Bulletin of Electrical Engineering and Informatics
Vol. 7, No. 2, June 2018, pp. 272~278
ISSN: 2302-9285, DOI: 10.11591/eei.v7i2.1177

Results of Fitted Neural Network Models on Malaysian Aggregate Dataset

Nor Azura Md Ghani[1], Saadi Bin Ahmad Kamaruddin[2], Ismail Musirin[3], Hishamuddin Hashim[4]
[1,2]Center for Statistical Studies and Decision Sciences, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA,40450 Shah Alam, Selangor, Malaysia
[3]Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, Selangor Darul Ehsan, Malaysia
[4]Pilihsepakat Sdn. Bhd., Lot 5431, Tingkat 1(B), Jalan J9, Fasa 6, Taman Melawati, Kuala Lumpur, Malaysia

Article Info

Article history:

Received Jan 25, 2018
Revised Mar 26, 2018
Accepted Apr 09, 2018

Keywords:

Aggregate
Backpropagation
Neural network
Ordinary Nonlinear Autoregressive (NAR)
Ordinary Nonlinear Autoregressive Moving Average (NARMA)

ABSTRACT

This result-based paper presents the best results of both fitted BPNN-NAR and BPNN-NARMA on MCCI Aggregate dataset with respect to different error measures. This section discusses on the results in terms of the performance of the fitted forecasting models by each set of input lags and error lags used, the performance of the fitted forecasting models by the different hidden nodes used, the performance of the fitted forecasting models when combining both inputs and hidden nodes, the consistency of error measures used for the fitted forecasting models, as well as the overall best fitted forecasting models for Malaysian aggregate cost indices dataset.

Corresponding Author:

Nor Azura Md Ghani,
Center for Statistical Studies and Decision Sciences,
Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA,
40450 Shah Alam, Selangor, Malaysia.
Email: azura@tmsk.uitm.edu.my

1. INTRODUCTION

According to [1], input lags alone may not be adequate to precisely rough the underlying generation process. He added that one lag is not able to capture the underlying relationship or not sufficient to model a process, thus there is a need of adequate number of hidden neurons for optimal network performance. Both inputs and hidden nodes can significantly affect the learning of a neural network when forecasting a time series. This research is interested in accuracy of the models instead of parsimonious time series models, as suggested by [2]. For time series, inputs also include lags. As the input dimensionality increases, model complexity increases and learning becomes more difficult, leading to poor convergence. The challenge here is to select or find the best combinations of input lags with adequate number of hidden nodes that will lead to a superior model, specifically to a particular dataset of a certain scenario. If the training set is sufficiently large, ANN will generalise accurately and will produce accurate outputs for inputs not in the training set. When neural networks are properly trained, they can generalise and extrapolate additional details of the function mapping the inputs to outputs. Nevertheless, no prior assumptions of dependent and independent variables need to be made since the neural network is trained on observed data [3], [4]. This infers neural systems are ordinarily used as "black-box" devices that are no earlier information about the method was assumed; the goal was to develop a process model based only on observations of its input-output behaviour

[5-7]. By definition, the black-box is a modelling technique that constructs the model just based on the information gained from the framework, and it doesn't depend on upon other information about the system [8]. No early presumptions about the model structure are made, and rather, the modelling technique's concern is to make a generic model that maps the information yield relationship of the dataset [9, 10]. Revelation models are known to be viable and versatile, as the assistant information about the system may not be promptly accessible. From this investigation, one shortcoming lies in the way that that the model complexity expands [7] Also, another favorable position of the black-box identification is that it can demonstrate flow that are inadequately caught by mathematical models [12], [13].

2. METHODOLOGY

In this research, Malaysian Aggregate monthly data from January 1980 to December 2013 were adapted. These datasets are secondary data collected from three different sources, UKAS, from Malaysian Prime Minister's Office, Malaysian CIDB, and Malaysian Statistical Department. In this research, only Malaysian central region aggregate price indices were adapted which comprise of 3.9% outliers. The total N=408 (12 months x 34 years) from January 1980 to December 2013 (base 1980=100). The mean of Malaysian Aggregate is 113.7731. Based on the Jarque-Bera test for normality, the variable aggregate (J-B=0.873, p=0.000) is highly significant at 99% confidence interval.

