ICIC Express Letters
Part B: Applications
Volume **16**, Number **5**, May **2025**

ICIC International ©2025 ISSN 2185-2766
pp. **549–557**

DEEP LEARNING PREDICTING POWER COSTS FOR POWER PLANNING

Pimolrat Ounsrimung[1], Walaiphan Phuanphannid[2,*] and Supakit Nootyaskool[2]

[1]Department of Information Science
Faculty of Informatics
Mahasarakham University
Khamriang Sub-District, Kantarawichai District, Mahasarakham 44150, Thailand
pimolrat.o@msu.ac.th

[2]School of Information Technology
King Mongkut's Institute of Technology Ladkrabang
1 Chalong Krung, Ladkrabang, Bangkok 10520, Thailand
supakit@it.kmitl.ac.th
*Corresponding author: 64607059@kmitl.ac.th

Received July 2024; accepted September 2024

Abstract. *The current global situation, especially in Thailand, is showing recovery following the easing of COVID-19 restrictions. With a growing population and industrial expansion, there is an increased demand for electricity. This rise in demand, along with factors like fuel costs and fluctuating currency rates, has influenced global power prices. The authors are keen on studying power's variable price prediction using machine learning techniques. The objective is to analyze factors associated with power costs and their interrelationships. It is essential to have accurate data for strategic power planning. Research has been conducted on theories related to variable power prices, time series, traditional statistical forecasting, machine learning, and deep learning. It was found that three main factors affecting variable power prices were identified: natural gas prices, exchange rates, and inflation rates, with natural gas prices and inflation rates having a strong correlation. Traditional statistical forecasting is less efficient for highly volatile power's variable price. Deep learning models outperform conventional machine learning for this dataset.*
Keywords: Variable power prices, Ft price, Time series, Traditional forecasting method, Machine learning method, Deep learning method

1. **Introduction.** After reducing the COVID-19 situation living free, the economy in each country is growing and increasing as to the electric power needed having increment not only the public but also the industry. Factors are related to the calculation of demand for electricity such as the oil price, inflation rate, and money exchange rate. In 2021, Thailand's average exchange rate at 32 Thai baht (THB) per USD was increasing over 36THB. These factors are directly related to the increasing power cost of the electricity produced.

The Energy Regulation Commission (ERC) of Thailand planned the power needed from January to April 2023 at 67,833 million units [1], increasing by 3,724 million units when compared with the four months ago. The trends of need for electricity are usually the raised cost in the industry up to 8% to 10%. Therefore, this research is interested in the usage of techniques to predict a parameter factor inside the calculation of the electricity price, which will help the heavy industries to predict the electric price coming and make a plan of making products to reduce the cost. The research uses various techniques from traditional methods, machine learning, and deep learning.

DOI: 10.24507/icicelb.16.05.549

This research has two main objectives: 1) to find the parameters and factors such as inflation rate, exchange rate, and oil price, which are related to the power electric cost, and 2) to find a good model by comparing the deep learning model to learning the factor data in power cost. When the research finds a good model, it helps the industry to set the production plan to reduce the effect of high costs that are coming. The scope of this research is the creation and evaluation of the predicting model by using the historical data collected from 2016 to 2023.

Presentation in each section by topic in the second presented the background and theory for power cost calculation, gas price, and Gross Domestic Product (GDP) chain volume measures. The third section introduced a detail of data collection, data exploration, data preprocessing, and model design with training models. The fourth is the experiment results comparing the traditional forecasting, machine learning algorithms, and deep learning. The fifth is the conclusion giving a summary of the research result and the summarized knowledge from the research.

2. **Background and Theory.** Electrical-unit calculation in Thailand combines three components in a unit as Equation (1), E is an electrical unit in 1000W per hour, B is a basis price controlled by three electrical organizations, the Electricity Generative Authority of Thailand (EGAT), the provincial electricity authority (PEA), and the metropolitan electricity authority (MEA). The F is the fuel adjustment charge called Ft price, and T is value added tax.

$$\dot{\mathrm{E}} = \mathrm{B} + \mathrm{F} + \mathrm{T} \tag{1}$$

The cost of the electrical comes from the cost of building the electric plant, and the cost of the wire cable system. Thailand buys electricity from LAOS and private electricity companies. The assumption of power needs for a 3-to-5-year future by giving more weight to any type of electricity supply is more important. The factor that mentioned makes a buying contact in the long times. Therefore, B, or the basis price has a stable cost. A factor or attribute for the electric cost in a short time is the F that comes from the oil price. We know the oil price sometimes changes every day, and the government organizer takes a substitution when the oil price moves up.

