

ANALYSIS OF LOAD FORECASTING ACCURACY BASED ON HO CHI MINH CITY DATA

TRẦN THANH NGỌC

*Khoa Công nghệ điện, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh,
tranthanhngoc@iuh.edu.vn*

Abstract. Short term load forecasting is one of the fundamental parts of the electric system. Among exponential smoothing methods, the Holt-Winters method is widely used to forecast the short-term load since it is easy and simple to use, and it has high ability to adapt to the forecast of different time horizons. This paper presents a new approach by combining Holt-Winters and Walk-Forward Validation methodology to forecast the maximum power demand for Ho Chi Minh City, Vietnam. The data is divided into the training and test sets in many cases. The forecast accuracy of the mean absolute error (MAE) and the mean absolute percentage error (MAPE) are used to analyze the characteristic of forecast for each day of the week.

Keywords. Holt-Winters, Short-term load forecasting, Walk-Forward Validation, forecast accuracy.

1 INTRODUCTION

Load forecasting is an important part of electric power system, including the generation, transmission, distribution and retail of electricity. Depending on different forecast horizons and resolutions, load forecast problems can be divided into 3 groups: long-term, mid-term, and short-term. Long-term forecasts of the peak load are necessary for capacity planning and maintenance scheduling. Mid-term demand forecasts are applied for power system operation and planning. Short-term load forecasts are required for the control and scheduling of power systems [1-5].

There are several ways used for short-term load forecasting, for that the exponential smoothing method is considered as one of the most popular approaches due to the simplicity to apply to yield forecasts for real data with a level of accuracy comparable to that of alternative complex methods. The most general form of exponential smoothing methods is named as Holt-Winters consisting of level, trend, and seasonal components in the time series [6-15].

In order to apply Holt-Winters method, the common way is to split the data into training and test sets, which are used to build the forecast model and to measure the accuracy of forecast values, respectively. And it is easier to see that the training set and the forecast model is fixed during forecast operation. Unlike the traditional way, the Walk-Forward Validation Methodology allows to retrain the forecasting model as new data becomes available, and to get the best forecasts at each time step [16-17]. Furthermore, in the case of applying the Holt-Winters method for short-term load forecasting, the results reported in literature are mostly concentrated on the total forecast accuracy as values of MAE, MAPE for one week, a few weeks or one month [6-15], while the forecast accuracy for each day of week has not considered yet. Obviously, the load demands for days of a week are not the same, for instance, it could be highest on working days and lowest on weekends. Thus, the accuracy for each day of a week is essential and its understanding will be useful for in real load forecasting.

In the present work, the Holt-Winters method and Walk-Forward Validation are combined to evaluate the accuracy of load forecasting for each day of a week based on the maximum power demand data of Ho Chi Minh city. This paper will be organized as follows. Section 2 presents the basic theories including Exponential smoothing method, Walk-Forward Validation Methodology and the forecast accuracy. Section 3 provides predictions and discussion. The conclusions are provided in Section 4.

2 FORECASTING METHODS

2.1 Exponential smoothing method

Exponential smoothing method is one of the most promising forecasting strategies for time series with the simplest form called as simple exponential smoothing or single exponential smoothing (SES). Then Holt extended SES method allowing to forecast data based on a trend called Holt's linear method. After that Holt and Winters continued to improve to get the well-known Holt-Winters method used for capturing seasonality [18-19].

A. SES method

SES method is applied for non-seasonal and trend time series, the only component considered here is the level l_t . The equations of SES method are given as follows [18-19]:

$$\begin{aligned} F_{t+h|t} &= l_t \\ l_t &= \alpha y_t + (1 - \alpha)l_{t-1} \end{aligned} \quad (1)$$

B. Holt's linear trend method

Holt's linear trend method is suitable for non-seasonal data, which contains the trend b_t and the level l_t components. The Holt's linear trend method is expressed in the following equations [18-19]:

$$\begin{aligned} F_{t+h|t} &= l_t + hb_t \\ l_t &= \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \end{aligned} \quad (2)$$

C. Holt-Winters method

Holt-Winters method is appropriate for data with trend and seasonal components. There are two basic models for the Holt-Winter method [18-19]:

i. Additive Seasonal Model

$$\begin{aligned} F_{t+h|t} &= l_t + hb_t + s_{t+h-m(k+1)} \\ l_t &= \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \\ s_t &= \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m} \end{aligned} \quad (3)$$

This is called additive because the seasonal component is added to level and trend components.

ii. Multiplicative Seasonal Model

This is called multiplicative because seasonal component is multiplied by the total level and direction components.