First and foremost, this research discusses the performance of the fitted forecasting models on Malaysian aggregate cost indices data by each set of input lags and error lags used. Table 1 shows performance results of fitted forecasting models of two-layer tansig/linear transfer functions on Malaysian aggregate cost indices data. For comparison purposes, this research used input lags from 5 to 40 and hidden nodes from 5 to 45 for BPNN-NAR model, while for the BPNN-NARMA model, this research used the error lags ranging from 5 to 40, together with input lags from 5 to 40 and hidden nodes from 5 to 45. The most common approach to decide this property is experimentation or trial-and-error [14-16], albeit different strategies (rule-of-thumb) and algorithms (pruning and growing) are additionally accessible [17]-[19]. It is to be noticed that BPNN-NAR is a feedforward neural network type model, while BPNN-NARMA is a recurrent neural network type model [18]. This exploration reports the error measures on the test dataset, which is the most vital characteristic, reflecting ANN's generalisation ability [19], [20]. Table 2 shows the data partitioning for network pre-processing. Figure 1 shows the flowchart of the BPNN-NAR and BPNN-NARMA.

Table 1. Combinations of Input Lags, Error Lags and Percentage of Outliers for BPNN - NAR and BPNN - NARMA Models

No.	Data Type	Notation	Outliers	ANN Models			Hidden Nodes
				NAR	NARMA		
				Input Lags	Input Lags	Error Lags	
1.	Aggregate	Agg	3.9%	5, 10, 15, 20, 25, 30, 35, 40	5, 10, 15, 20, 25, 30, 35, 40	5, 10, 15, 20, 25, 30, 35, 40	5, 10, 15, 20, 25, 30, 35, 40, 45

Table 2. Size of Data Partitioning for Training, Validation and Testing Sets for Each Data Used in This Research

No.	Data Type	Total Sample Size (N)	In-sample Data		Out-of-sample Data
			Training (70%)	Validation (15%)	Testing (15%)
1.	Monthly Malaysian Aggregate Price Indices Data	408	286	61	61

