theoretical research on how many hidden layers are adequate for a network. The hidden layer actually determines the size of the network. The bigger the size of the network, the more time consumed to train the network. Selection of the number of neurons in the hidden layer is very important as well, although complicated. To date, there are no favorable analytical formulas to define appropriate numbers. All researchers in Table I have selected the number of hidden neurons using the method of constructive method except Z. Li [17] who developed the network neuron size based on the equation found in their research. However, some of the papers do not discuss on how the selection of the number of hidden layer and neurons are made.

C. Data Pre-processing

Among the papers of energy forecasting, only a few mentioned the normalization of the data done in the simulation. Commonly, the normalization lies between 0 to 1 or -1 to 1 as discussed in papers [11],[5],[17]. Peharda et.al [7] suggested another range for normalization other than the two common techniques mentioned previously. Numbers of training and validation data meanwhile gives a significant difference on the network performance. Papers reviewed mostly have divided the data according to 70/30% of data for training and validation for the network development as shown in Table I.

D. Learning Algorithm

Learning is the process of adapting or modifying the connection weights between neurons as a result of the mismatch between the actual and the desired output of the neural network in response to an input presented to the input layer [6].

Back-propagation algorithm is the most popular algorithm over for decades up until now. The BP learning algorithm is an iterative gradient descent procedure [13]. Another popular learning algorithm is the LM which calculates performance with respect to the weight and bias variables. J.J. Moré, [21] discussed more on this learning algorithm. A few other researchers [9],[12], chose to use different algorithm which is the the Kalman filter, a set of mathematical equations that provides an efficient computational means to estimate the state of a process, in a way that minimizes the mean of the squared error [22].

Saini [18] uses the adaptive learning BP where the learning rate is varied according to whether or not iteration decreases the performance index. Only Kizilaslan and Karlik[5] came to a conclusion to use Quick Propagation algorithm (QP) and Conjugate Gradient Descent algorithm (CGD) for the network model. QP algorithm treats the weights as if they were quasi independent and attempts to use a single quadratic model while CGD search linearly to find the optimal network weights' change and corrections of weights is conducted once per iteration [5].

III. Energy Prediction Neural Network Model

A metering system consists of a turbine meter, measuring equipments (pressure transmitter, temperature transmitter), gas chromatography, and a flow computer that calculates the energy consumption of the sales of gas production. The system calculates the energy consumption based on value of inputs from all the measuring equipments. The inputs are gross volume, pressure, temperature, calorific value, and gas components. The reliability of the system must be ensured so that it would not affect the billing integrity between the distributors and consumer.

In the oil and gas industry it is important to have a reliable and accurate metering system for billing purposes. A wrong quantification of product selling or buying will cause lost of income to the company. To achieve this objective, the neural network prediction model will be used to predict the energy consumption as well as to construct a more reliable metering system for billing integrity. The ANN model will learn the relationship between the input parameters and controlled and uncontrolled variables by utilizing previously recorded data. The model will then predict the output based on the earlier trained data for other input.

A. *Neural Network Structure*

In this paper, a structure of multi-layer feedforward neural network with three-layer (input layer-hidden layer-output layer) nodes and sigmoid activation function for the hidden layer and linear activation function for output layer. This structure is trained using the Levenberg-Marquardt algorithm [21]. The inputs are selected based on the energy consumption calculation in the metering system. There are five inputs altogether which are the gross volume (Vg), temperature (T), pressure (P), calorific value (CV) and specific gravity (sg). The output is simply the energy (E) at the output layer.

B. *Inputs Pre-processing*

All inputs are scaled or normalized to lie in [-1,1] range. Scaling of data is necessary under certain circumstances such as when variables span in different ranges. It will predict the short-term energy consumption of sales gas which is the hourly energy forecasting. Data was taken for duration of one year on an hourly basis from the flow computer and gas chromatography in the metering system. Data which is invalid or behave irrelevant are filtered out such as out of ranges or abnormal (spike, zero reading). There are 3303 data altogether. The initial weights are assigned randomly. Training data set and validation data set is divided into a few sets of data division as in Table II to investigate which set is more tolerable for the model.