The F changing up or down depends on the fuel and gas having a ratio in the electricity generator at 70 to 80% of which 50% is the natural gas [2]. The EGAT purchases the natural gas from two source types, independent power producers (IPP), and small power producers (SPP). Additionally, in the personnel homes with PV solar systems, the government allowed the electricity to be sold PEA or MEA. The mentioned above comprehend many sources and factors related to the calculation of an electricity unit. Economic growth associated with the GDP-chain volume measures (GDP-CVM) [3] and the inflation rate being parameters should be considered in the model. The CVM measures the products and services given to customers, by calculating every year in the average value. This work used GDP-CVM and inflation rate as inputs to the recognized model.

The specific of data given to the model is time series data, by Ugurlu et al. [4] in 2018, applying a recurrent neural network (RNN) to forecasting the electricity price. The research did not apply the RNN, but the experiment compared neural networks, long short-term memory (LSTM), gate-recurrent unit (GRU), and hidden Markov models, applying a lot of experimentation. The research shows the RNN, GRU, and LSTM give good results over the statistical techniques. Chaudhury et al. [5] in 2020 studied LSTM predicting the energy price in the open electricity market. The research compared LSTM, simple linear regression (SLR), and support vector machine (SVM) with SLR, in the work after applying cross-validation to reducing the overfitting problem. Likiaporn et al. [6] in 2020 used LSTM and bidirectional-LSTM (Bi-LSTM) to forecast the stock market. The Bi-LSTM gives the result over the LSTM that both with and without applying wavelet transform. Moreover, LSTM in [7] is applied to predicting stock market behavior and industry profitable. In 2022, [8] applied the transformer and looked back window in the time deep learning having the result both mean average error (MAE) and root mean square error (RMSE) lower. Zadransk [9] in 2019 implemented the deep neural network by comparing the modified RNN, which gave the result over the simple RNN. Ouyang [10] gives more detail for artificial neural network (ANN) to recognize the time series data. Boonmana and Kulvanich [11] applied ARIMA with ANN and autoregressive integrated moving average (ARIMA) with SVM by the experiment showing ARIMA with SVM gave lower RMSE. From the review above, various models can be recognized in the time series data, in the work of Zhang et al. [12] presenting XGBoost. The XGBoost is a multi-model combining and making a system for recognition, called ensemble learning. In this work, we focus on the selection between deep learning and simple machine learning models by implementing on Python with machine learning libraries.

3. **Implementation.** Implementation detail has four subtopics: data collection, data exploration, data preprocessing, and detail of design model for training.

3.1. **Data collection.** The data from the industry company and EGAT [1] to [3] presented on the official websites, collected from January 2016 to March 2023. The data variable is divided into two groups: independent variable and dependent variable.

The independent variable has 5 variables:

- GDP-CVM is the mass product and service by using volume measures shown in a unit in a million Thai baht;
- PG is the power generation totally by a unit in a gigawatt per hour;
- FX-rate is the foreign exchange rate by the ratio of a Thai baht to a United States Dollar;
- Inflation rate presented in the percentage;
- Natural gas price is the global natural gas price given to IPP and SPP that is presented in a Thai baht per one million British thermal units (MMBTU).

The dependent variable has a parameter:

- Ft price is the fuel adjustment rate presented in one-satang-baht (one-hundredth of a Thai baht) per electric unit.

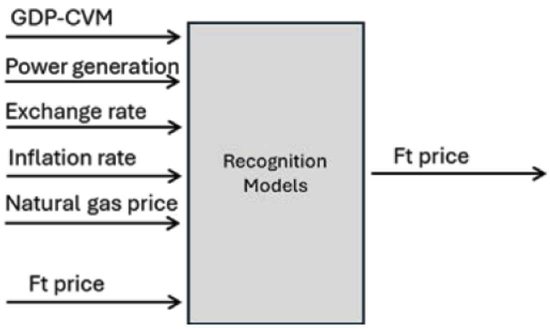


Figure 1. Structure of variables diagram

3.2. **Data exploration.** Figure 2 shows the Ft price from 2016 to 2023 which grew by an approximation of 200%, but 2019 to 2021 is a COVID-19 situation that has low demand and low power consumption. When comparing the Ft price in each month it is found that the beginning of the year has a high power usage and the end of the year has the median to the usage of low power. In Thailand, the summer season is March to May with average generated power between 17,000GWh to 20,000GWh and the cool season is November to February with an average generated power between 14,500GWh to 17,500GWh. The

