

Applying Neural Network And Levenberg - Marquardt Algorithm for Load Forecasting in IA-Grai District, Gia Lai Province

Quang Huy Giap*, Dang Luu Nguyen, Thi Thanh Quynh Nguyen, Thi Minh Dung Tran

Abstract—Electrical energy is fundamental to the life and economic development. Accompany by the increasing demand for electricity by people and businesses, the electricity industry must always ensure sufficient supply for production activities. This poses an urgent need to have an accurate load forecasting tool for electricity production and distribution strategy. If the demand forecast is lower than the reality, the phenomenon of power supply shortage will occur. While if the forecast exceeds the demand for use by people and businesses, it will cause economic losses because it must mobilize power sources by using a backup as a diesel generator. For these reasons, the problem of electrical load forecasting becomes very important in terms of efficient power supply. In this paper, a load forecasting tool using Neural Network and Levenberg - Marquardt algorithm is presented.

Index Terms—Neural network; Levenberg-Marquardt algorithm; electrical load forecasting.



1. Introduction

IN the context of a rapidly increasing demand for electricity, due to the increasing demand of businesses and households, Vietnam's power system is facing the risk of supply shortage [1]. On the other hand, the high temperature for many months combined with the development of people's living standards makes the task of balancing electric energy increasingly difficult. This poses a problem of load forecasting to optimize the production and distribution of electricity. To balance power systems, an accurate load forecasting tool is essential for a wide range of applications: short-term forecasting for power distribution planning, medium-term forecasting for maintenance planning, and Long-term forecasting for planning system expansion to match demand [2]. Predicting power demand requires the implementation of an advanced forecasting technique.

In the past time, load forecasting has received much research attention. In particular, the method of using artificial neural networks is especially noticed [3], [4]. Because the neural network can process and learn the associations in the input data. In this method, many different training algorithms are studied to improve the

prediction results (increasing prediction accuracy). One of the algorithms that are highly appreciated for its accuracy and often applied is the Levenberg-Marquardt algorithm [5]. But with the characteristics of this algorithm, the probability of convergence at the local extreme with an error close to the global extreme is very low. This article proposes a solution to apply Neural Network with Levenberg - marquardt algorithm in load forecasting in IA-Grai district, Gia Lai province by considering the appropriate input data and the convergence of Levenberg - Marquardt algorithm to improve the accuracy of load forecasting.

2. Overview of load forecasting in IA-GRAI District, GIA LAI Province

The nature of the load forecasting problem is very suitable for forecasting technology using artificial neural networks (ANNs) because it is capable of simulating complex nonlinear correlations through a training process on input data set.

The objective of this paper is to develop an accurate medium-term load forecasting tool for the Ia-Grai district, Gia Lai province. The forecasting of electricity production in the Ia-Grai district is for the following purposes [6] : To provide appropriate operation plans and power outages scheduled for maintenance and repair; to support technical management with necessary measures to reduce power loss; To calculate and design accurately to make investment plans at the right time and effectively; To provide a forecast of electricity output to ensure the power source for socio-economic development and people's daily life.

An important part of the problem is determining the input data capable of predicting the output data (load

Quang Huy Giap is with the University of Danang - University of Science and Technology, Danang, Vietnam (e-mail: gqhuy@dut.udn.vn).

Dang Luu Nguyen is with the FPT Group (e-mail: dangluuqthw@gmail.com).

Thi Thanh Quynh Nguyen is with the University of Danang - University of Science and Technology, Danang, Vietnam (e-mail: nt-tquynh@dut.udn.vn).

Thi Minh Dung Tran is with the University of Danang - University of Science and Technology, Danang, Vietnam (e-mail: ttmdung@dut.udn.vn).

*Corresponding author: Quang Huy Giap (e-mail: gqhuy@dut.udn.vn)
Manuscript received May 03, 2022; revised May 29, 2022; accepted June 20, 2022.

Digital Object Identifier 10.31130/ud-jst.2022.240ICT

by day). This relationship depends largely on the characteristics of each unit in the forecasted area. Because Ia Grai is a district where most of the area is used for farming purposes, a large percentage of the load is used for irrigation purposes for agricultural products (coffee, pepper, rubber tree, etc.). In reality, irrigation activities for agriculture will significantly affect the electricity load, it is greatly influenced by weather conditions such as rainfall, temperature, humidity, etc. For this reason, the proposed forecasting model uses the following inputs: daily rainfall, daily maximum temperature, daily minimum temperature, day and month.