TABLE II
DATA DIVISION IN DIFFERENT SETS

Data Set	Data Division
Set A	75% training, 25% validation
Set B	50% training, 50% validation
Set C	25% training, 75% validation

The measure of the neural network performance is defined from the Root Means Square Error (RMSE). The error is calculated between the energy predicted (yp) and current energy (y) for each training data set and validation data set. N is the number of data. The equation of the RMSE is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{p_i} - y_i)^2} \qquad (1)$$

C. *Parameter Selection-Learning Algorithm*

There are a few parameters investigated in developing the model and to be considered in order to optimise the performance which are the learning algorithm, the activation function, the training and validation data division and the number of neurons. The neural network is trained using a few numbers of learning algorithms which are listed in Table III. It shows the RMSE value for validation and training.

For every learning algorithm, three types of activation function are tested (refer Table III). In this investigation, the number of neurons in the hidden layer is set to 10 neurons. All learning algorithm are tested in the neural network model with all activation function combinations. The data set used is Set B. From the observation, it can be seen that the learning algorithm LM and LMBR give more promising result rather than the other learning algorithms. The RMSE obtained from CGBPB, GDB and RB is not consistent, shows no real trend and gives quite high RMSE. Between the LM and LMBR, the RMSE are both small but LMBR has smaller RMSE. Meanwhile, using the activation function tan-sigmoid for hidden layer always give better result than log-sigmoid. Out of these five learning algorithms, the most compromising result is obtained from the learning algorithm of Levenberg-Marquadt with Bayesian Regularization with activation function of tan-sigmoid for hidden layer and pure-linear for output layer as highlighted in Table III.

TABLE III
RESULT OF ANN ENERGY PREDICTION FOR LEARNING ALGORITHM

Learning Algorithm	Activation Function (hidden layer)	Activation Function (output layer)	RMSE (train data)	RMSE (valid data)
Levenberg-Marquardt (LM) [23]	Tan-sigmoid	Pure-linear	57.08	160.06
		Tan-sigmoid	61.02	173.79
		Log-sigmoid	190.49	179.02
	Log-sigmoid	Pure-linear	55.74	232.94
		Tan-sigmoid	52.63	342.12
		Log-sigmoid	193.18	278.61
Levenberg-Marquardt with Bayesian Regularization (LMBR) [24]	Tan-sigmoid	Pure-linear	45.58	122.34
		Tan-sigmoid	51.21	187.73
		Log-sigmoid	213.17	292.00
	Log-sigmoid	Pure-linear	46.48	204.64
		Tan-sigmoid	134.13	154.37
		Log-sigmoid	194.78	317.80
Conjugate Gradient Backpropagation with Powell-Beale restarts (CGBPB) [23]	Tan-sigmoid	Pure-linear	209.80	525.27
		Tan-sigmoid	159.73	221.34
		Log-sigmoid	219.85	160.12
	Log-sigmoid	Pure-linear	194.86	499.26
		Tan-sigmoid	135.39	163.54
		Log-sigmoid	212.00	164.98
Gradient descent with Adaptive Learning Rate Backpropagation (GDB) [25]	Tan-sigmoid	Pure-linear	599.70	449.50
		Tan-sigmoid	288.74	531.35
		Log-sigmoid	221.64	309.12
	Log-sigmoid	Pure-linear	571.40	475.26
		Tan-sigmoid	219.15	530.91
		Log-sigmoid	237.69	316.21
Resilient Backpropagation (RB)	Tan-sigmoid	Pure-linear	188.30	149.20
		Tan-sigmoid	178.43	143.46
		Log-sigmoid	218.93	364.66
	Log-sigmoid	Pure-linear	142.45	165.64
		Tan-sigmoid	144.77	149.38
		Log-sigmoid	217.72	154.15

