11

Optimizing Feed Forward Backpropagation Neural Network Based on Teaching-Learning-Based Optimization Algorithm for Long-Term Electricity Forecasting

Widi Aribowo[1]*

[1]*Departement of Electrical Engineering, Universitas Negeri Surabaya, Indonesia*
* Corresponding author's Email: widiaribowo@unesa.ac.id

Abstract: The improvement of the economy country along with developments in each sector will be followed by technological progress. This has led to an increase in electricity demand in Indonesia every year. Electricity load forecasting has an important role in the energy management system. The purpose of electricity load forecasting is as an effort to balance between electricity demand and electricity supply. Power system management can be said to be good if the planning carried out will make a major contribution to the development of electric power systems. In this study, the Feed Forward Backpropagation Neural Network method (FFBNN) is optimized using the Teaching-Learning-Based Optimization Algorithm (TLBO) to predict long-term electricity load. The results were measured and validated using the Mean Absolute Percentage Error (MAPE) method. The performance of the proposed method will be compared with the actual data, feed forward backpropagation neural network (FFBNN) and cascade forward backpropagation neural network (CFBNN). It was found that the TLBO-FFBNN method had an average MAPE value of 0.00004936%. The Comparison results show the effectiveness of the proposed method. The proposed method shows that is adapted and has a good MAPE value.

Keywords: Feed forward backpropagation neural network, Teaching-learning-based optimization, Artificial intelligence, Load forecasting, Metaheuristic.

1. Introduction

Electrical energy has become a major requirement for human life today. The energy of electricity always increases over time [1]. Increased electricity load requires energy service providers to increase electricity capacity electrical load, so as to improve the resolution and quality [2]. Electrical energy cannot be stored on a large scale. Therefore, this energy must be available when is needed. As a result, problems arise in dealing with the need for electricity that is not fixed from time to time, how to operate a system electric power that can meet the demand for power at any time, with good quality [3]. When power was sent from generating much more than the demand for power at load, then there will be a problem of energy waste in the company electricity [4]. Meanwhile, if the power generated and transmitted is lower or does not meet the needs of the

consumer's burden then there will be a local blackout at the expense, which consequently harms the consumer.

Electricity load forecasting [5] is divided into 3 types, namely short-term forecasting which has a period ranging from hourly to per day, medium-term load forecasting namely electricity load forecasting with periods per month to per year, and long-term load forecasting that predicts loads in periods over per year.

The application of "traditional" methods for estimating costs continues. On the other hand, machine learning methods are increasingly becoming a favorite and are widely used in electrical load forecasting research. In recent years, research based on artificial intelligence methods has been published. While countless techniques include support vector machines (SVM), ANN has recently become the most popular research area. The application of ANN
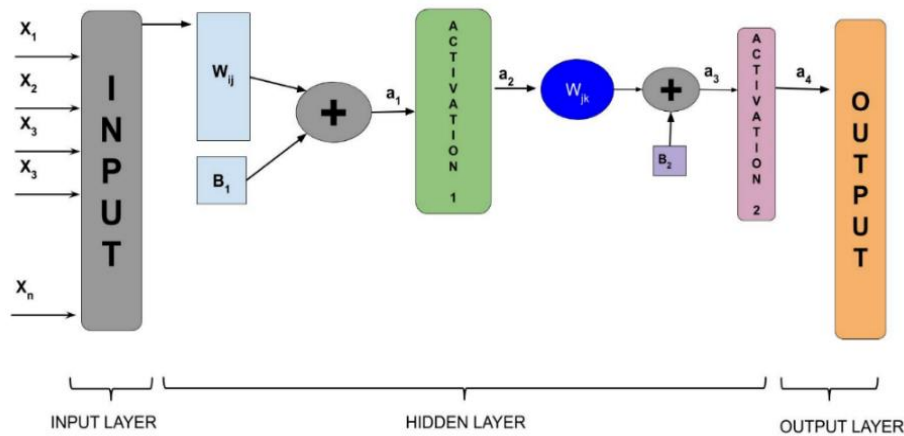
Figure. 1 Feed forward backpropagation neural network structure

methods is increasing with the development of science, and various techniques have been adopted according to specific purposes.

This makes the usage of data that is increasingly large and often called big data. This makes the applying of ANN increasingly sophisticated. Kong et al [6] conducted a study based on a long short-term memory (LSTM) recurrent neural network (RNN) framework for a short-term load forecasting task for individual residential households. Zheng et al [7] explores Long Short-Term-Memory (LSTM) based Recurrent Neural Network (RNN) to solve Load Forecasting in Smart Grid. Zhang et al [8] integrate macro and micro output by continuous and discrete-time series to produce multiple time series (MTS) using a recurrent neural network (RNN). Zahid et al [9] use Enhanced Convolutional Neural Network (ECNN) and Enhanced Support Vector Regression (ESVR) for electricity load and price forecasting in smart grids. Dong et al [10] intergrating convolutional neural network with K-means clustering is using for short-term load forecasting with increased scalability. Tian et al [11] integrated the hidden feature of the CNN model and the LSTM model is using to improve the forecasting accuracy in a real-world case. Song et al [12] presented a method for short-term load forecasting based on wavelet transform, extreme learning machine (ELM), and partial least squares regression to improve forecasting performance, a wavelet-based ensemble strategy is introduced into the forecasting model. Sami et al [13] suggested extreme learning machine (ELM) and regular ELM (RELM) which are used to forecast electrical loads. Teo et al [14] was presented forecasting of photovoltaic power using Artificial Neural Network (ANN) which improved using Extreme Learning Machine (ELM).