$$\begin{aligned} F_{t+h|t} &= (l_t + hb_t) s_{t+h-m(k+1)} \\ l_t &= \alpha \frac{y_t}{s_{t-m}} + (1 - \alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} \\ s_t &= \gamma \frac{y_t}{l_{t-1} - b_{t-1}} + (1 - \gamma)s_{t-m} \end{aligned} \quad (4)$$

where in equations (1), (2), (3), (4):

- h is step-ahead forecast, $h = 1, 2, \dots$; α is the smoothing parameter.
- m is the frequency of the seasonality, for quarterly data $m=4$, for weekly data $m=7$, for monthly data $m=12, \dots$
- k is the integer part of $(h-1)/m$
- α, β, γ is the smoothing parameters.
- l_t, b_t, s_t are the level, trend, and seasonal components.
- Y_t is observed value and F_t is forecasting value at time t

The Additive Seasonal Model and Multiplicative Seasonal Model give the same results in most cases. The Multiplicative Seasonal Model is utilized in this paper.

2.2 Walk-Forward Validation Methodology

In load forecasting practice, it is better to retrain the forecasting model as new data becomes available. The Walk Forward Validation methodology gives the load forecasting model with the best opportunity to make good forecasts at each time step. The sequential operation of the Walk Forward Validation methodology is shown in Table 1 below. Firstly, using the history data (Weeks) for training, the forecasting model makes a load forecasting for the next week (Week1) and then stored or evaluated against the known value. Continuously, the training data is expanded to include the know value (Weeks + Week1), the forecasting model is updated and the next week is forecasted (Week2). The process is repeated to the end [16-17].

Table 1: The rolling of data in the Walk Forward Validation methodology.

History data	Predictive data
[Weeks]	Week1
[Weeks + Week1]	Week2
[Weeks + Week1 + Week2]	Week3
....	...

2.3 The forecast accuracy

To measure the accuracy of the forecasting data, the criteria Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) have been chosen. The equations of MAE and MAPE are given by [7,8,10]:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - F_i| \quad (5)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - F_i}{Y_i} \right| \quad (6)$$

where:

- Y_i is the actual observed values
- F_i is the forecasting values
- n is the number of observed values

2.4 The framework for Walk-Forward Validation Methodology based on Holt-Winters method

The framework for Walk-Forward Validation Methodology based on Holt-Winters method is shown in Figure 1. Firstly, the data is split into history data [Weeks] and testing data [Week1, Week 2, ..., Weekn]. The history data was used in training process of HW method and forecast the values for the first week [Week1]. Then the obtained data of the week 1 is added into the history data, and therefore the history data now includes Weeks + Week 1 values. Next, the training process of HW method is performed again to forecast the values for the second week [Week2]. The process is sequentially repeated by n steps, where the n^{th} week of testing values is added into history data, and therefore we have the forecast value of n^{th} Week. Finally, the forecasting accuracy (MAPE, MAE) for each day of a week is calculated based on the forecasting values and the testing values.