D. *Parameter Selection-Training and Validation Data Division*

The investigation continues to find ratio of validation and training data that is sufficient to predict the energy consumption by means of comparing the RMSE value of each ratio tested. For this purpose, the data is tested with three sets of ratios as in Table II. At the same time, the investigation will also consider performance with a different number of neurons. The model is tested from 1 neuron to 12 neurons. The learning algorithm used is the Levenberg-Marquardt with Bayesian Regularization. The activation function for hidden layer is the tan-sigmoid while for output layer is the pure-linear. Fig. 1 shows the RMSE result for each set of data division. From the result and analysis, it is found that for Set A, the optimum number of neurons ranges from 1 to 12 neurons that give the least RMSE value of 10 neurons.

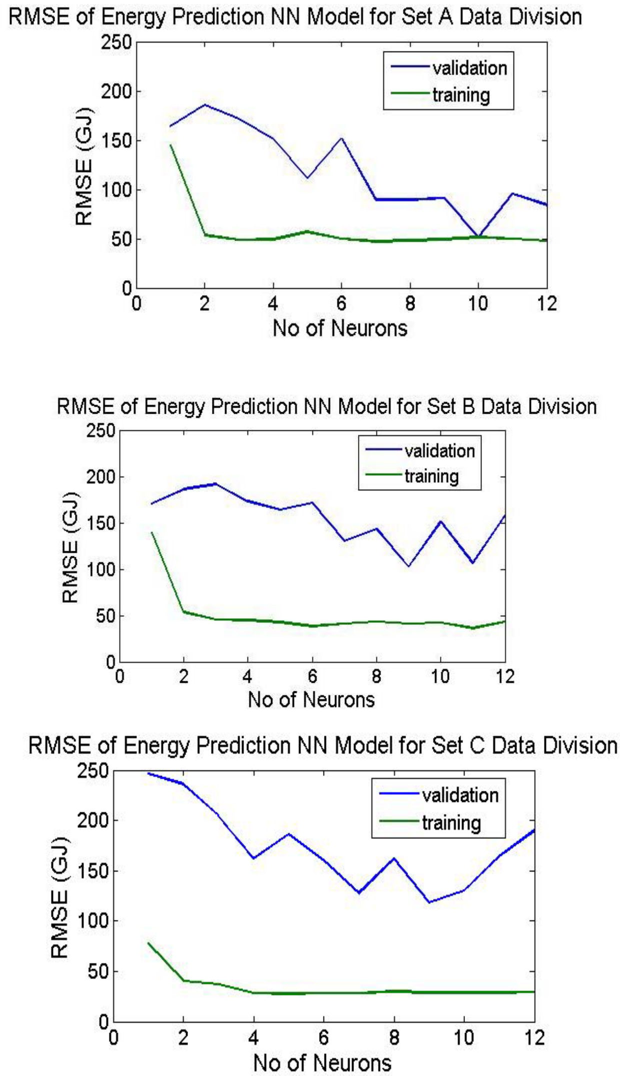

Fig. 1 The training and validation RMSE trends for Set A, Set B and Set C data division

The RMSE for training is 51.34 and 50.62 for validation. The RMSE value was observed to decrease with the increment of number of neurons. After 10 neurons, the RMSE value started to increase

For Set B, the optimum number of neurons that gives the least RMSE value is 9 neurons as shown in the Fig. 1. The RMSE for training data is 41.30 and for validation data is 102.59. As for Set C, the optimum number of neurons that gives the least RMSE value is 9 neurons. The RMSE for training data is 28.76 and for validation data is 118.05. The training RMSE for Set C is very small because the training data is very few compared to the validation data. A small number of training data might give small RMSE but it is actually not enough to predict or validate the data precisely. This has caused the validation of RMSE to be exceedingly higher than other sets of data division.Each set of data shows that as the number of neurons increases, the RMSE for training and validation decreases. After it reaches the optimum number of neurons, the RMSE starts to increase again. From the result it can be observed that for Set A, the model is very well trained as it gives low validation RMSE value as compared to Set B. This is because the number of data for training is sufficient for the model to learn from the example and predict the output as close as the target data.