Salkuti et al [15] presented the radial basis function neural networks for solving the short-term load forecasting problem. The results are tested on Pennsylvania–New Jersey–Maryland (PJM) interconnection. Cecati et al [16] presented radial basis function (RBF) networks for 24-h electric load forecasting.

In previous studies, the authors have conducted load forecasting research using CFBNN [17] which is compared with the FFBNN method. It was concluded that the CFBNN method was better than FFBNN with epoch variations and time-training. The weakness of this research is the non-optimal weighting value which has a random initial value. In this paper, the research is conducting long-term load forecasting by optimizing Feed Forward Backpropagation Neural Networks (FFBNN) using Teaching-Learning-Based Optimization Algorithms (TLBO-FFBNN). The advantages of TLBO algorithm are the lack of algorithm parameters to run it apart from population size and maximum number of iterations. TLBO algorithm is also easy to apply and does not require computational memory [18]. The study is comparing several methods, namely, Feed Forward Backpropagation Neural Network (FFBNN), Cascade Forward Backpropagation Neural Network (CFBNN), and Feed Forward Backpropagation Neural Network optimization based on Teaching-Learning-Based Optimization Algorithms (TLBO-FFBNN). The paper also provides a comparative analysis of different methods such as a number of hidden layers and training functions.

This paper consists of the following: the second part provides a complete study of a brief description of neural networks, the Teaching-Learning-Based Optimization Algorithm method, and the Proposed TLBO-FFBNN Algorithm. Section 3 presents the results and performance comparisons. In the last part is the conclusion.

Table 1. Symbol list of FFBNN

Parameter	Function
a_1	Output from X_n
B_1	Bias layer 1
W_{ij}	The weights from X_n
X_n	Input Data
a_2	Output from activation 1
B_2	Bias layer 2
W_{jk}	The weights from Hidden layer
δ_i	Weight control factors of hidden layers
α	Learning Rate
δ_k	Weight control factors of output layer

2. Literature review

2.1 A feed forward backpropagation neural network (FFBNN)

Artificial neural network is a computing system that duplicates the skill of the brain. It has a biological process through the use of interconnections with regular artificial neurons [19]. Technically, each neuron is connected to the other by a direct link. Each link is associated with a weight that represents the information used by the network to solve the problem. FFBNN method has been widely applied.

This algorithm is also widely used in control applications because the training process is based on a simple interconnection, namely: if the output gives the wrong result, the weight is corrected so that the error can be minimized. The NN response is expected to be closer to the correct value. FFBNN is also capable of improving the weighing on hidden layers. FFBNN structure can be seen in Fig 1.

In Feed Forward Stage, Input data (X_n) will be processed in hidden layer. Input data is multiplied by the weighting then summed with bias 1 and bias 2. The result will be applied to the existing activation function in the hidden layer. The Feed Forward Stage equation can be formulated as follows:

$$a_1(t) = \sum_{i=1}^{j} W_{ij} X_n(t) + B_1 \qquad (1)$$

$$a_2(t) = f\big(a_1(t)\big) = \frac{1}{1+exp^{a_1}} \qquad (2)$$

$$a_3(t) = \sum_{j=1}^{k} W_{jk} a_2(t) + B_2 \qquad (3)$$

$$a_4(t) = f\big(a_3(t)\big) = \frac{1}{1+exp^{a_3}} \qquad (4)$$

In Backpropagation Algorithm Steps, each output neuron receives a target pattern that corresponds to training input patterns. The error information calculation is obtained by multiplying by the derivation of the activation function

$$\delta_k = (t_i - a_4)f'(a_3) \qquad (5)$$

The weight correction used to repair the next W_{jk},

$$\Delta W_{jk} = \alpha \cdot \delta_k \cdot a_2 \qquad (6)$$

Where α is Learning Rate, which regularly ranges from 0.1 to 0.5. Learning rate parameters highly affect the sharply of the training process. Each weight (W_{jk}) connected to the output unit with a hidden unit multiplied with δ_k.

$$\delta_j = \sum_{j=1}^{k} \delta_k \cdot W_{jk} \qquad (7)$$

The error is obtained by multiplying δ_in_l by the derivative of the activation function

$$\delta_i = \delta_j f'(a_1) \qquad (8)$$

The weight correction used to correct W_{jk}

$$\Delta W_{ij} = \alpha \cdot \delta_i \cdot X_i \qquad (9)$$

Each output neuron improves weight (W_{jk})

$$W_{jk}(new) = W_{jk}(old) + \Delta W_{jk} \qquad (10)$$

Each hidden layer neuron improves weight (W_{jk})

$$W_{ij}(new) = W_{ij}(old) + \Delta W_{ij} \qquad (11)$$

The description of the FFBNN symbol is explained in Table 1

2.2 Teaching-learning-based optimization algorithm (TLBO)

Teaching-Learning-Based Optimization (TLBO) is an algorithm that duplicates the learning method carried out by a teacher that affects the output of students in the classroom. The output is reflected in the results or values. A teacher is mostly regarded as a great knowledgeable person who contributions knowledge to students [20].