3. Application of neural network for load forecasting

3.1. Introduction to artificial neural networks

The nature of electrical load data makes them difficult to accurately represent and forecast based on simple calculations. Today, the development of artificial intelligence allows us to create a system that can train itself on an existing data set. The application of artificial neural networks (ANNs) [7] in the electricity industry is increasingly popular. ANNs were developed to mimic the human nervous system for example-based learning. Neural networks learn based on the relationship between input and output variables. In general, ANNs are composed of a combination of many basic elements called neurons across multiple layers.

The model in this paper is built using Matlab Neural Networks Toolbox. It is trained on historical input data to determine optimal values of weights and biases, allowing it to predict output consumption load values. The input data set is taken from the DMS telemetry program on feeder 471/110 Dien Hong from January 1, 2017, to September 30, 2020, and weather data is obtained from the Hydro-meteorological Center. Input data includes 6 aspects: day, month, the highest temperature of the day, the lowest temperature of the day, precipitation of the day, and an average temperature of the day. The back-propagation algorithm used is Lavenberg-Marquardt, with fast training and high convergence to find the optimal structure of prediction accuracy.

3.2. Build the training algorithm

The selected neural network structure has a 4-layer structure: 1 input layer, 1 output layer, and 2 hidden layers shown in Fig. 1.

The most important problem of load forecasting using an artificial neural network is how to train the network to give accurate forecast output. By adjusting the weights and biases parameters, the model will be improved over many parameter updates. The algorithm chosen for use in this prediction model is Levenberg-Marquardt. Levenberg-Marquardt algorithm flowchart is shown in Fig. 2.

The Levenberg-Marquardt algorithm [8], developed independently by Kenneth Levenberg and Donald Marquardt, is a combination of steepest descent and the Gauss-Newton algorithm. It allows for minimizing the

error function of a nonlinear function. It has the fast convergence of the Gauss-Newton algorithm and the stability of the Steepest gradient.

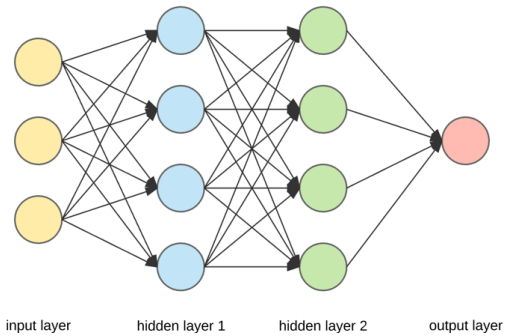


Fig. 1: Diagram of an artificial neural network with 2 hidden layers

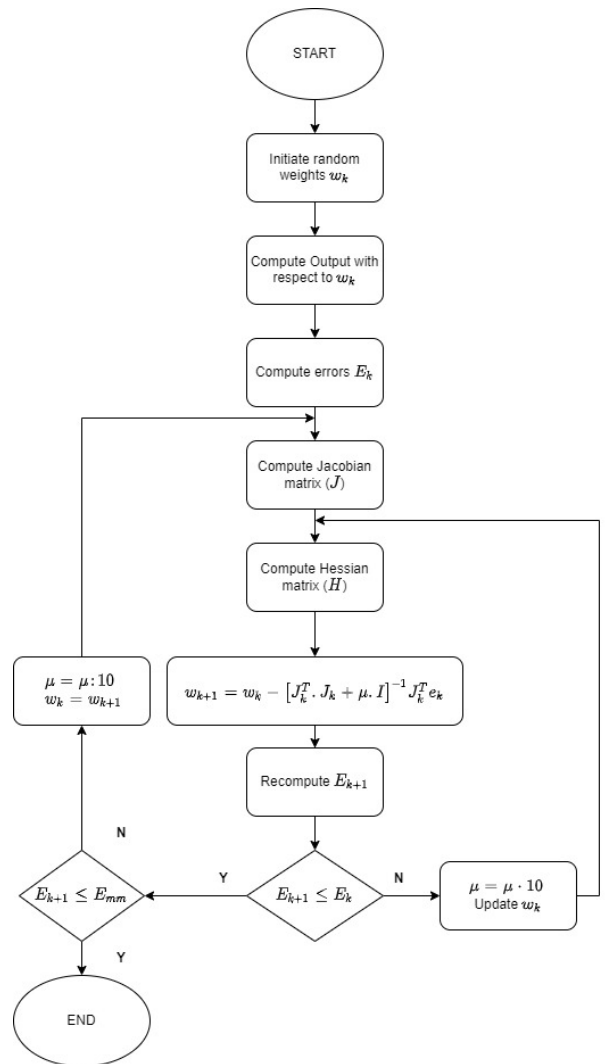


Fig. 2: Levenberg-Marquardt algorithm

- The variance of the network is presented as:

$$E(x_k) = \sum_{k=1}^m [e(e_k)]^2 = \sum_{k=1}^m [d_k - z_k] \quad (1)$$

In which:

- x_k is the input data vector.