In Set C, the training RMSE is good but the validation is poor because the number of training data is inadequate for the model to predict the output accurately. All sets of data shows that the number of neurons appropriate for the model is 9 or 10 neurons. The training data gives a good performance as the RMSE values are approaching a steady state line as can be seen in Fig. 1.

E. Parameter Selection-Number of Neurons

Another objective of this investigation is to determine the number of neurons that gives the least RMSE value for both training and validation data set.

The number of neurons manipulated only for the hidden layer and the output layer is made constant to one neuron since there is only one output. The trend for this neural network model is observed and it shows that as the number of neuron increases, the RMSE value decreases. After it reached the optimum number of neurons, the RMSE started to increase. Result from the latter investigation is used for this evaluation. The results are plotted in Fig. 1. Fig. 1 (Set A) shows the plotted RMSE for both training and validation data with number of neurons. The training RMSE reaches a steady state line while the validation RMSE is decreasing with increasing number of neurons. The validation RMSE is almost the same as the training RMSE when the number of neurons is 10.

As can be seen, for Set B, the RMSE for validation started to increase after 9 neurons. The training RMSE also become steady but somehow the divergence between the training RMSE and the validation RMSE is still large.

The RMSE for Set C is also plotted as shown in Fig. 1 (Set C). Looking at the trending, the RMSE for validation is higher compared to other division of data. The RMSE for training on the other hand is smaller compared to the other two data division results. This is because a smaller number of data for training is used.

In all the three sets of neural network model with different data division percentage, one obvious similarity is that all models has the least value of RMSE when 9 or 10 neurons are used in the model. Although out of the three models the 10 neurons neural network model gives the least RMSE, but considering that the other two sets of model obtained the same number of neurons best resulting the least RMSE, it is recommended that the 9 neurons neural network energy prediction model to be chosen as the best model with 75 percent training data and 25 percent validation data.

IV. Conclusion

This paper presents the development of Neural Network model for energy consumption with adequate performances. Based on the performance the developed model was constructed based on a 3-layer structure with 9 neurons at the hidden layer. Meanwhile the activation function for the hidden layer and the output layer is Tan-sigmoid and pure-linear respectively. The investigation shows that the learning algorithm, Lavenberg-Marquardt with Bayesian Regularization is able to provide the best performance as compared to other learning algorithm. Data division also plays an important role in providing a reliable insight of the system. The developed model was obtained with 75 percent of the data for training exercise and the remaining 25 percent for validation purposes. The paper also manages to highlight a critical aspect that neural network designer's face in developing a neural network model that is to choose an appropriate network size for a given application. Network size involves in the case of layered neural network architectures, the number of layers in a network, the number of nodes per layer, and the number of connections. A reliable neural network model for energy prediction is very important in forecasting accurately and precisely. Therefore, the most important part in developing an ANN model is the parameter selection in order to optimize the performance. ANNs are being used increasingly for the prediction and forecasting of load and energy consumption. Analysis on the ANN performance will actually determine the robustness of the system.

Acknowledgment

M.J. Ismail, R. Ibrahim and I. Ismail would like to thank UTP for the graduate assistant scheme program to fund the research and PETRONAS for the project collaboration.

References

[1] H. R.Maier and G. C.Dandy, "Neural network for the prediction and forecasting of water resources variables: a review of modelling issues and applications," *Environmental Modelling & Software*, vol. 15, 2000, pp. 101-124.

[2] P. Benardos and G. Vosniakos, "Optimizing feedforward artificial neural network architecture," *Engineering Applications of Artificial Intelligence*, vol. 20, Apr. 2007, pp. 365-382.