TLBO is also inspired by natural organic life like the other methods. TLBO has the characteristic of using population solutions to move on to global solutions. On TLBO, different design variables will be modeled with different subjects that will be

Table 2. Symbol list TLBO

Parameter	Function
S_i	The row vector of student
$S_{teacher}$	The Teacher
T_f	Teaching factor
M	the mean knowledge level of all student
α	An uniform distributed random number between 0 and 1
NP	The population size
S_j	The student is selected randomly.
$f(S_i)$	the fitness value of S_i
$f(S_j)$	the fitness value of S_j

offered to students. And learning outcomes of students are called "fitness". A teacher can be considered the best solution obtained so far. TLBO has two parts. The first part is the Teaching Phase and the second part is the Learning Phase. The Teaching Phase is learning from the teacher and the Learning Phase is learning through interaction between students.

Teaching phase

This algorithm phase is the mathematics model of student learning conducted by the teacher. A teacher transmits knowledge among students. The teacher also tries to increase the grade average of the class. The best students in the population are assumed to be teachers. The teacher will do maximum effort to increase knowledge from all classes, but students will get knowledge in accordance with the quality of teaching delivered by the teacher and the quality of students in the class [21].

In TLBO, The class An increase in the subject's average grade point and the difference in the teacher's grade point and the subject's average grade point was assessed as

$$S_i^{new} = S_i + \alpha \cdot \left(S_{teacher} - T_f \cdot M\right) \quad (12)$$

$$M = \frac{1}{NP}\sum_{i=1}^{NP} S_i \quad (13)$$

$$T_f = round(1 + rand(1,0)) \quad (14)$$

Learning phase

Students interact randomly with other students by group discussions, presentations, formal communication, etc. Students learn something new if other students have more knowledge than themselves [20]. Steps of learning phase are as follows

$$S_i^{new} = S_i + \alpha \cdot \left(S_i - S_j\right), if\ f(S_i) < f(S_j) \quad (15)$$

$$S_i^{new} = S_i + \alpha \cdot \left(S_j - S_i\right) \quad (16)$$

TLBO symbols are explained in Table 2

2.3 The proposed TLBO-FFBNN algorithm

In Table 3, it can be seen the economic data. The data used are data with a span of 2000-2008. Electricity load forecasting data from the electricity company is a reference. The data starts from 2009 to 2019. It is function as a target. In the second step of this research is to use these data into the TLBO-FFBNN algorithm. Input parameters used as input. Meanwhile, the estimated burden from the electricity company is used as a target. The results will be skilled by entering the desired algorithm. The proposed research model for the TLBO-FFBNN can be seen in Fig. 3.

The proposed TLBO is to train FFBNN. This is to overcome the weaknesses possessed by FFBNN such as errors at the beginning, initial weight values, and several other parameters. Therefore, the TLBO algorithm is used to find the optimal weights in the FFBNN training process. The FFBNN weight will be updated continuously until the desired error value is reached.

The training on the TLBO method is programmed by entering the algorithm with the maximum number of iterations of 1000 and the population size is 50 on all models. At the beginning of training, the weighted value of the FFBNN is set with a random initial value. The training process is repeated by executing input vectors into the neural network and updating the network each time until the desired criteria are met. The study consisted of 3 steps that were used to predict the electrical load, namely entering the dataset, training using the TLBO-FFBNN method, and validating it by making comparisons with other methods. The flow chart can be seen in Fig. 4.

TLBO - FFBNN is proposed to start with data processing. Then, the process of initializing the population of students and FFBNN weighting units randomly. Weighting in the population is initialized randomly between values -1 to 1, and they will act as potential candidates for the FFBNN weighting for the classification of individual datasets. The aim is to get an optimal weighting for the FFBNN network which has better classification accuracy. Individual weighting units are set to FFBNN, and networks are trained with particular dataset. This is depends on the output obtained by the network and the results of the target, the network error is calculated. The data in this study was used economic data, population, and electricity as input parameters. This is data from 2000 to 2008.
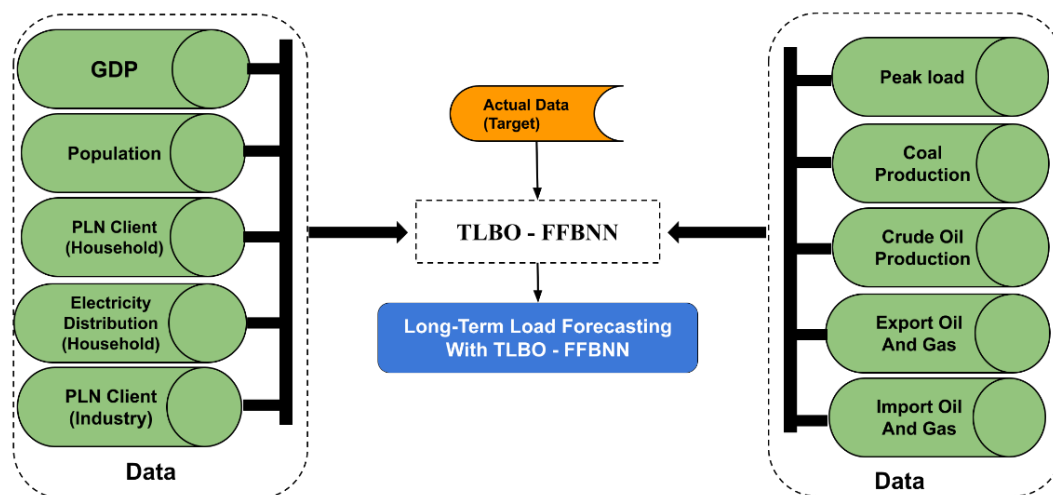
Figure. 3 The proposed forecasting using TLBO-FFBNN