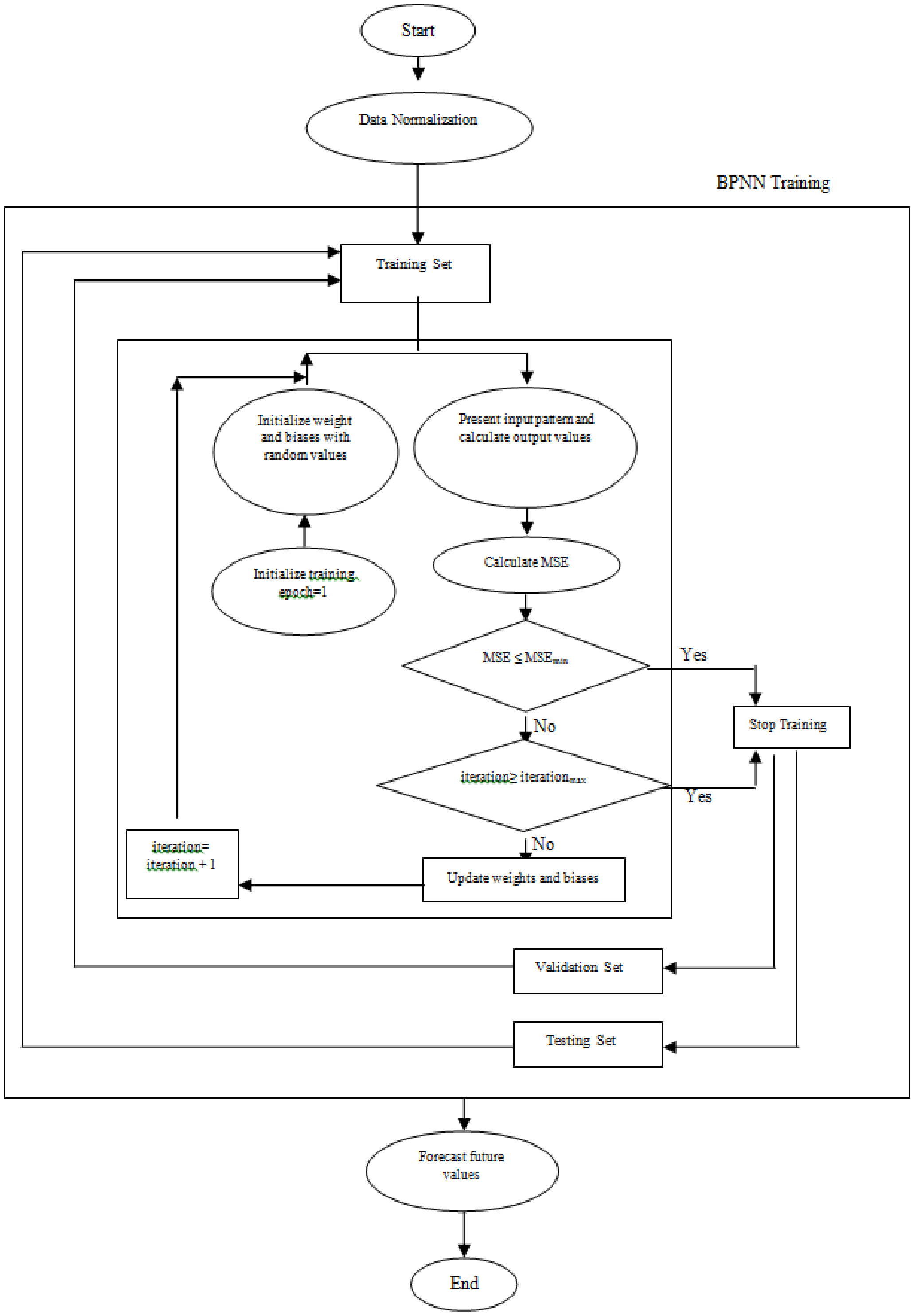

Figure 1. Flowchart of Backpropagation Neural Network NAR and NARMA Mechanism

3. RESEARCH FINDINGS

Figure 2 shows the performance of BPNN-NAR with respect to input lags and hidden nodes on Malaysian aggregate cost indices data based on RMSE. Whereas, Figure 3 shows the performance of BPNN-NARMA with respect to input lags and hidden nodes on Malaysian aggregate cost indices data based on

RMSE. From the radar diagrams of both figures, this research can clearly see that overall of the lines of BPNN-NAR are more towards the center of zero values, meaning zero error compared to BPNN-NARMA.