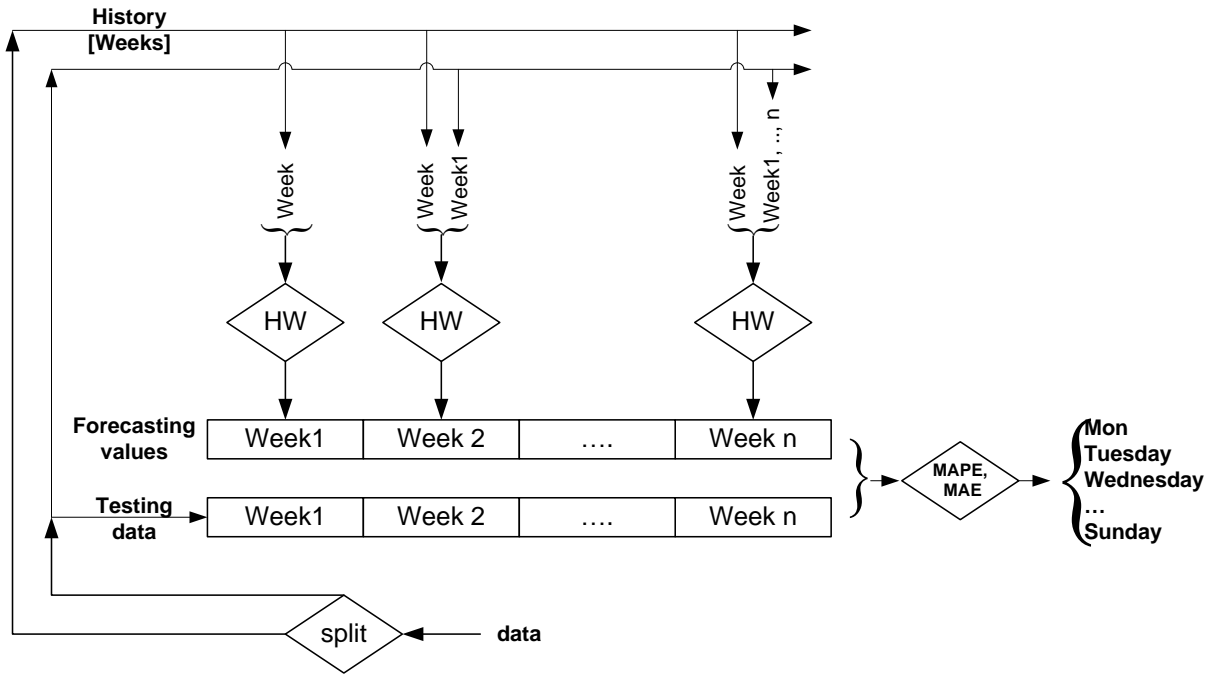


Figure 1: The framework for Walk-Forward Validation Methodology based on Holt-Winters method

3 PREDICTION RESULTS & DISCUSSION

3.1 Data description

Ho Chi Minh City is the largest City in Vietnam, and is also one of Vietnam's most important economic centers. With a population of over 10 million people and a concentration of industrial clusters, the electricity demand for Ho Chi Minh City is extremely necessary and important. In the paper, we use the data of maximum power demand (P_{max}) of Ho Chi Minh City. The dataset starts from Monday, January 10, 2011 and ends on Sunday, December 30, 2018. The dataset consists of 8 years, 52 weeks per year and 7 days per week, including $8 \times 52 \times 7 = 2912$ days. A typical week starts from Monday to Sunday. Table 2 below shows the value of dataset for the first week and the last week.

Table 2: The first week and last week of max Power demand in Ho Chi Minh city

Day of week	Date	P_{max} (MW)	Day of week	Date	P_{max} (MW)
2	1/10/11	2260.37
3	1/11/11	2247.6	2	12/24/18	3695.2
4	1/12/11	2241.3	3	12/25/18	3675.1
5	1/13/11	2227.45	4	12/26/18	3646
6	1/14/11	2246.1	5	12/27/18	3636.7
7	1/15/11	2145.06	6	12/28/18	3494.1
CN	1/16/11	1709.45	7	12/29/18	3387.9
....	CN	12/30/18	2564.2

Figures 2-5 below show the plotting of dataset in 8 years, the last year, the last month and the last week of dataset, respectively. In Figures 2 and 3, there are some values that dramatically decrease, corresponding to the Luna New Year time. Figure 4 clearly indicates that the data have seasonality with the period of 7 days. It is obviously seen that the P_{max} will decrease on weekends (Saturday and Sunday).

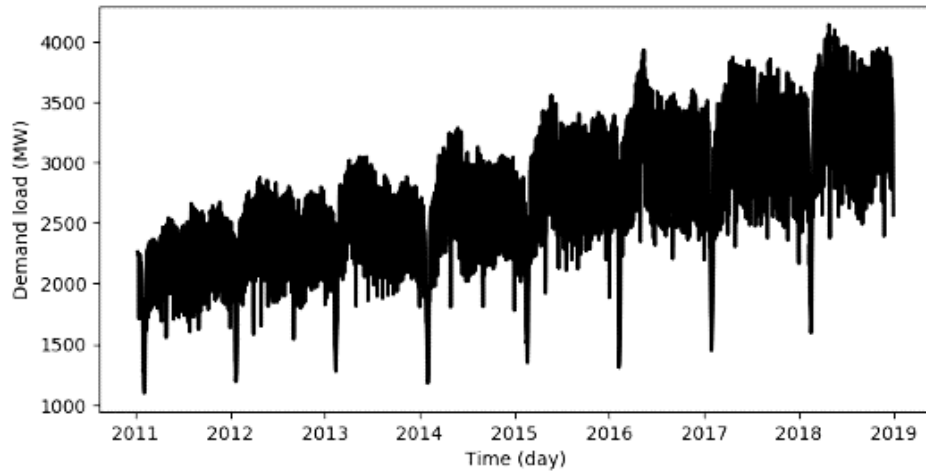


Figure 2: The data from 2011-2018.