Table 3. Data economic for forecasting [17]

Data / Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	Unit
GDP	195648	6541078	7267470	7942705	10610081	12675545	15029733	17851380	216784	100 K (Rp)
Population	203256	206100	212003	215276	217854	219852	222192	225642	228523.3	10 Juta
PLN Client (Household)	26796675	27905482	28903325	29997554	31095970	32174922	33118262	34684540	36025071	User
PLN Client (Industry)	44337	46021	46824	46818	46520	46475	46366	46818	47536	User
Electricity Distribution (Household)	30563319	33361591	34393967	35753158	38591235	41184272	43753223	47324905	50184187	MW
Electricity Distribution (Industry)	34014123	35600196	37180353	36494853	40328206	42448363	43615446	45802511	47968859	MW
Coal Production	434368	432588	351949	339100	354351	341202	313037	305137	314221	TON
Crude Oil Production	67105	71072	105539	113525	128479	149665	162294	188663	178930	Barel
Export Oil And Gas	62124	56320	57158	61058	71584	85659	100798	114100	137020	US $
Import Oil And Gas	33514	30962	31288	32550	46524	57700	61065	74473	129197	US $
Peak load	20850	21052	21114	21207	21459	22515	24846	25212	25574	MW

The results are explored using Mean Absolute Percentage Error (MAPE) methods. MAPE is a measure of relative accuracy used to determine the percentage of deviation from forecasting results. The testing has the good achievement if the MAPE percentage is below 10%. MAPE (%) has a formula as follows

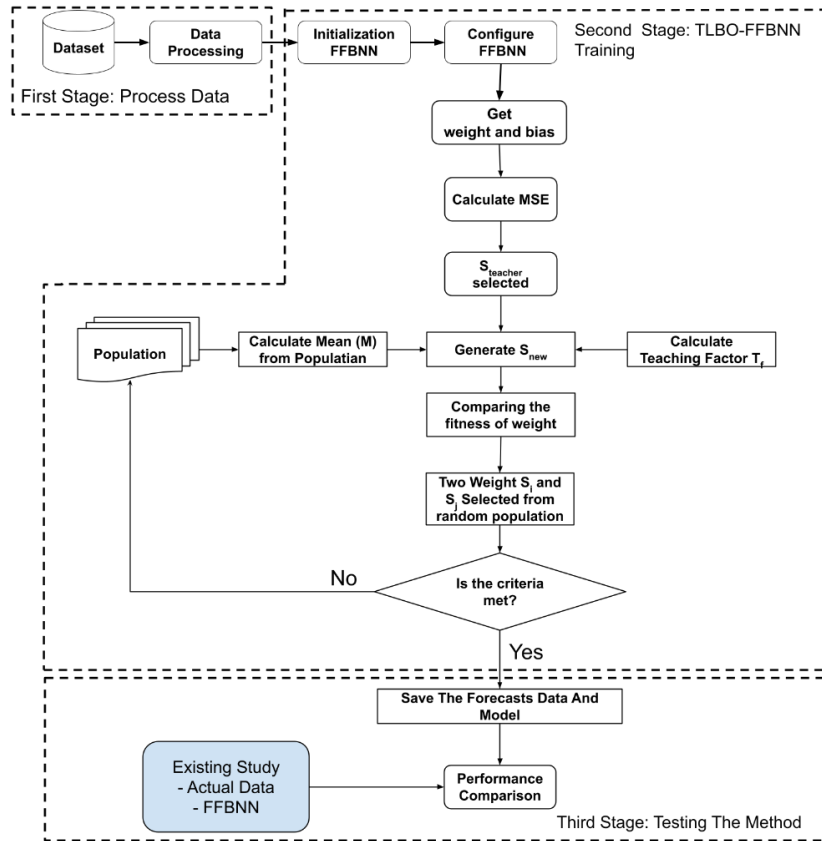$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \frac{|R_i - F_i|}{R_i} x100 \qquad (17)$$