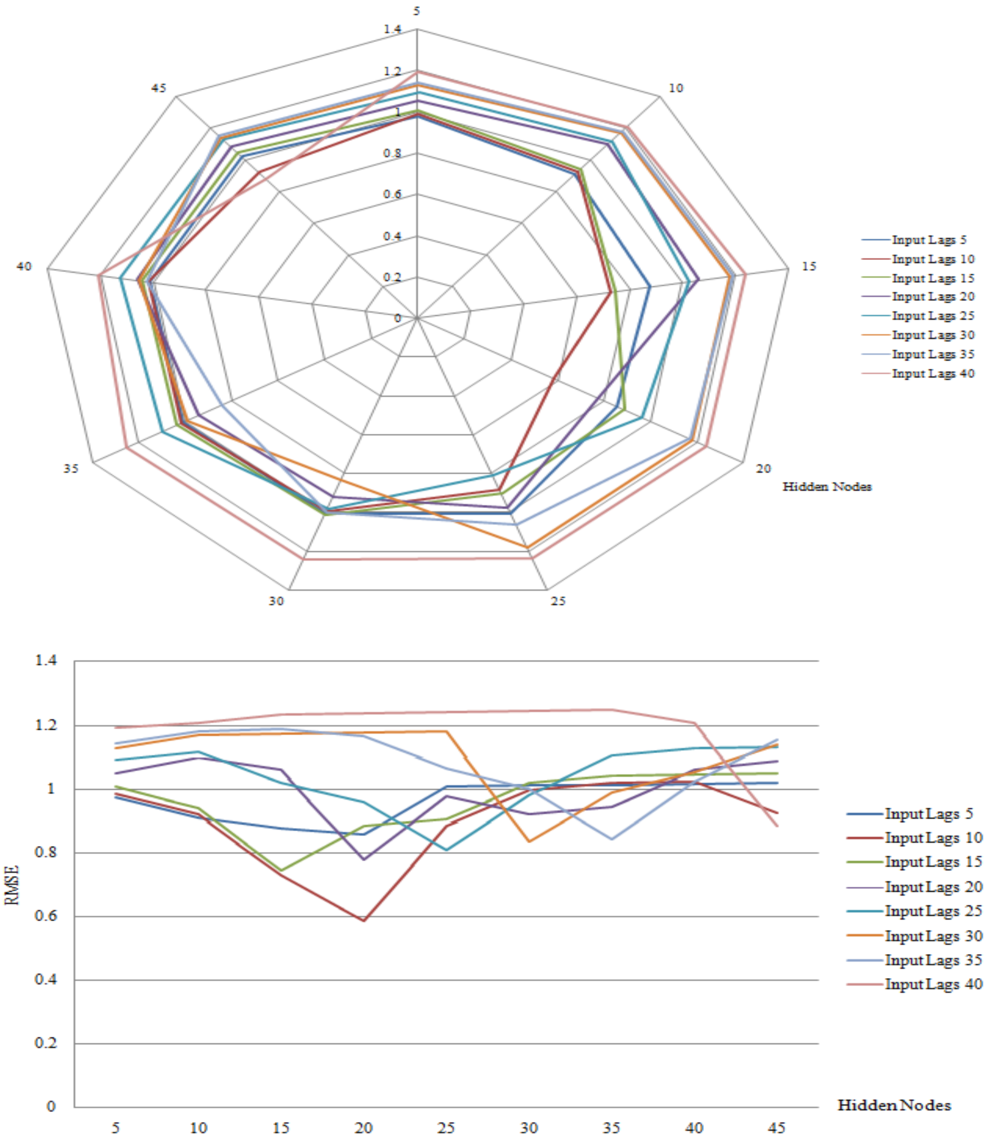

Figure 2. Performance of BPNN-NAR with respect to input lags and hidden nodes on Malaysian Aggregate cost indices data based on RMSE

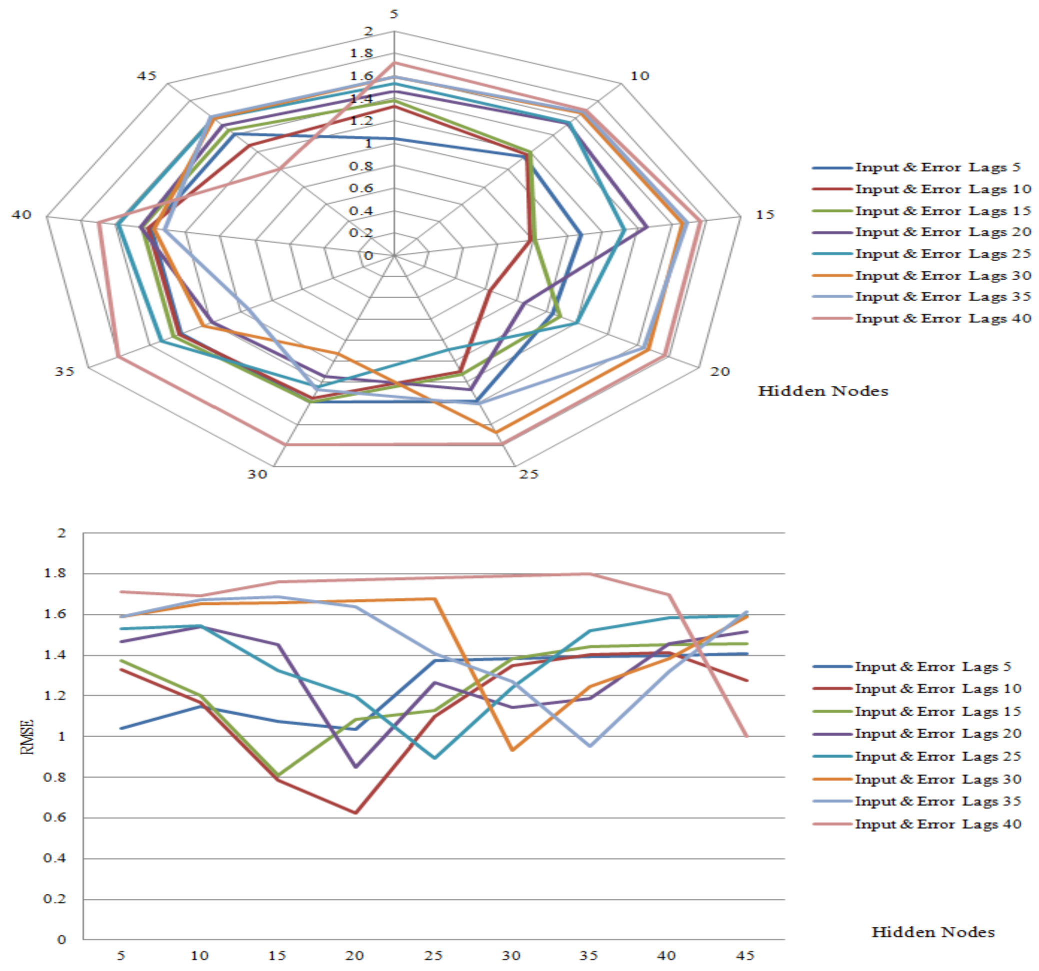

Figure 3. Performance of BPNN-NARMA with respect to input lags, error lags and hidden nodes on Malaysian aggregate cost indices data based on RMSE