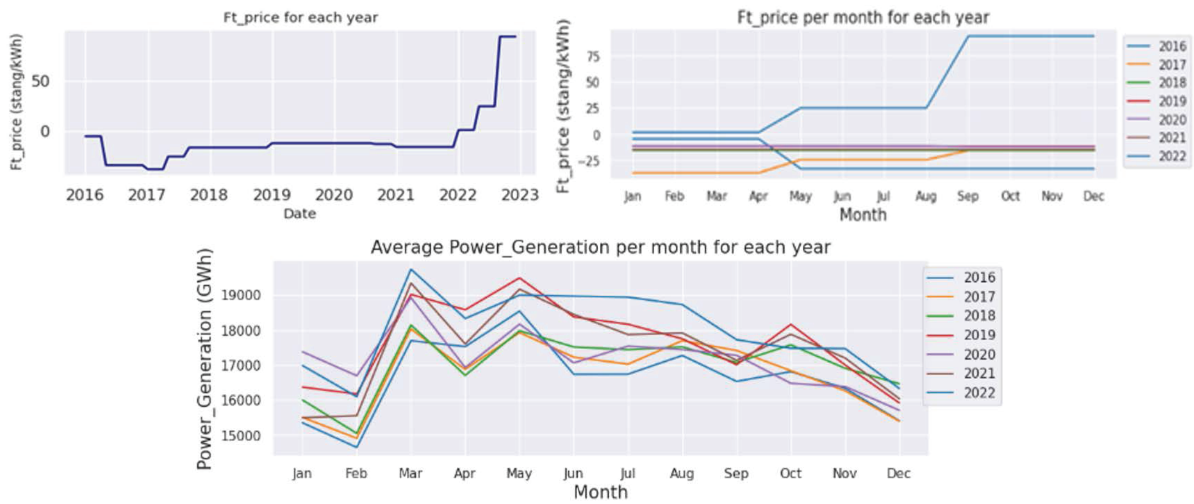


FIGURE 2. (color online) The top left is the Ft price from 2016 to 2023, the top right is the Ft price shown in each month, and the bottom is the average generated power in each month.

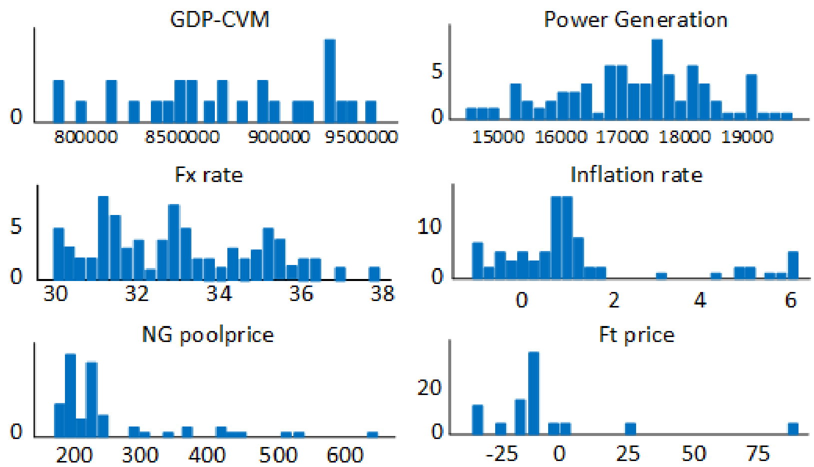


FIGURE 3. Data distribution comparison of each variable

difference in power usage varies by season, with 2016 showing a higher average power demand compared to other years.

From the six data brought to compare the distribution as Figure 3, the Ft price has both positive and negative parts with most of the data being in -50 to 0. The NG pool price averages between 200 to 300. The inflation rate is in 0 to 1%.