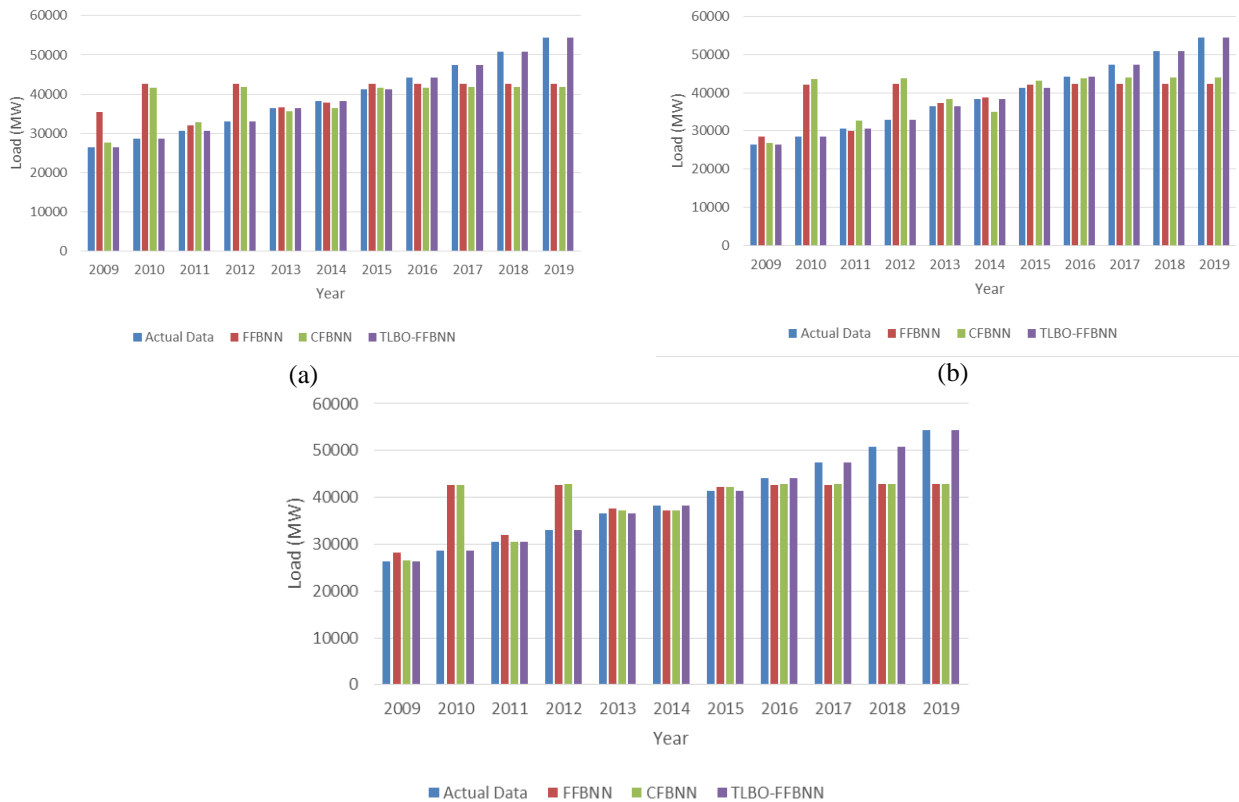Figure. 4 Flowchart of TLBO-FFBNN

(a)

(b)

(c)

Figure. 5 Chart of load electricity forecasting with bayesian regularization: (a) 4 hidden layer, (b) 8 hidden layer, and (c) 10 hidden layer

Table 4. MAPE of load electricity forecasting with bayesian regularization

Layer	MAPE FFBNN (%) [17]	MAPE CFBNN (%) [17]	MAPE TLBO-FFBNN (%)
4 Hidden Layer	17.336557	15.093320	0.000001
8 Hidden Layer	14.478653	15.238102	0.000004
10 Hidden Layer	14.843100	13.609991	0.000003