- d_k is the training data $\{d_k; 1 \leq k \leq m\}$.
- z_k is forecast data $\{z_k; 1 \leq k \leq m\}$
- Error: $e(x_k) = d_k - z_k$

- Taylor expansion of a function $f(x)$ in the neighborhood of Δx_k

$$\begin{aligned} E(x_{k+1}) &= E(x_k + \Delta x_k) \\ &\approx E(x_k) + G(x_k)\Delta x_k + \frac{1}{2}\Delta x_k H(x_k)\Delta x_k \end{aligned} \quad (2)$$

In which:

- $\Delta x_k = x_{k+1} - x_k$
- $G(x_k)$: Gradient of function $E(x)$
- $H(x_k)$: Hessian matrix of $E(x)$

- The partial derivative of equation 2 with respect to Δx_k will be:

$$\begin{aligned} G(x_k)\Delta x_k + H(x_k)\Delta x_k &= 0. \\ \iff \Delta x_k &= -H(x_k)^{-1}G(x_k) \end{aligned} \quad (3)$$

- Gradient and Hessian matrix of function $E(x)$ can be rewritten as:

$$G(x_k) = 2J^T(x_k)E(x_k). \quad (4)$$

$$H(x_k) = 2J^T(x_k)J(x_k) + 2S(x_k) \quad (5)$$

In which:

- $J(x_k)$: Jacobian matrix
- $S(x_k) = \sum_{i=1}^m e_i(x_k)\nabla^2 e_i(x_k)$

- Considering the value of $S(x_k)$ is so small, then we can approximate the Hessian matrix to:

$$H(x_k) \approx 2J^T(x_k)J(x_k) \quad (6)$$

- Substituting eq.3 and eq.5 into eq.3.2 we get:

$$\Delta x_k = -[2J^T(x_k)J(x_k)]^{-1}(2J^T(x_k)E(x_k)) \quad (7)$$

- However, the limitation of this algorithm is that the Hessian matrix can be difficult or impossible to inverse. To overcome this limitation, an approximation of the Hessian matrix is used:

$$H_s(x_k) \approx H(x_k) + \mu.I \quad (8)$$

Where:

- I : Unit matrix
- μ : Coherence coefficient

- Thus, the Levenberg-Marquardt algorithm uses approximation and updates the weights for the Hessian matrix as follows:

$$w_{k+1} = w_k - [J^T.J + \mu.I]^{-1}J^T E \quad (9)$$

Where:

- J : the Jacobian matrix that contains first derivatives of the network errors with respect to the weights w and biases b

- When μ is zero, the Levenberg-Marquardt algorithm becomes Newton's algorithm. When μ is large, this algorithm becomes a Gradient Descent algorithm with a small step size. Newton's method is faster and

more accurate near an error minimum, so the aim is to shift toward Newton's method as quickly as possible. Thus, μ is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this way, the performance function is always reduced at each iteration of the algorithm for ensuring the fast convergence.

- The disadvantage of Levenberg-Marquardt algorithm is to require much memory. The size of the Jacobian matrix J is $Q \times n$, where Q is the number of training sets and n is the number of weights and biases in the network. Details about the Levenberg-Marquardt algorithm and how to reduce memory storage can be found in the Matlab manual.

4. Improvement of load forecasting accuracy

From the algorithm flowchart, we can see that if we only use the Levenberg-Marquardt algorithm to train the artificial neural network, there will be points that need to be improved. Because at the initialization of the training process, the Levenberg-Marquardt algorithm only initializes one set of weights. After the training process, the algorithm will converge to the initial starting point at some local extreme. But, in a forecasting model with many weights and biases, the nonlinear relationship between the output power and the input data is very complex, leading to the existence of many local minima. The minimum probability that the algorithm converges to a small error is very low. So, in the proposed algorithm, the number of initialization points will be increased. After applying the method of increasing the number of initial initialization points of the Levenberg-Marquardt algorithm, we have an overview algorithm diagram as shown in Fig. 3.