3.3. **Data preprocessing.** All data in the difference range in the unit, hundred, and thousand should be converted to the same range before training as Equation (2).

$$X_{scaled} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{2}$$

Then, consider the relation between the variable and Pearson correlation to find the feature selection with filter methods. As Figure 4, the variable NG pool price correlates with 0.8 by the reason that the natural gas price has more effect on the cost of electricity generative.

Comparison between Ft price and inflation rate has the relation between 0.76 and 0.89 with NG pool price. From finding the feature importance with embedded methods it is found that FX rate, inflation rate, and NG pool price are the features more related to the model.

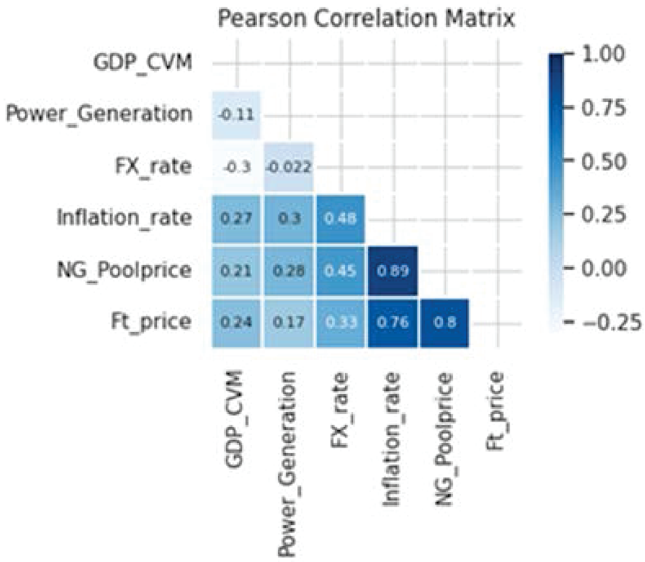


Figure 4. Pearson correlation comparison between each variable

3.4. **Design model and training.** The model has three measurements: traditional forecasting, machine learning techniques, and deep learning.

3.4.1. *The traditional forecasting.* Traditional forecasting is a common approach using historical data and extracting trends, seasonal, and residual by using the decomposing method, as some results in Figure 5. The data after extraction in the trend increasing from 2016 to 2022 and the seasonality has the same pattern occurring every year.

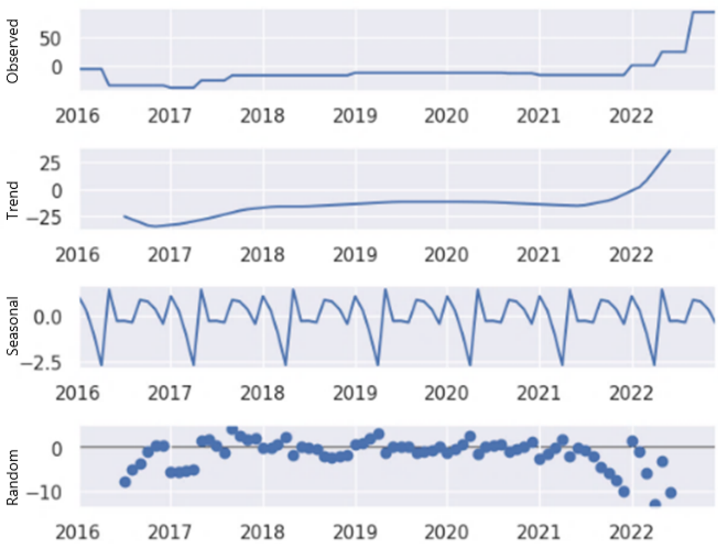


Figure 5. An example of raw data after decomposing into trend, seasonal, and random

3.4.2. *Machine learning.* Machine learning uses historical data after extracting training to machine learning models. This work used 4 methods, linear regression (LR), decision tree (DT), random forest (RF), and XGBoost (XGB) [12]. The specific data is the time series that this work used the TimeSeriesSplit for Walk-Forward Validation (TSS) similar to the K-fold cross-validation, in which the TSS changes the sample data by selecting a small group to train the model that helps reduce overfitting problem. The performance measurement used the mean square error (MSE) Equation (3) and root mean square error (RMSE) Equation (4).

$$\text{MSE}(y, \hat{y}) = \frac{\sum_{i=0}^{N-1} (y_i - \hat{y}_i)^2}{N} \tag{3}$$

$$\text{RMSE}(y, \hat{y}) = \sqrt{\frac{\sum_{i=0}^{N-1} (y_i - \hat{y}_i)^2}{N}} \tag{4}$$

3.4.3. *Deep learning.* Deep learning is developed multi-layer perceptron neural network having many models for training images and signals. This work used the long short-term memory (LSTM) design to recognize time series data. The LSTM has the parameters as Table 1, setting some units at 30 and 80 units. The optimizer method used Adaptive Moment Estimation (Adam) and stochastic gradient descent (SGD), and number of bath sizes has 16, 32 and 64.