Table 5. MAPE of load electricity forecasting with resilient

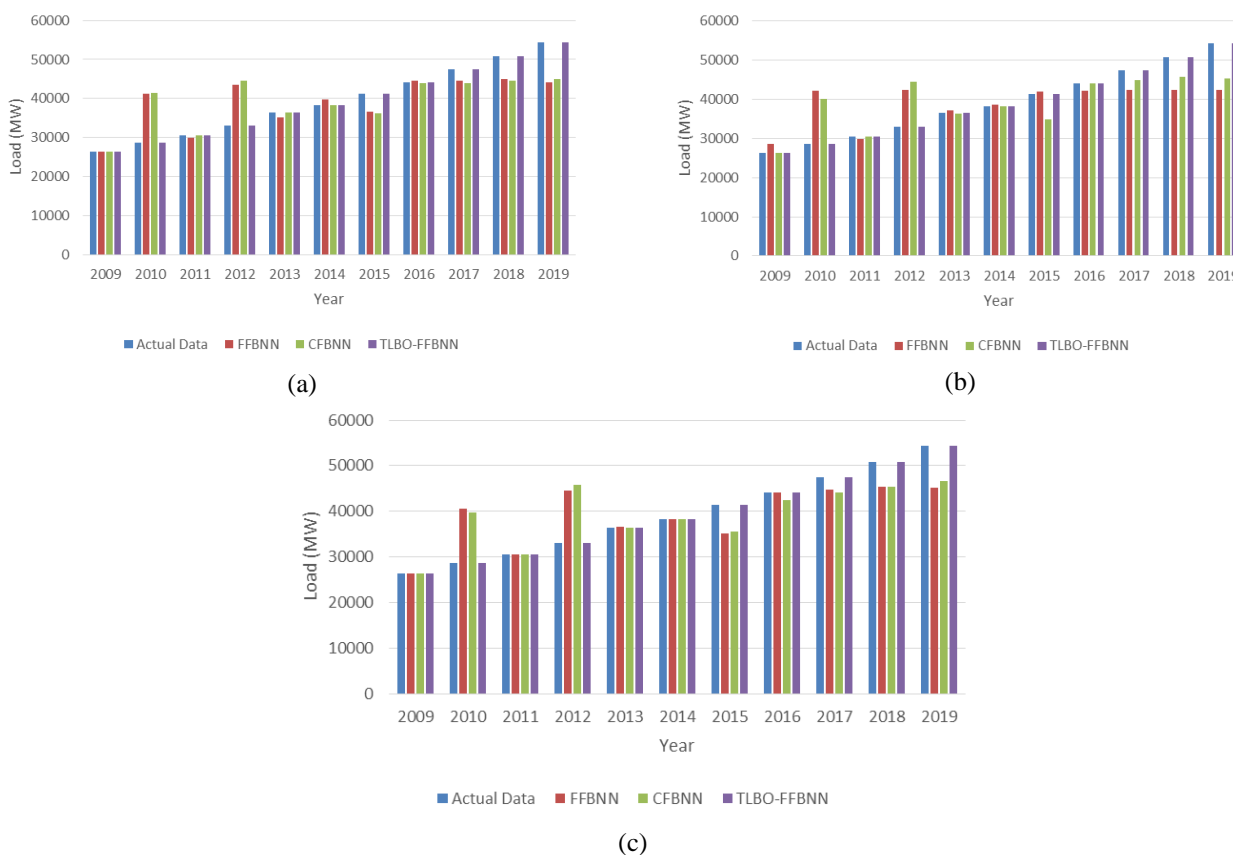| Layer	MAPE FFBNN (%) [17]	MAPE CFBNN (%) [17]	MAPE TLBO-FFBNN (%)
4 Hidden Layer	13.469493	12.928413	0.000013
8 Hidden Layer	14.478653	12.418553	0.000031
10 Hidden Layer	12.565948	12.795088	0.000022

Figure. 6 Chart of load electricity forecasting with resilient:
(a) 4 hidden layer, (b) 8 hidden layer, and (c) 10 hidden layer

Where is N the amount of data, R_i describes the actual load, and F_i describes the forecast load.

3.   Results and discussion

In the future, the usage of data is very important in all fields. Technological advances are developing rapidly, this makes data an important role. When the data can already be integrated, this makes it easy to group and process for further analysis

In this part, the paper are presenting the estimated result of the proposed algorithm. In the beginning, the working is processing and clustering the dataset. The data is covering the years 2009 until 2019. Performance measurement for forecast is used the MAPE method. The data is processed using 3 neural network methods, 3 model types of neural network training functions, and 3 hidden layer variations. This is to measure the performance of the proposed method. Neural network methods used are FFBNN, CFBNN, and TLBO-FFBNN. The training function model used is Bayesian regularization, Resilient and