In the algorithm diagram above, n_{max} is the initial number of initial points of the LM algorithm. In this simulation, we choose $n_{max} = 5$. The value of n is taken from 1 to 5 corresponding to Random_num in Fig. 6

With two parameters i, j is the number of nodes of the first hidden layer and the second hidden layer, respectively. Two loops i, j will be inserted into each other. The purpose of this is to have the model run iteratively with many different structures, then select the model with the highest accuracy.

5. Achievement results

5.1. Input data and output data

The input data set is composed of the weather parameters by day, taken from the date 01/01/2017 to 30/09/2020 including:

- Maximum temperature T_{max} ($^{\circ}C$)
- Minimum temperature T_{min} ($^{\circ}C$)
- Average temperature T_{tb} ($^{\circ}C$)
- Rainfall L (mm)
- Day
- Month (Mo)
- Daily electricity load Ang (MWh) is the output parameter for the ANN model.

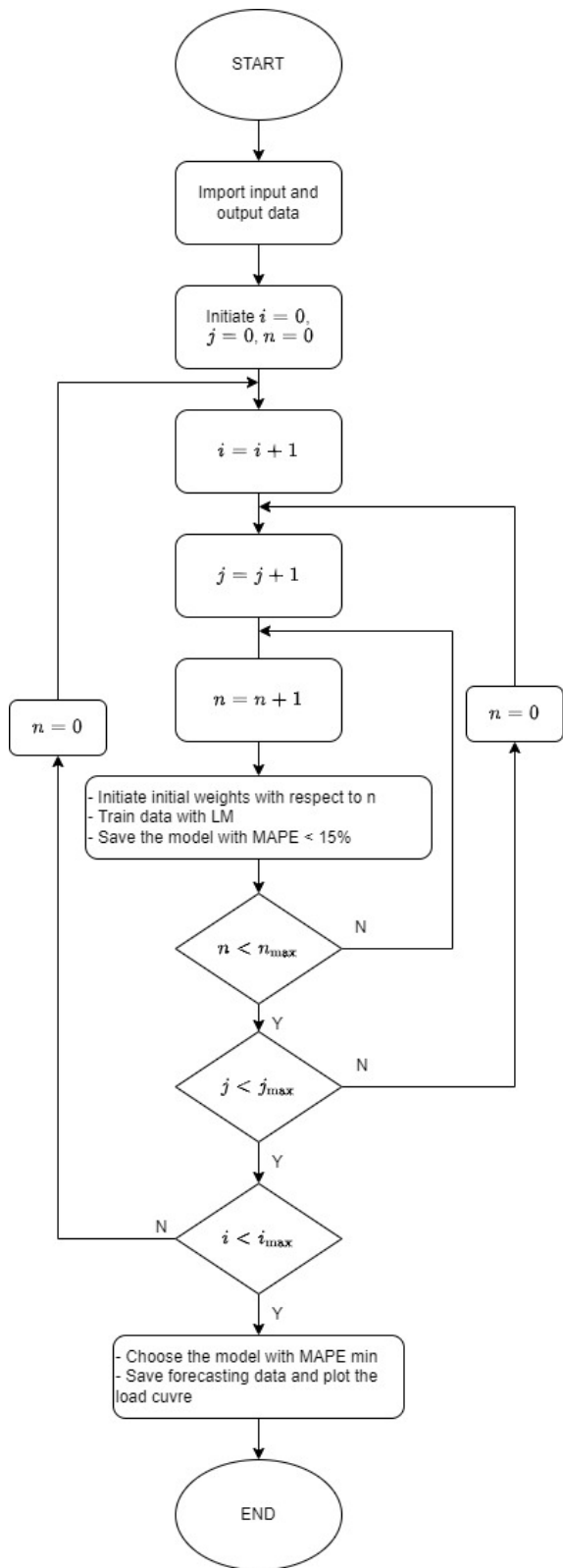


Fig. 3: Accuracy improvement procedure

The input data set is divided into 3 parts:

- Data from 01/01/2017 to 31/03/2019 and from 01/07/2019 to 31/12/2019 are used for the training stage.
- Data from 01/04/2019 to 30/06/2019 is used for

the assessment stage.

- Data from 01/01/2020 to 30/09/2020 is used for testing.

TABLE 1: Input data.