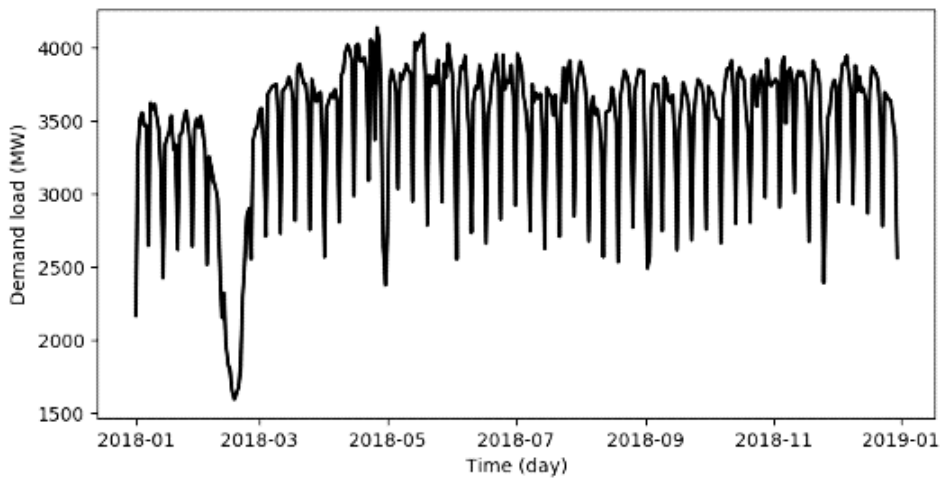


Figure 3: The last year's data

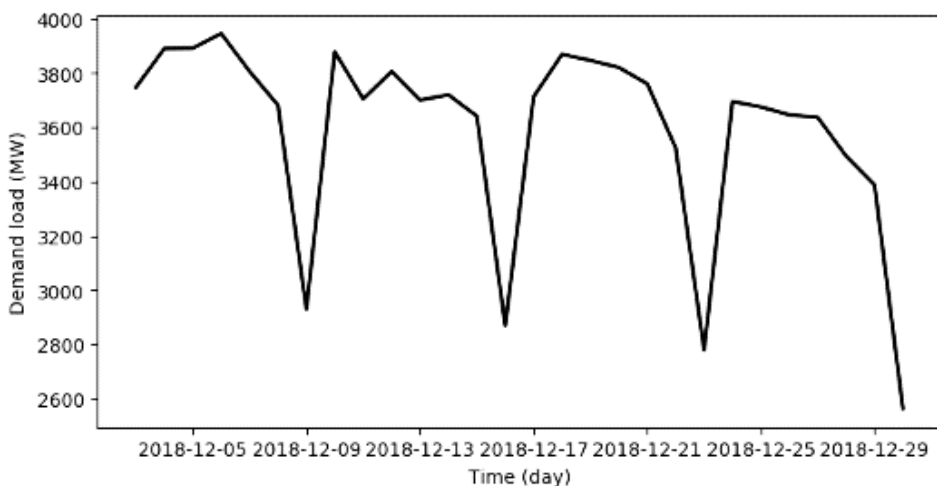


Figure 4: The last month's data.

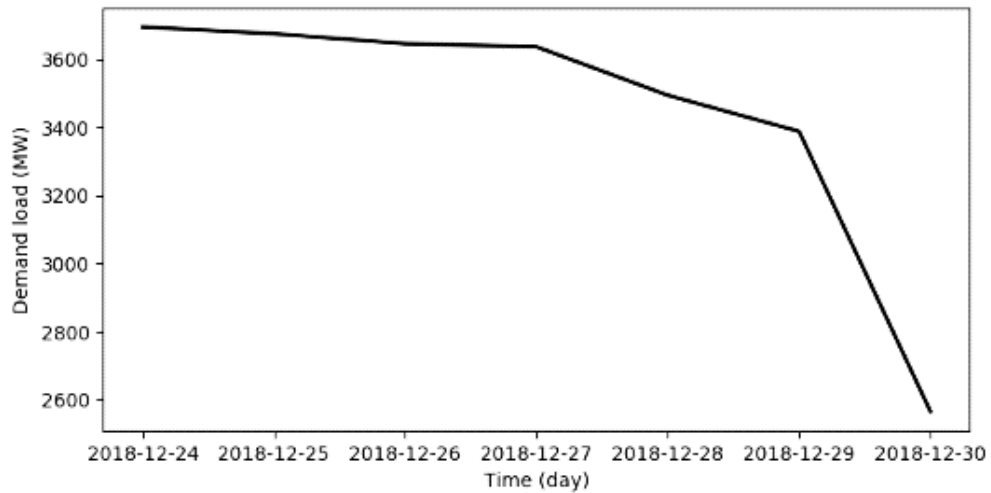


Figure 5: The last week's data.

Figure 6 shows the decomposing analysis results for the last year with the train, seasonal and residual components of the observed data. Figure 6 clearly shows that the data have seasonal with period of 7 days.