[3] M. N Jamal, M. E Ibrahim, and A. N Salam, "Multilayer Perceptron Neural Network (MLPs) For Analyzing the Properties of Jordan Oil Shale," *World Applied Sciences Journal*, vol. 5, 2008, pp. 546-552.

[4] D. Ivezić, "Short-Term Natural Gas Consumption Forecast," *FME Transactions*, vol. 34, 2006, pp. 165-169.

[5] R. Kizilaslan and B. Karlik, "Comparison neural networks models for short term forecasting of natural gas consumption in Istanbul," *Applications of Digital Information and Web Technologies, 2008. ICADIWT 2008.*, 2008, pp. 448-453.

[6] K. Kavaklioglu, H. Ceylan, H.K. Ozturk, and O.E. Canyurt, "Modeling and prediction of Turkey's electricity consumption using Artificial Neural Networks," *Energy Conversion and Management*, vol. 50, Nov. 2009, pp. 2719-2727.

[7] D. Peharda, M. Delimar, and S. Loncaric, "Short term hourly forecasting of gas consumption using neural networks," *Information Technology Interfaces, 2001. ITI 2001. Proceedings of the 23rd International Conference on*, 2001, pp. 367-371.

[8] A. Khotanzad and H. Elragal, "Natural gas load forecasting with

combination of adaptive neural networks," *Neural Networks, 1999. IJCNN '99. International Joint Conference on*, 1999, pp. 4069-4072 vol.6.

[9] R. Brown and I. Matin, "Development of artificial neural network models to predict daily gas consumption," *Industrial Electronics, Control, and Instrumentation, 1995., Proceedings of the 1995 IEEE IECON 21st International Conference on*, 1995, pp. 1389-1394 vol.2.

[10] P. Musilek, E. Pelikán, T. Brabec, and M. Simunek, "Recurrent Neural Network Based Gating for Natural Gas Load Prediction System," *Neural Networks, IJCNN'06*, 2006, pp. 3736-3741.

[11] Nguyen Hoang Viet and J. Mandziuk, "Neural and fuzzy neural networks for natural gas consumption prediction," *Neural Networks for Signal Processing, 2003. NNSP'03. 2003 IEEE 13th Workshop on*, 2003, pp. 759-768.

[12] R. Brown, P. Kharouf, Xin Feng, L. Piessens, and D. Nestor, "Development of feed-forward network models to predict gas consumption," *Neural Networks, 1994. IEEE World Congress on Computational Intelligence., 1994 IEEE International Conference on*, 1994, pp. 802-805 vol.2.

[13] A. Khotanzad, H. Elragal, and T. Lu, "Combination of artificial neural-network forecasters for prediction of natural gas consumption," *Neural Networks, IEEE Transactions on*, vol. 11, 2000, pp. 464-473.

[14] L. Xu, W. Zhou, X. Li, and S. Tang, "Wet Gas Metering Using a Revised Venturi Mete and Soft-Computing Approximation Techniques," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, Mar. 2011.

[15] J. Fidalgo and M. Matos, "Forecasting Portugal Global Load with Artificial Neural Networks," *Artificial Neural Networks – ICANN 2007*, Springer Berlin / Heidelberg, 2007, pp. 728-737.

[16] A. Bakirtzis, J. Theocharis, S. Kiartzis, and K. Satsios, "Short term load forecasting using fuzzy neural networks," *Power Systems, IEEE Transactions on*, vol. 10, 1995, pp. 1518-1524.

[17] Z. Li, G. Zhang, D. Li, X. Liu, S. Mei, and J. Wu, "Neural network prediction of energy demand and supply in China," *Proceedings of the Institution of Civil Engineers - Energy*, vol. 160, 2007, pp. 145-149.

[18] L.M. Saini, "Peak load forecasting using Bayesian regularization, Resilient and adaptive backpropagation learning based artificial neural networks," *Electric Power Systems Research*, vol. 78, Jul. 2008, pp. 1302-1310.