Year	Mo	Day	T_{min}	T_{tb}	T_{max}	L	Ang
2017	1	1	17.6	21.9	27.3	0.01	81.1
2017	1	2	20.1	23.2	28.1	0.01	72.9
2017	1	3	20.3	22.9	27.1	0.01	100.1
2017	1	4	18.1	22.4	27.9	0.01	83.8
2017	1	5	16.1	21.7	27.2	0.01	81.4
2017	1	6	20.3	22.8	28.2	0.01	76.8
2017	1	7	16.1	21.3	28.2	0.01	75.7
...
2020	9	26	22.1	24.0	27	12	67.9
2020	9	27	21.3	24.4	28.4	0.01	69.2
2020	9	28	21	25.1	31.9	0.01	70.3
2020	9	29	22.1	24.0	28.3	5.5	68.6
2020	9	30	22.2	24.6	29.3	0.01	70.5

5.2. Forecasting stage

The forecasting tool used here is the Matlab neural network toolbox. After configuring the input data set. The model will be trained and produce the prediction results. Models with an average percentage error of $MAPE \leq 15\%$ will be saved to an Excel file, then the model with the smallest MAPE will be selected and the parameters of that model will be stored for future forecasting. Where MAPE is Mean Absolute Percentage Error computed by:

$$MAPE = \frac{1}{N} \sum_{k=1}^N \left| \frac{d_k - z_k}{d_k} \right| \quad (10)$$

5.3. Achievement Results and comparison with case that is without the increase of initial points

When there is no method to increase the initial number of initialization points:

- + The best-achieved model:

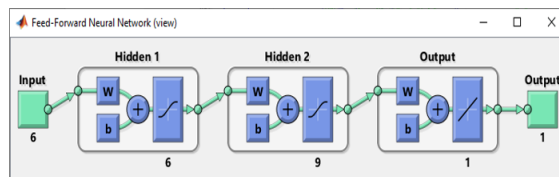


Fig. 4: The results when there is no improvement method

- Neurons in hidden layer of ANN2 model - layer 1: 6, layer 2: 9
- With MAPE test min: 11.7306 % (r: 0.92971, RMSE: 16995.7881, MAE: 12303.0641).
- With MAPE valid: 10.3671 %

When the method to increase the initial number of initialization points is applied:

- + The best achieved model:

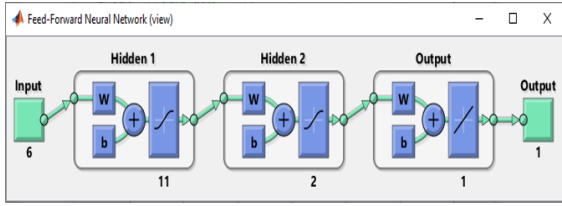


Fig. 5: After applying the improvement algorithm

TABLE 2: Comparison between with/without the increase of the number of initial points

Type of Model	Achieved model	MAPE test
Without the increase of the number of initial points	$i = 6; j = 9$	11.73%
With the increase of the number of initial points	$i = 11; j = 2$	8.513%

- Neurons in hidden layer of ANN2 model – layer 1: 11, layer 2: 2
- With MAPE test min: 8.5132% (r: 0.92861, RMSE: 15547.878, MAE: 9840.7684).
- With MAPE valid: 9.0942%

6. Conclusions

The usage of artificial neural networks and load forecasting has attracted many researchers. In this paper, the authors have proposed a solution apply Neural Network with Levenberg - marquardt algorithm in load forecasting in IA-Grai district, Gia Lai province. An improvement to increase the output accuracy of the model has been considered. The proposed solution has been proven by the achievement results. Furthermore, a comparison is also executed between 2 cases, with and without using the improvement method. As can be seen, a procedure with an improvement has been proven to be more effective when it comes to reducing the prediction error. In the future, improvement algorithms as well as an increase in the amount of input data will make the prediction tool using artificial neural networks more powerful and contribute to the management and operation of the network in the Power industry.

Acknowledgement

This work was supported by The University of Danang - University of Science and Technology, code number of Project: T2022-02-35.