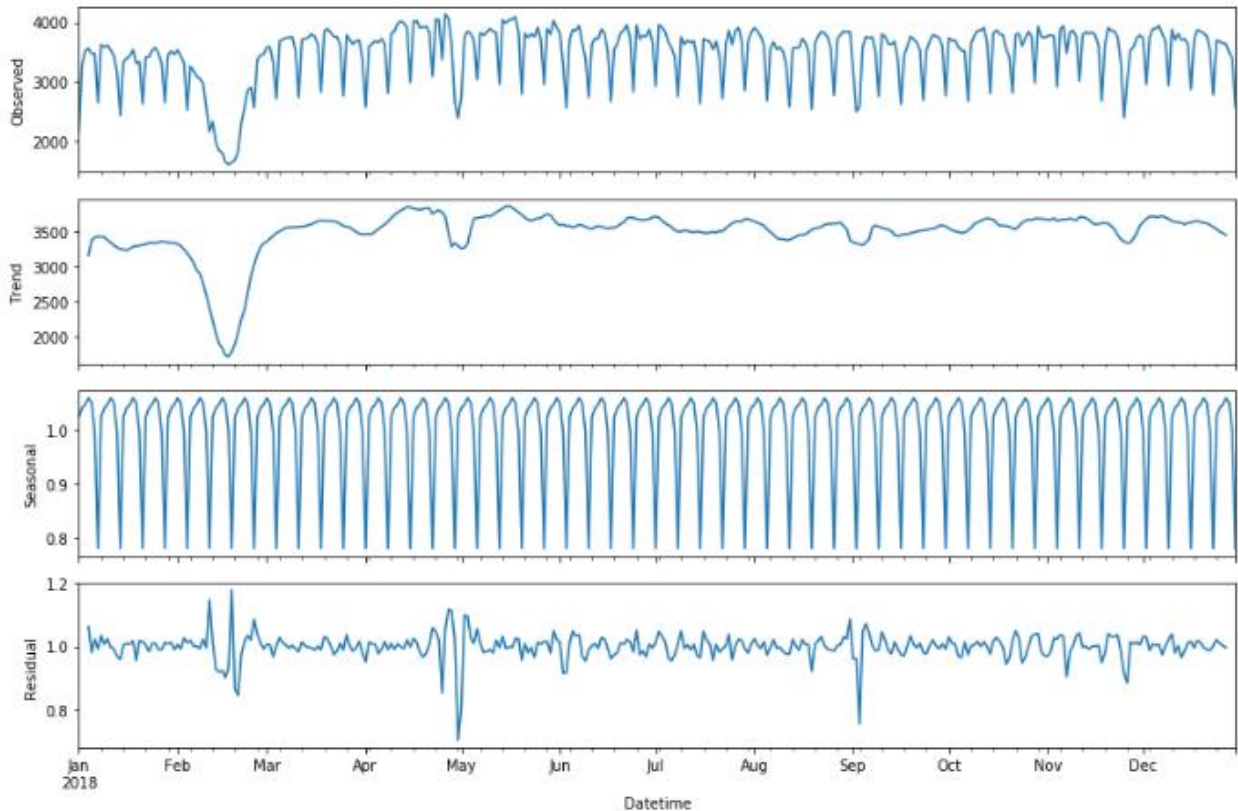


Figure 6: decomposing analysis for components of the last year.

3.2 Results and Discussion

In the paper, we use the data of max power (P_{\max}) of Ho Chi Minh city as described above to analyze the characteristics of forecasting's error for days of a week. The Walk Forward Validation methodology is applied to make sure that the training data is updated with available data from test set after one week ahead forecast step. The load data are split into training set and test set in several cases as shown in Table 3 below:

Table 3: The training and testing set data.

Training set (days)	Test set (days)
7*52*7	1*52*7
6*52*7	2*52*7
5*52*7	3*52*7
4*52*7	4*52*7
3*52*7	5*52*7
2*52*7	6*52*7
1*52*7	7*52*7

Table 4 shows the MAE values in case of the test time from 1 year ($1 \times 52 \times 7 = 364$ days) to 7 years ($7 \times 52 \times 7 = 2548$ days) for each day of a week along with the average value.

Table 4: MAE values for day of the week and the average value

Test set (days)	MAE (MW)							
	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Ave
1*52*7	156.7	148.8	181.5	181.6	205.7	222.1	187.5	183.4
2*52*7	147.7	158.2	176.2	190.4	198.9	229.5	177.7	182.7
3*52*7	140.2	152.7	170.0	179.4	198.3	230.6	169.8	177.3
4*52*7	129.9	150.0	168.3	178.1	198.8	216.9	162.7	172.1
5*52*7	124.2	145.2	167.4	176.2	191.1	205.1	150.7	165.7
6*52*7	119.6	140.6	162.4	167.2	182.4	192.3	143.4	158.3
7*52*7	117.6	139.3	155.7	162.0	176.2	185.6	141.2	153.9

Table 5 shows the MAPE values in case of the test time from 1 year ($1 \times 52 \times 7 = 364$ days) to 7 years ($7 \times 52 \times 7 = 2548$ days) for each day of a week along with the average value.