Based on input and error lags 5, the optimal number of hidden nodes was 20. Based on input and error lags 10, the optimal number of hidden nodes was 20. Based on input and error lags 15, the optimal number of hidden nodes was 15. Based on input and error lags 20, the optimal number of hidden nodes was 20. Based on input and error lags 25, the optimal number of hidden nodes was 25. Based on input and error lags 30, the optimal number of hidden nodes was 30.

Based on these results, BPNN-NARMA model performed worse than the BPNN-NAR model. By right, BPNN-NARMA should perform better than BPNN-NAR [21]-[23]. This is perhaps due to the outliers in the data. In most cases, as the number of input and error lags increased, the errors are also increased.

Furthermore, this research discusses on the performance of the fitted forecasting models by different hidden nodes used. From Table 1, over the input and error lags, if the network is assigned with fewer nodes or huge load of nodes, the performance of the network will get worse. This happens due to the network overfitting problem [24]-[28]. Therefore, an adequate number of hidden nodes is important in neural network modelling [29]-[31]. It can be concluded that if the network is given an adequate number of hidden nodes, the performance of the network will get better and model errors will be reduced, and the chance of getting a more accurate model can be achieved.

In terms of the consistency of error measures used for the fitted forecasting models, this research concludes that RMSE performed consistently. This is supported by recent works in the forecasting arena such as [32-36]. Moreover, in most cases, many researchers used RMSE to evaluate their forecasting models such as [37, 38]. Therefore, RMSE will be the used to assess the performance of the models throughout this research. The best results for this section based on RMSE can be summarized in Table 2.

Lastly, it can be concluded that the best fitted model for Malaysian aggregate cost indices data was the two-layer tansig/linear transfer functions BPNN-NAR model with 10-20 configurations (RMSE=0.584, MSPE=0.341, MAPE=55.378, MAD=0.429, and GRMSE=0.782), can be seen in Table 3.

Table 3. Best Results of Ordinary BPNN-NAR and BPNN-NARMA Models on Malaysian Aggregate Cost Indices Data based on Different Lags

Input Lags	Error Lags	Hidden Nodes	RMSE BPNN-NAR	RMSE BPNN- NARMA
5	5	20	0.857	1.036
10	10	**20**	**0.584**	0.626
15	15	15	0.744	0.808
20	20	20	0.766	0.850
25	25	25	0.807	0.892
30	30	30	0.834	0.931
35	35	35	0.843	0.950
40	40	45	0.882	1.003

4. CONCLUSION

The investigation of BPNN-NAR on MCCI data demonstrates that higher input lags implies higher RMSE. Correspondingly, the higher or the lesser hidden nodes to the input lags, the higher the RMSE, and the higher the input lags, the higher the RMSE.

Despite the fact that NARMA is superior to NAR, when there are outliers in the dataset, the NARMA display appeared to separate, and the NAR demonstrate outflanked the NARMA show. Here, this examination proposes that the NAR model ought not be amplified or gone before by NARMA demonstrate when the dataset comprises of outliers.

The NARMA shows performed more terrible than NAR when there exist outliers in the time arrangement datasets. NARMA model should perform better contrasted with NARMA demonstrate. This can be demonstrated by the aftereffects of both models when fitting Malaysian sand cost indices data, which don't comprise of outliers issue. From the results, it is clear that that there is a need to modify NAR and NARMA models so that they may handle outliers issue effectively, or else the expansion of NAR model to NARMA model ought not be continued since the NARMA model will tend deliver bigger errors, and the outcomes are not solid for further utilize. In the future, the results will be compared to the existing works of [39], [40].