[19] B. Kermanshahi and H. Iwamiya, "Up to year 2020 load forecasting using neural nets," *International Journal of Electrical Power & Energy Systems*, vol. 24, Nov. 2002, pp. 789-797.

[20] H. Hippert, C. Pedreira, and R. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *Power Systems, IEEE Transactions on*, vol. 16, 2001, pp. 44-55.

[21] J.J. Moré, "The Levenberg-Marquardt algorithm: Implementation and theory," *Numerical Analysis*, Springer Berlin / Heidelberg, 1978, pp. 105-116.

[22] G. Welch and G. Bishop, *An Introduction to the Kalman Filter*, University of North Carolina Chapel Hill, 2006.

[23] D.J.C. MacKay, "Bayesian Interpolation," *Neural Computation*, vol. 4, 1992, pp. 415-447.

[24] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: the RPROP algorithm," *Neural Networks, 1993., IEEE International Conference on*, 1993, pp. 586-591 vol.1.

[25] P. Baldi, "Gradient descent learning algorithm overview: a general dynamical systems perspective," *Neural Networks, IEEE Transactions on*, vol. 6, 1995, pp. 182-195.

Maryam Jamela was born in Pahang, 1986, holds a B.Hons in Electrical & Electronics Engineering, Universiti Teknologi PETRONAS, Perak, Malaysia in 2009, currently pursuing M.Sc in Electrical & Electronics Eng, from the same university. She is also currently working as an Instrument Engineer at Ranhill Worley Parsons. Her published articles are Selection of Network Architecture and Input Sensitivity Analysis for a Neural Network Energy Prediction Model (Proceeding of 2010 International Conference on Intelligent and Advanced Systems (ICIAS 2010), Kuala Lumpur, IEEE, 2010), Adaptive Neural Network Prediction Model for Energy Consumption (Proceeding of 2011 International Conference on System Engineering and Modeling (ICSEM 2011), China, IEEE, 2011). Parameter Selection on Neural Network Prediction Model of Energy Consumption (International Conference on Computer, Electrical and Systems Sciences and Engineering, Bali, 2011). Her research interest is in Neural Network. Maryam Jamela Ismail is a member of the Board of Engineers, Malaysia.

Rosdiazli Ibrahim received the BEng. Electrical Engineering. from Univeristi Putra Malaysia in 1996. In 2000, he awarded a Master of Science in Automation and Control from the University of Newcastle Upon-Tyne, UK with Distinction. He was conferred a PhD degree in Electrical & Electronic Eng. from the University of Glasgow, UK in 2008.

He joined Universiti Teknologi PETRONAS in 1999 and currently is a Senior Lecturer in the Electrical & Electronic Engineering department. Formerly, he was an Instrument Engineer at PETRONAS Fertilizer (Kedah) Sdn. Bhd. He has presented and published several articles in the area of intelligent control and nonlinear multivariable process modeling which is very much in line with his research interest.

Idris Ismail received a B.S. Electrical Eng. from Wichita State Uni, Kansas in 1986, M.Sc. in Control System for the University of Sheffield, UK in 2000 and PhD in Electrical & Electronic Eng. from the University of Manchester, UK in 2009.

He is a Senior Lecturer in the Electrical & Electronic Engineering department at Universiti Teknologi PETRONAS which he joined in 1998. He holds a position as the Head of Student Industrial Internship Program and is a Senate member of the university. From 1986 to 1998, he was working with PETRONAS Penapisan (T) Sdn. Bhd. and Ethylene Polyethylene Malaysia where his last position was as an Instrument Manager. His research interests lie in the area of multiphase flow measurement, process tomography and plant process control and instrumentation. From 2004 to 2007 he had the opportunity to work with Schlumberger Cambridge Research, UK on a joint research work with the University of Manchester.

Dr I Ismail is a registered professional Electrical Engineer with Board of Engineers, Malaysia and member of the Institute of Engineers, Malaysia.