Table 5: MAPE values for days of week and the average value

Test set (days)	MAPE (%)							
	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Ave
1*52*7	5.5	4.7	5.6	5.6	6.5	7.2	7.3	6.1
2*52*7	5.0	5.0	5.5	6.0	6.4	7.8	7.2	6.1
3*52*7	4.8	4.9	5.4	5.8	6.6	8.1	7.0	6.1
4*52*7	4.5	5.0	5.6	6.0	7.0	7.8	6.8	6.1
5*52*7	4.5	5.0	5.9	6.4	7.0	7.6	6.5	6.1
6*52*7	4.4	5.0	5.9	6.1	6.8	7.3	6.4	6.0
7*52*7	4.5	5.2	5.8	6.1	6.7	7.2	6.5	6.0

Figure 7 presents the results in case of the test time for 1 years, $1 \times 52 \times 7 = 364$ days. Figure (7a) shows the testing and forecasting values. Figures 7b and 7c show MAE and MAPE values of Monday, Tuesday, Wednesday, Thursday, Friday, Saturday and Sunday, respectively.

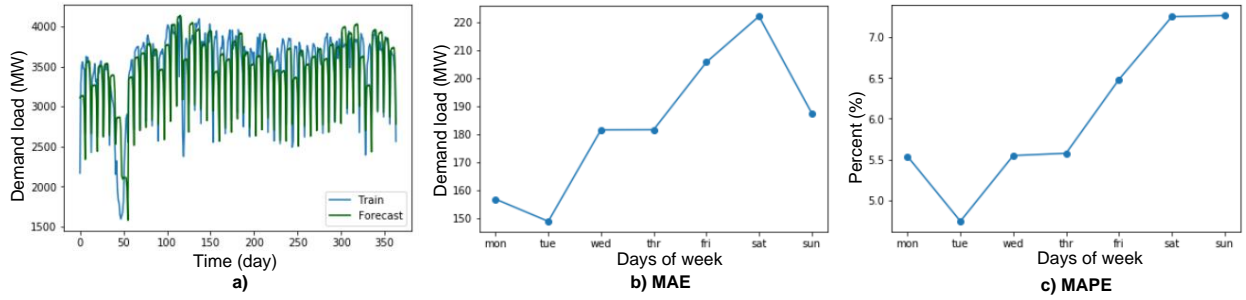


Figure 7: The results in case of $n_test = 1 \times 52 \times 7$ days.

Figures 8 - 13 show the results in case of the test time from 2 years ($2 \times 52 \times 7 = 728$ days) to 7 years ($7 \times 52 \times 7 = 2548$ days), respectively.

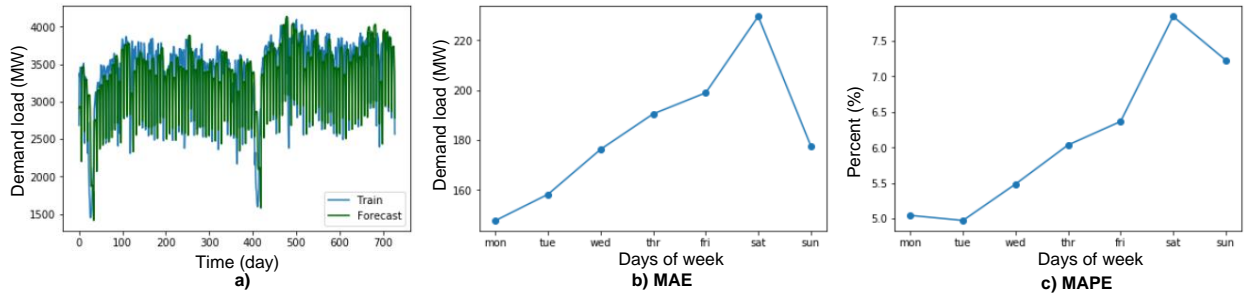


Figure 8: The results in case of $n_test = 2 \times 52 \times 7$ days.