ACKNOWLEDGEMENT

We want to express our gratefulness and appreciation to Unit Kerjasama Awam Swasta (UKAS) over Prime Minister's Office, Construction Industry Development Board (CIDB) then Malaysian Statistics Department. Appreciation to Universiti Teknologi MARA for financial support under the Grant Scheme No. 600-RMI/DANA 5/3/CIFI (65/2013).

REFERENCES

[1] Samarasinghe, S, “Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition”, CRC Press. 2006.
[2] Zhang, G. P, “Business forecasting with artificial neural networks: An overview”, *Neural networks in business forecasting*, 2004. 1-22.
[3] Kim, D, “Neural Networks for Trip Generation Model”, *Journal of the Eastern Asia Society for Transportation Studies*, 2001; 4(2): 201-208.
[4] Ghosh, S. & Konar, A, “Call Admission Control in Mobile Cellular Networks”, 2012, Vol. 437, Springer.
[5] Qiu, F. & Jensen, J. R, “Opening the black box of neural networks for remote sensing image classification”, *International Journal of Remote Sensing*, 2004; 25(9): 749-1768.
[6] Heinert, M, “Artificial neural networks–how to open the black boxes.Application of Artificial Intelligence“ in *Engineering Geodesy*, 2008; 1: 42-62.
[7] Hametner, C. & Jakubek, S, “Nonlinear Identification with Local Model Networks Using GTLS Techniques & Equality Constraints”, *IEEE Trans Neural Networks*, 2011; 22(9): 1406-1418.
[8] Barbosa, B. H. G., *at. al*, "Black and Gray-Box Identification of a Hydraulic Pumping System”, *IEEE Trans. Control Systems Technology*, 2011; 19(2): 398-406.
[9] Ahmadi, S. S. & Karrari, M., "Nonlinear identification of synchronous generators using a local model approach", *Przeglad Elektrotechniczny (Electrical Review)*, 2011; 8:166-170.
[10] Talmon, R., *at. al*, "Parametrization of linear systems using diffusion kernels", *IEEE Transactions on Signal Processing*. 2012; 60(3): 1159-1173.
[11] Olvera M. A. G. & Tang, Y, "Black-Box Identification of a Class of Nonlinear Systems by a Recurrent Neurofuzzy Network", *IEEE Trans Neural Networks*, 2010; 21(4): 672-679.
[12] Dong, N. & Chen, Z., “A novel data based control method based upon neural network and simultaneous perturbation stochastic approximation”, *Nonlinear Dynamics*, 2012; 1(67):. 957-963.
[13] Safak, C., *at. al.*, “Pneumatic motor speed control by trajectory tracking fuzzy logic controller”, *Sadhana*, 2010; 35(1): 75-86.