TABLE 1. Parameters for training LSTM

Parameter	Range
LSTM units	30 to 80
Drop rate	0.1 to 0.5
Optimizer	Adam, SGD
Batch sizes	16, 32, 64
Epochs	300
Verbose	0
Number of trials	50

4. **Experiment Result.**

4.1. **Traditional forecasting results.** The usage of traditional forecasting found the predicted Ft with non-stationarity data having the p-value at 0.9982 over 0.05 in an acceptance point. Therefore, the model ARIMA and Seasonal Autoregressive Integrated Moving Average (SARIMA) have the feature importance at the lowest MAE at 13.38 and 18.38 from RMSE. The best parameter for the grid search is SARIMA by configuration at $(2, 2, 1)$, $(2, 1, 1, 12)$ as shown in Figure 6.

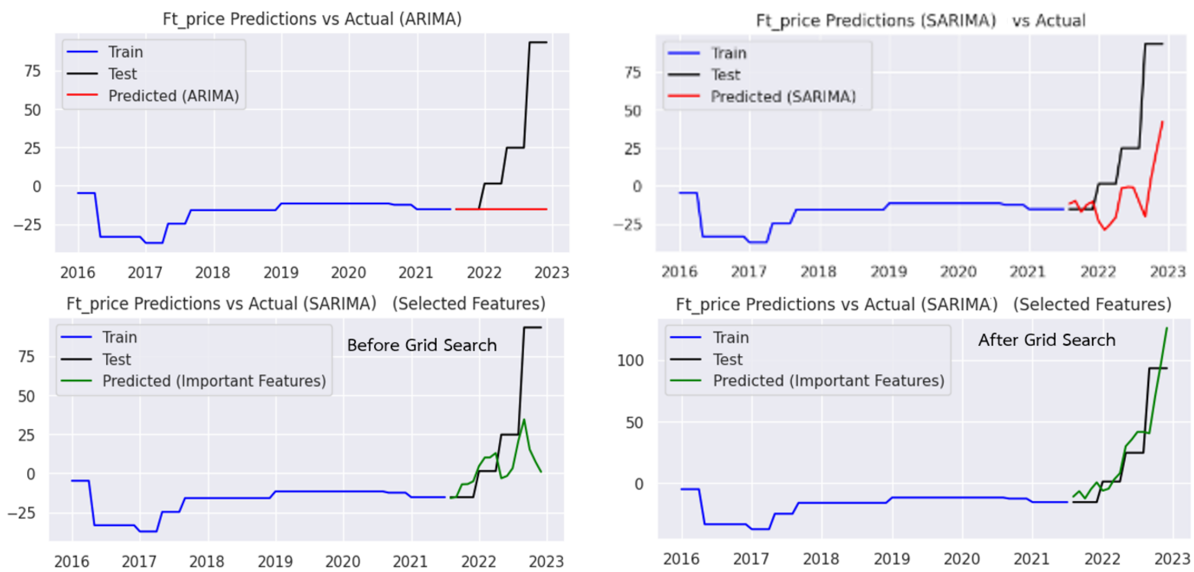


FIGURE 6. (color online) Comparison of the predicted result from ARIMA and SARIMA

4.2. **Machine learning results.** The machine learning from the 4 models, LR, DT, RF, and XGB without identifying feature importance, by the result found that the XGB has the lowest MAE at 8.069 and 10.259 for RMSE as Table 2.