18

Table 6. MAPE of load electricity forecasting with Levenberg-Marquardt

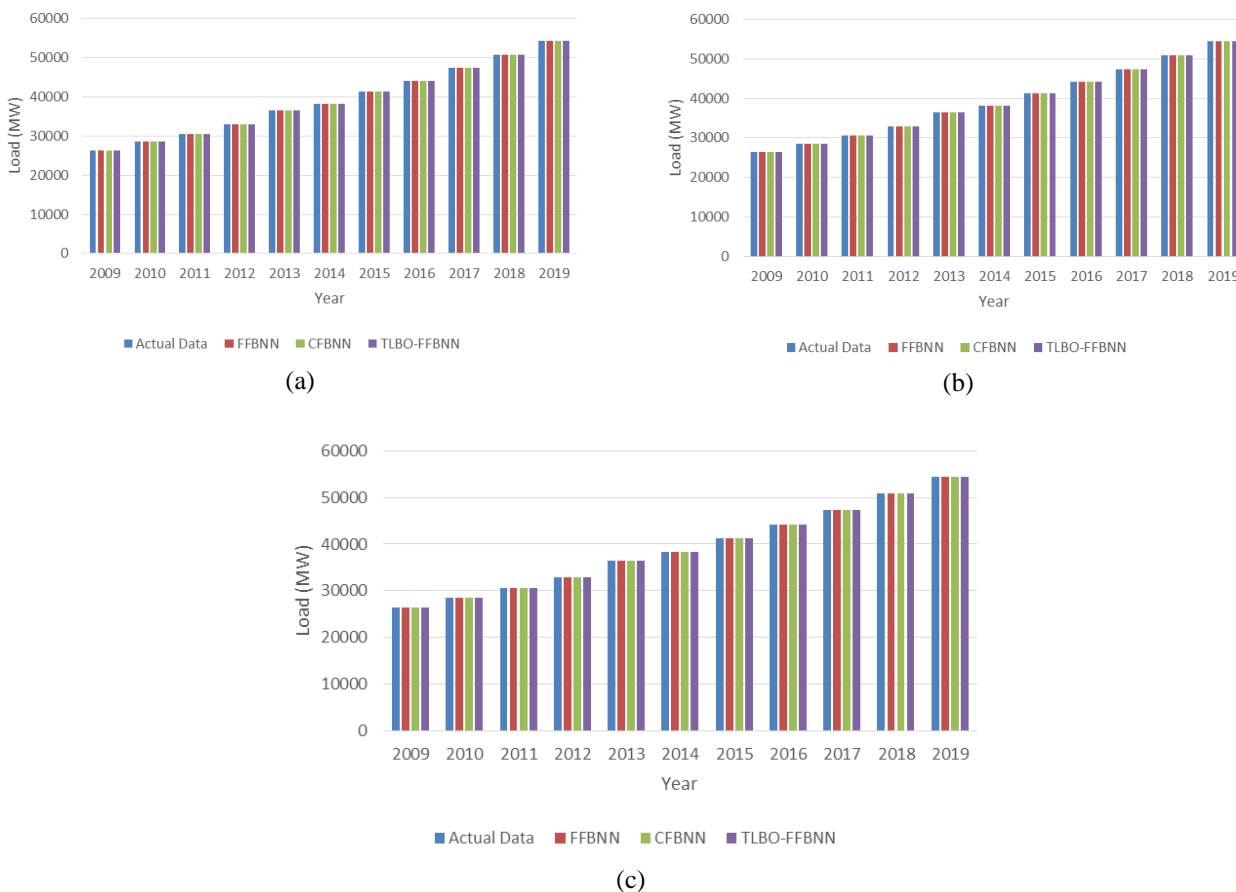| Layer	MAPE FFBNN (%) [17]	MAPE CFBNN (%) [17]	MAPE TLBO-FFBNN (%)
4 Hidden Layer	0.00000027	0.00000593	0.00001176
8 Hidden Layer	0.00000027	0.00000593	0.00003158
10 Hidden Layer	0.00000033	0.00000064	0.00003074

(a)

(b)

(c)

Figure 7. Chart of load electricity forecasting with Levenberg-Marquardt:
(a) 4 hidden layer, (b) 8 hidden layer, and (c) 10 hidden layer

Levenberg-Marquardt backpropagation. The hidden layer variations are 4, 8 and 10.

In the first experiment, the data were tested using the Bayesian regularization training function. Each data is grouped and tested in stages with hidden layer variations. The test results can be seen in Fig. 5. The MAPE results are shown in Table 4. In the first experiment, the highest MAPE value is using of the FFBNN method with hidden layer 4. It has a value of 17.336%.

The second experiment, testing data is using the resilient training function. The results of the training can be seen in Fig. 6. The results of MAPE can be seen in Table 5. In this experiment, the highest MAPE value was 14.48%. This value is obtained when using the FFBNN method with hidden layer 8.

The last experiment was using the Levenberg-Marquardt training function. The results of the training is obtained a very small average MAPE value for each method. The highest MAPE value is 0.00003158%. This is using the TLBO-FFBNN method and hidden layer 8. The results of the training can be seen in Fig. 7. While the MAPE values are shown in Table 6.

4. Conclusion

In this study, the TLBO-FFBNN algorithm is used to make long-term load predictions. Experiments using various training functions and hidden layers. In the first experiment, it was found that the average MAPE value of the proposed method was 0.0000024988%. This value is the lowest compared to the FFBNN and CFBNN methods. The average MAPE value of the TLBO-FFBNN method

in the second experiment was 0.000022% and became the lowest value. The third experiment using the Levenberg-Marquardt training function, the average MAPE value of the method used has a very small value and almost the same value. The lowest value is owned by the FFBNN method, which is 0.00000087. In this study, which used 3 experiments with variations of the training function and hidden layer, the average MAPE values of FFBNN, CFBNN and TLBO-FFBNN were 29.06%, 27.36% and 0.000004936%. From the research, it was concluded that the TLBO-FFBNN method was very adaptive and had a very good average MAPE value.

Conflicts of interest

The author declare no conflict of interest

Author contributions

Widi Aribowo is contributed to resource data and analysis data, implementation of algorithms, the conduct of experiments, and the formation of the paper.

References

[1] W. Aribowo, B. Suprianto, I. G. P. Asto Buditjahjanto, M. Widyartono, and M. Rohman, "An Improved Neural Network Based on Parasitism – Predation Algorithm for an Automatic Voltage Regulator", *ECTI-EEC*, Vol. 19, No. 2, pp. 136-144, 2021.