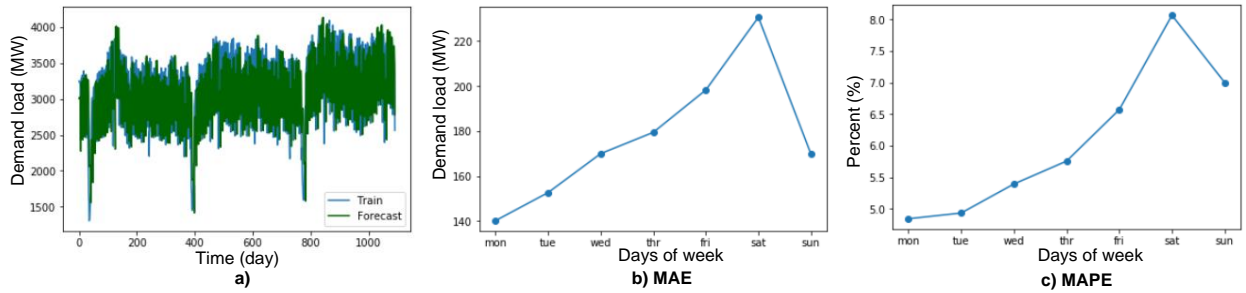


Figure 9: The results in case of $n_test = 3 \times 52 \times 7$ days.

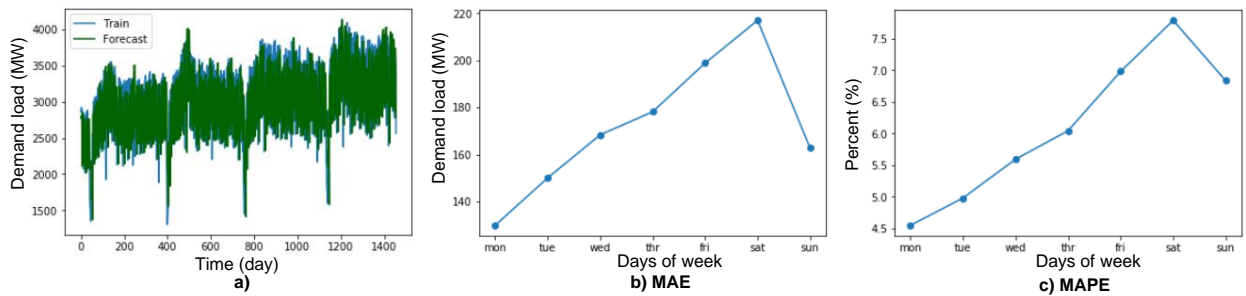
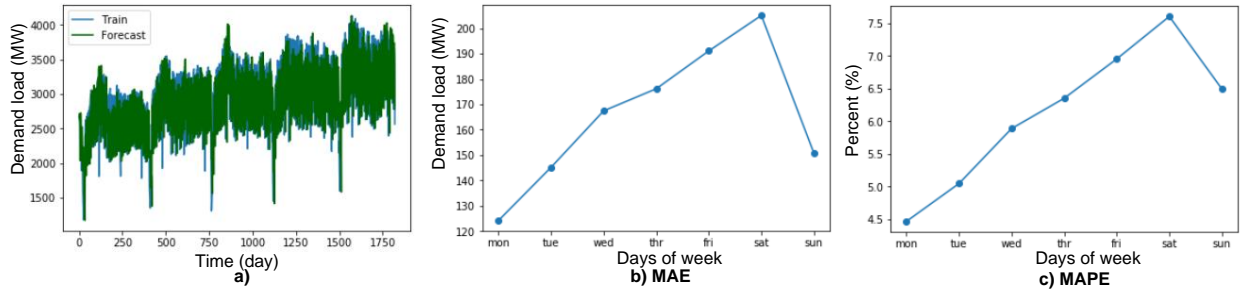
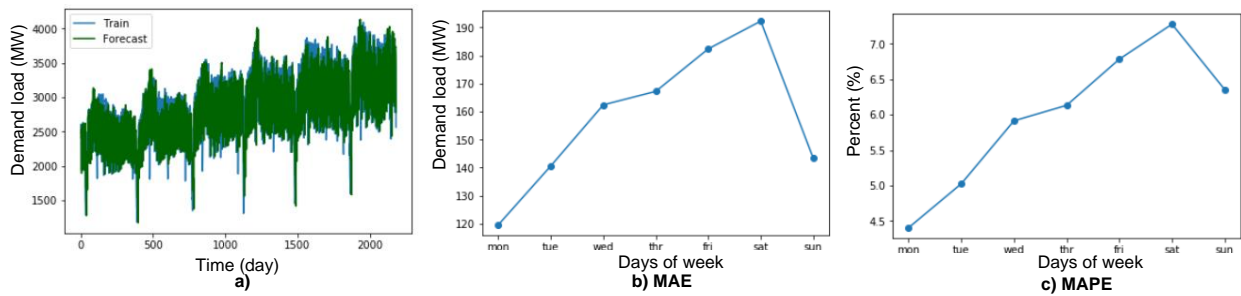
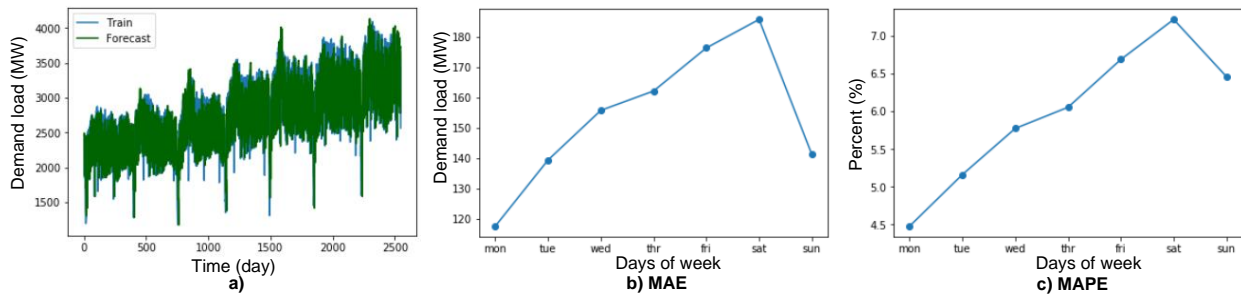