TABLE 2. Comparison results from the machine learning models

Models	MAE	RMSE
Linear regression	12.076	14.922
Decision tree	8.358	11.257
Random forest	10.800	12.465
XGBoost	8.069	10.259

After fine-tuning XGBoost with GridSearchCV using the parameter as Table 3, the result is a fitting of 10 folds with 2,304 candidates, totaling 23,040. The fit parameters include column sample at 3, grammar at 0, learning rate at 0.3, and max depth at 7. The result has MAE at 8.0664 and RMSE at 10.2625 as the result shown in Figure 7, closing before doing the adjustment.

TABLE 3. Fine-tuning XGBoost parameters

Parameter	Values
Number of estimators	50, 100, 150, 200
Learning rate	0.01, 0.05, 0.1, 0.3
Maximum depth	3, 5, 6, 7
Number of subsamples	0.7, 0.9, 1.0
Number of column sample	0.7, 0.9, 1.0
Gamma	0, 0.1, 0.2, 0.3

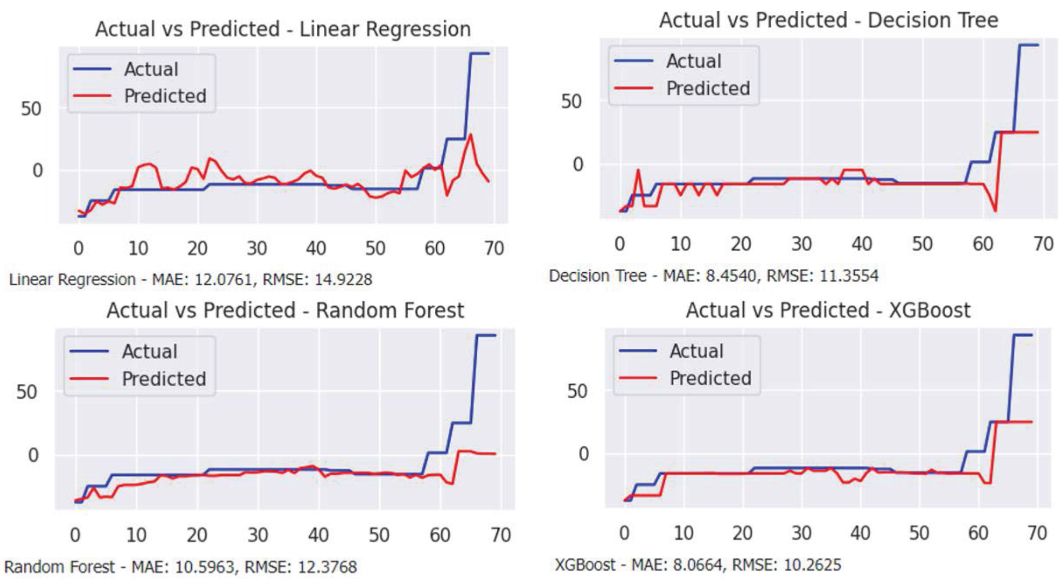
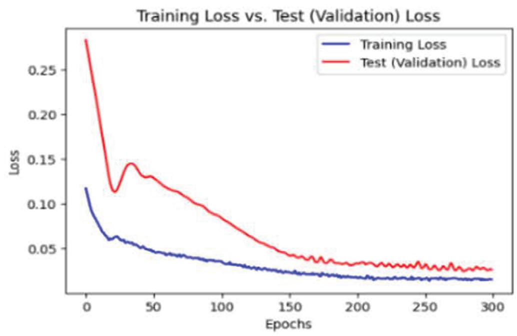


FIGURE 7. (color online) Comparison of the predicted result of LR, DT, RF, and XGBoost

4.2.1. *Deep learning results.* The deep learning in the work selected LSTM giving the lowest validation loss at 0.024, MAE at 0.026, and RMSE at 0.075, as Table 4 and Figure 8 show.

TABLE 4. LSTM parameters giving a low validation loss

Parameter	Values
LSTM units	72
Dropout rate	0.101994
Optimizer	Adam
Batch sizes	64
Best validation loss	0.024483



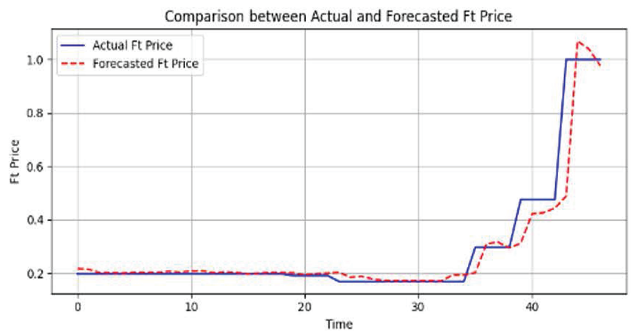


FIGURE 8. (color online) The left-hand side is the validation loss between training and testing on LSTM, and the right-hand side compares the Ft price between the actual and the forecasted result.