References

- [1] V. C. N. Nguyen, H. T. Dung, N. S. H. Nguyen, and T. H. Nguyen, "Green technology for sustainable development: Practice and experience of renewable energy in vietnam," *Research Nepal Journal of Development Studies*, vol. 4, no. 2, pp. 79–92, 2021.
- [2] A. Azeem, I. Ismail, S. M. Jameel, and V. R. Harindran, "Electrical load forecasting models for different generation modalities: a review," *IEEE Access*, 2021.
- [3] H. H. Aly, "A proposed intelligent short-term load forecasting hybrid models of ann, wnn and kf based on clustering techniques for smart grid," *Electric Power Systems Research*, vol. 182, p. 106191, 2020.
- [4] P. Singh, P. Dwivedi, and V. Kant, "A hybrid method based on neural network and improved environmental adaptation method using controlled gaussian mutation with real parameter for short-term load forecasting," *Energy*, vol. 174, pp. 460–477, 2019.
- [5] A. Mikhaylov and S. Tarakanov, "Development of levenberg-marquardt theoretical approach for electric networks," in *Journal of Physics: Conference Series*, vol. 1515, no. 5. IOP Publishing, 2020, p. 052006.
- [6] N. Abu-Shikhah and F. Elkarmi, "Medium-term electric load forecasting using singular value decomposition," *Energy*, vol. 36, no. 7, pp. 4259–4271, 2011.
- [7] M. Saviozzi, S. Massucco, and F. Silvestro, "Implementation of advanced functionalities for distribution management systems: Load forecasting and modeling through artificial neural networks ensembles," *Electric Power Systems Research*, vol. 167, pp. 230–239, 2019.
- [8] M. T. Hagan, H. B. Demuth, and M. Beale, *Neural network design*. PWS Publishing Co., 1997.

First hidden layer	Second hidden layer	Random num	r_t	r_v	MAPE_t	MAPE_v
1	1	2	0.78097	0.284963	13.81179	13.78681
1	3	4	0.842081	0.308889	10.24411	13.11894
2	8	3	0.747374	0.080394	14.94608	14.66061
2	8	5	0.784983	0.462558	12.97045	11.89152
...
10	4	5	0.813002	0.566298	14.59138	14.21857
10	5	1	0.865962	0.554205	12.77778	13.50746
10	5	2	0.814929	0.727155	14.21789	12.45008
10	5	5	0.886586	0.639885	14.90813	14.3713
10	6	1	0.938757	0.855869	12.61866	10.69169
10	6	2	0.93213	0.771054	13.51635	13.6545
10	6	4	0.84576	0.691359	13.426	12.04777
10	10	5	0.944395	0.899493	12.70153	9.734361
11	2	1	0.893573	0.805252	14.3193	12.48623
11	2	2	0.928608	0.906218	8.513222	9.094224
11	2	3	0.936401	0.832943	12.99914	14.15411
11	2	4	0.7861	0.626073	14.02043	14.25135
11	3	5	0.946255	0.88146	12.96427	12.60703

Fig. 6: Models with MAPE <15%

And the highlighted model is the best model achieved. It has a MAPE error on the Test set of 8.513% and has 11 nodes in the first hidden layer, and 2 nodes in the second hidden layer.

5.4. A comparison between the two cases above

From the training results of two models using the same input data set, one model applies the method of increasing the number of initial initialization points for the Levenberg-Marquardt algorithm, while the other does not.

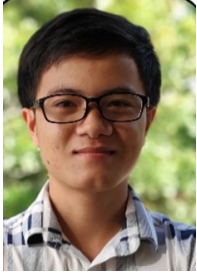
From the table above, we can see that the application of the method of increasing the initial initialization points has brought about higher accuracy prediction results, improving 3.2% of the prediction error on the Test set.



Quang Huy GIAP is a Lecturer-researcher at the University of Danang - University of Science and Technology. He received his Ph.D. in Automation- Production in 2011 from Grenoble INP, France. His research is situated in the field of Automation, with a special focus on fault detection and diagnostics technologies, load forecasting, renewable energies and energy management.



Thi Thanh Quynh Nguyen has obtained her Ph.D. in Electrical Engineering from the Doctoral School Electronics, Electrical Energy, Automatic Control, Signal Processing (EEATs), University Grenoble Alpes in 2019. She joined the automation division of the Electrical Engineering faculty at the University of Danang - University of Science and Technology in 2020 as a lecturer. Her research interests include load forecasting, distributed data management for smart grids and distributed control and management in microgrid.



Dang Luu Nguyen received his B.E. (2020) at the University of Danang - University of Science and Technology (DUT), Vietnam. He is working at FPT Group. His research interests include load forecasting, AI.



Thi Minh Dung Tran received a Specialist Diploma of Engineering in the specialty of Biotechnical systems Automation in Moscow State University of Applied Biotechnology, Russia in 2009; a Doctoral Degree in Automation and Production in Gipsa-Lab, Universities of Grenoble, France in 2015. Her research interests include Control Systems, Manufacturing Automation and Energy Management, Control theory and Optimization techniques, and Smart Building.