[2] M. Massaoudi, S. S. Refaat, I. Chihi, M. Trabelsi, F. S. Oueslati, and H. Abu-Rub, "A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for short-term load forecasting", *Energy*, Vol. 214, p. 118874, 2021.

[3] H. Bo, Y. Nie and J. Wang, "Electric Load Forecasting Use a Novelty Hybrid Model on the Basic of Data Preprocessing Technique and Multi-Objective Optimization Algorithm", *IEEE Access*, Vol. 8, pp. 13858-13874, 2020, doi: 10.1109/ACCESS.2020.2966641.

[4] A. Zainab, D. Syed, A. Ghryeb, H. Abu-Rub, S.S. Refaat, M. Houchati, O. Bouhali, and S.B. lopez, "A Multiprocessing-Based Sensitivity Analysis of Machine Learning Algorithms for Load Forecasting of Electric Power Distribution System", *IEEE Access*, Vol. 9, pp. 31684-31694, 2021, doi: 10.1109/ACCESS.2021.3059730.

[5] T. Li, Z. Qian, and T. He, "Short-term load forecasting with improved CEEMDAN and GWO-based multiple kernel ELM", *Complexity*, Vol. 2020, Art. no. 1209547, 2020,

[6] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network", *IEEE Trans. Smart Grid*, pp. 1–8, 2017.

[7] J. Zheng, C. Xu, Z. Zhang, and X. Li, "Electric load forecasting in smart grids using Long-Short-Term-Memory based recurrent neural network", *in Proc. 51st Annu. Conf. Inf. Sci. Syst. (CISS)*, pp. 1–6, 2017.

[8] B. Zhang, J.-L. Wu, and P.-C. Chang, "A multiple time series-based recurrent neural network for short-term load forecasting", *Soft Comput.*, Vol. 22, No. 12, pp. 4099–4112, 2018.

[9] M. Zahid, F. Ahmed, N. Javaid, R. Abbasi, H. Zainab Kazmi, A. Javaid, M. Bilal, M. Akbar, and M. Ilahi, "Electricity price and load forecasting using enhanced convolutional neural network and enhanced support vector regression in smart grids", *Electronics*, Vol. 8, No. 2, p. 122, 2019

[10] X. S. Dong, L. J. Qian, and L. Huang, "Short-term load forecasting in smart grid: a combined CNN and K-means clustering approach", In: *Proc. of IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 119-125, 13–16, 2017.

[11] C. Tian, J. Ma, C. Zhang, and P. Zhan, "A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network", *Energies*, Vol. 11, No. 12, pp. 1–13, 2018.

[12] Li, Song, L. Goel, and P. Wang, "An ensemble approach for short-term load forecasting by extreme learning machine", *Applied Energy*, Vol. 170, pp. 22-29, 2016.

[13] S. Ekici, "Electric load forecasting using regularized extreme learning machines", *International Journal of Industrial Electronics and Electrical Engineering*, Vol. 4, No. 6, pp. 119-122, 2016.

[14] T. T. Teo, T. Logenthiran, and W. L. Woo, "Forecasting of photovoltaic power using extreme learning machine", In: *Proc. of Smart Grid Technologies - Asia (ISGT ASIA), 2015 IEEE Innovative*, pp. 1–6, 2015.

[15] S. R. Salkuti, "Short-term electrical load forecasting using radial basis function neural networks considering weather factors", *Electrical Engineering*, Vol. 100, No. 3, pp. 1985-1995, 2018.

[16] C. Cecati, J. Kolbusz, P. Różycki, P. Siano, and B. M. Wilamowski, "A novel RBF training algorithm for short-term electric load forecasting and comparative studies", *IEEE*

20

Trans. Ind. Electron., Vol. 62, No. 10, pp. 6519–6529, 2015.

[17] W. Aribowo, S. Muslim, B. Suprianto, S. Isnur Haryudo, and Joko, "Tunicate Swarm Algorithm-Neural Network for Adaptive Power System Stabilizer Parameter", *Sci Tech Asia*, Vol. 26, No. 3, pp. 50-63, 2021.

[18] B. Mohanty and S. Tripathy, "A teaching learning based optimization technique for optimal location and size of DG in distribution network", *Journal of Electrical Systems and Information Technology*, Vol. 3, No. 1, pp. 33-44, 2016.

[19] W. Aribowo, S. Muslim, B. Suprianto, S. Isnur Haryudo, and A. C. Hermawan, "Intelligent Control of Power System Stabilizer Based on Archimedes Optimization Algorithm–Feed Forward Neural Network", *International Journal of Intelligent Engineering and Systems,* Vol. 14, No. 3, pp. 43-53, 2021.

[20] R. V. Rao, V. J. Savsani, and D. P. Vakharia, "Teaching–learning-based optimization: A novel method for constrained mechanical design optimization problems", *Computer-Aided Design*, Vol. 43, No. 3, pp 303-315, 2011.

[21] P. J. Pawar and R. V. Rao. "Parameter optimization of machining processes using teaching–learning-based optimization algorithm", *The International Journal of Advanced Manufacturing Technology*, Vol. 67, pp. 995-1006, 2013.