4.2.2. *Result of finding an important feature.* The feature importance finding used the linear regression and the Lasso technique found that a multi-collinearity applied with Lasso gave a good coefficient. The result can be interpreted as Lasso adding a penalty term into linear regression that has a coefficient close to zero and able to reduce number of the features. Additionally, this work was applied to the research of the real application of the Ft price calculator that used Steamlit applications as the front-end helping users enter the data and parameters directly on the website through Python.

5. Conclusions.

Forecasting the fuel adjustment chart or Ft has factors related to GDP, amount of electricity produced total, exchange rate, inflation rate, and natural gas price. The analysis found three factors, natural gas price, exchange rate, and inflation rate having high relationship. The experiment found that deep learning has high effectiveness over traditional time series methods and can predict complex data. Moreover, the experiment with deep learning has an appropriate recognition with the data set shown from the experiment result. However, the good result from deep learning should consider some data and parameter adjustments to give the acceptance tolerance in the scope. The developed model in this work can apply the Ft price planning for companies or organizations which should regularly consider another factor to adjustment in the future.

Acknowledgment. This research project was financially supported by Faculty of Informatics, Mahasarakham University.

REFERENCES

[1] Energy Regulatory Commission, *Report: Five Years Action Plan of Energy Controlling 2566 to 2570*, https://www.erc.or.th/th/energy-regulatory-action-plan, 2022.

[2] Thai Government, *Royal Gazette Criteria for Calculation Calculating the Price of Natural Gas under the Governance of the Board of Directors Energy Regulatory Commission*, no.139, pp.12-14, https://www.erc.or.th, 2021.

[3] *Report of Quarterly Gross Domestic Product-Chain Volume Measures (GDP-CVM) 2022*, Official of the National Economic and Social Development, https://www.nesdc.go.th/, 2022.

[4] U. Ugurlu, I. Oksuz and O. Tas, Electricity price forecasting using recurrent neural networks, *Energies*, vol.11, no.1255, 23, DOI: 10.3390/en11051255, 2018.
[5] P. Chaudhury, A. Tyagi and P. K. Shanmugam, Comparison of various machine learning algorithms for predicting energy price in open electricity market, *2020 International Conference and Utility Exhibition on Energy, Environment and Climate Change (ICUE)*, Pattaya, Thailand, pp.1-7, DOI: 10.1109/ICUE49301.2020.9307100, 2020.
[6] P. Likiaporn, S. Phanphocha and B. Puangkerd, Forecasting stock market indices using wavelet transform, LSTM, and Bi-LSTM, *The 16th RSU National Graduate Research Conference*, vol.16, 2020.
[7] J. Li, G. Guo and X. Hu, Prediction algorithm of industry rotation based on attention LSTM model, *International Journal of Innovative Computing, Information and Control*, vol.18, no.6, pp.1969-1977, DOI: 10.24507/ijicic.18.06.1969, 2022.
[8] J. Shi, M. Jain and G. Narasimhan, Time series forecasting using various deep learning models, *World Academy of Science, Engineering and Technology, International Journal of Computer and Systems Engineering*, vol.16, no.6, pp.224-232, 2022.
[9] L. Zadransk, *Time Series Forecasting Using Deep Neural Networks*, Master Thesis, University of West Bohemia, 2019.
[10] Y. Ouyang, Time series prediction with a non-causal neural network, *2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, London, UK, pp.25-31, DOI: 10.1109/CIFEr.2014.6924050, 2014.
[11] C. Boonmana and N. Kulvanich, A comparative prediction accuracy of hybrid time series models, *Thai Science and Technology Journal*, vol.25, no.2, pp.177-190, 2017.
[12] D. Zhang, L. Qian, B. Mao, C. Huang, B. Huang and Y. Si, A data-driven design for fault detection of wind turbines using random forests and XGBoost, *IEEE Access*, vol.6, pp.21020-21031, DOI: 10.1109/ACCESS.2018.2818678, 2018.