[14] Rashid, M. T., *at. al*. "Nonlinear model identification for Artemia population motion". Nonlinear Dynamics, 2012; 69(4): 2237-2243.
[15] Boëly, N. & Botez, R. M., "New Approach for the Identification and Validation of a Nonlinear F/A-18 Model by Use of Neural Networks", *IEEE Trans Neural Networks,* 2010; 21(11): 1759-1765.
[16] Ren, X. & Lv, X., "Identification of Extended Hammerstein Systems Using Dynamic Self-Optimizing Neural Networks", IEEE Trans Neural Networks, 2011; vol. 22(8): 1169-1179.
[17] Boëly, N. & Botez, R. M. "New Approach for the Identification and Validation of a Nonlinear F/A-18 Model by Use of Neural Networks", *IEEE Trans Neural Networks,* 2010; 21(11): 1759-1765.
[18] Arnerić, J., Poklepović, T., & Aljinović, Z, "GARCH based artificial neural networks in forecasting conditional variance of stock returns", *Croatian Operational Research Review*, 2014; 5(2): 329-343.
[19] Beliakov, G., Kelarev, A. & Yearwood, J. "Robust artificial neural networks & outlier detection. Technical repor",.2011.
[20] Kordos, M., & Rusiecki, A. "Reducing noise impact on MLP training". *Soft Computing*, 2015; 1(1): 1-17.
[21] Bhuyan, M. K., Mohapatra, D. P. & Sethi, S., "Prediction Strategy for Software Reliability Based on Recurrent Neural Network", *Computational Intelligence in Data Mining,* 2016; 1(2): 295-303.
[22] Adams, M. D. & Kanaroglou, P. S. "Mapping real-time air pollution health risk for environmental management: Combining mobile & stationary air pollution monitoring with neural network models", *Journal of environmental management*, 2016; 168: 133-141.
[23] Oliveira, T. P., Barbar, J. S., & Soares, A. S., "Computer network traffic prediction: a comparison between traditional & deep learning neural networks", *International Journal of Big Data Intelligence*, 2016; 3(1): 28-37.
[24] Jung, S. K. & Kim, T. W. "New approach for the diagnosis of extractions with neural network machine learning", *American Journal of Orthodontics and Dentofacial Orthopedics*, 2016; 149(1), pp. 127-133.
[25] Sudar, D., Vršnak, B. & Dumbović, M. (2016). Predicting coronal mass ejections transit times to Earth with neural network. *Monthly Notices of the Royal Astronomical Society*, *456*(2), pp. 1542-1548.
[26] Macas, M., Moretti, F., Fonti, A., Giantomassi, A., Comodi, G., Annunziato, M. & Capra, A. (2016). The role of data sample size and dimensionality in neural network based forecasting of building heating related variables. *Energy and Buildings*, 111, pp. 299-310.
[27] Sutcliffe, P. & Rennie, M. R. (2016). Neural Network Model Predictive Control of Wind Tunnel Test Conditions. *54th AIAA Aerospace Sciences Meeting, AIAA SciTech Forum*, 1(1), pp. 1-15.
[28] Pelletier, F., Masson, C., & Tahan, A. (2016). Wind turbine power curve modelling using artificial neural network. *Renewable Energy*, 1(89), pp. 207-214.
[29] Grahovac, J., Jokić, A., Dodić, J., Vučurović, D. & Dodić, S. (2016). Modelling and prediction of bioethanol production from intermediates and byproduct of sugar beet processing using neural networks. *Renewable Energy*, 85, pp. 953-958.
[30] Dey, S., Sultana, N., Kaiser, M. S., Dey, P. & Datta, S. (2016). Computational intelligence based design of age-hardenable aluminium alloys for different temperature regimes. *Materials & Design*, 92, pp. 522-534.
[31] Kaya, Z. (2016). Predicting Liquefaction-Induced Lateral Spreading by Using Neural Network and Neuro-Fuzzy Techniques. *International Journal of Geomechanics*, 16(4), pp. 401-509.
[32] Kim, M. S. (2016). Analysis of short-term forecasting for flight arrival time. *Journal of Air Transport Management*, *52*, pp. 35-41.
[33] Lahmiri, S. (2016a). Interest rate next-day variation prediction based on hybrid feedforward neural network, particle swarm optimization, and multiresolution techniques. *Physica A: Statistical Mechanics and its Applications*, *444*, pp. 388-396.
[34] Lahmiri, S. (2016b). Intraday stock price forecasting based on variational mode decomposition. *Journal of Computational Science*, *12*, pp. 23-27.
[35] Liu, S., Xu, L. & Li, D. (2016). Multi-scale prediction of water temperature using empirical mode decomposition with back-propagation neural networks. *Computers & Electrical Engineering*, 1(49), pp. 1-8.
[36] Solis, A. O. (2016). Forecasting Lumpy Dem & : Statistical Accuracy & Inventory Control Performance. *Supply Management Research*, 1(1), pp. 79-95.
[37] Feng, X., Li, Q., Zhu, Y., Hou, J., Jin, L. & Wang, J. (2015). Artificial neural networks forecasting of PM 2.5 pollution using air mass trajectory based geographic model and wavelet. *Atmospheric Environment*, 1(107), pp. 118-128.
[38] Wang, W. C., Chau, K. W., Qiu, L. & Chen, Y. B. (2015). Improving forecasting accuracy of medium & long-term runoff using artificial neural network based on EEMD decomposition. *Environmental research*, 1(139), pp. 46-54.
[39] Saadi Bin Ahmad Kamaruddin, Nor Azura Md Ghani, Norazan Mohamed Ramli,(2018). Consolidated Backpropagation Neural Network for Malaysian Construction Costs Indices Data with Outliers Problem. Pertanika J. Sci. & Technol., 26 (1): 353 – 366.
[40] Ghani, N. A. M., bin Ahmad Kamaruddin, S., Ramli, N. M., Musirin, I., & Hashim, H. (2017). Modified BPNN via Iterated Least Median Squares, Particle Swarm Optimization and Firefly Algorithm. Indonesian Journal of Electrical Engineering and Computer Science, 8(3), pp. 779-786.