Figure 10: The results in case of $n_test = 4 \times 52 \times 7$ days.

Figure 11: The results in case of $n_{\text{test}} = 5 \times 52 \times 7$ days.Figure 12: The results in case of $n_{\text{test}} = 6 \times 52 \times 7$ days.Figure 13: The results in case of $n_{\text{test}} = 7 \times 52 \times 7$ days.

Analyzing the Tables 4, 5 and Figures 7 – 13 indicates that:

- By applying Walk-Forward Validation Methodology to forecast the P_{max} value for one week ahead, the errors of forecasting are small. Especially, the average value of MAPE was obtained in the region of 6%. In this regard, the proposed method demonstrated by itself as a reliable forecasting tool.
- Curves of MAE and MAPE for each day of a week show an increasing trend from Monday to Saturday and an opposite trend (decreasing) on Sunday.
- The errors (MAE and MAPE) observed for Sunday, Tuesday and Wednesday are smaller than that for Thursday, Friday and Saturday. This means that the first three days are easier to forecast, while next three days are difficult to forecast.

4 CONCLUSIONS

In this paper, the combination of Holt-Winters method and Walk-Forward Validation methodology has been utilized to analyze for the Ho Chi Minh City maximum power data. The load forecasting errors of the mean absolute error (MAE) and the mean absolute percentage error (MAPE) are small, proving the high reliability of the proposed method. The analysis of the accuracy of load forecasting values for each day of a week indicated that MAE and MAPE increased from Monday to Saturday and decreased on Sunday. The obtained results can be useful for real load forecasting.

REFERENCES

- [1] J.W. Taylor, P.E. McSharry, "Short-term load forecasting methods: an evaluation based on European data", *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2213-2219, 2007
- [2] P. E. McSharry, S. Bouwman, G. Bloemhof, "Probabilistic forecasts of the magnitude and timing of peak electricity demand", *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1166-1172, May 2005.
- [3] E. Gonzalez-Romera, M. A. Jaramillo-Moran, D. Carmona-Fernandez, "Monthly electric energy demand forecasting based on trend extraction", *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1946-1953, Nov. 2006.
- [4] M. P. Garcia, D. S. Kirschen, "Forecasting system imbalance volumes in competitive electricity markets", *IEEE Trans. Power Syst.*, vol. 21, no. 1, pp. 240-248, Feb. 2006.
- [5] J. W. Taylor, "Density forecasting for the efficient balancing of the generation and consumption of electricity", *Int. J. Forecasting*, vol. 22, pp. 707-724, 2006.
- [6] Al-maamary, G. H. S. (2012). Short and Medium Iraqi Load Forecast Using Holt-Winter Method And Wavelet Transformation. 3(5), 225–228.
- [7] Abd Jalil, N. A., Ahmad, M. H., & Mohamed, N. (2013). Electricity load demand forecasting using exponential smoothing methods. *World Applied Sciences Journal*, 22(11), 1540–1543.
- [8] Abd. Razak, F., Shitan, M., Hashim, A. H., & Z. Abidin, I. (2009). Load Forecasting Using Time Series Models. *Jurnal Kejuruteraan*, 21(1), 53–62.
- [9] Badran, S. M. (2009). Short term electrical load forecasting. *Australian Journal of Basic and Applied Sciences*, 3(3), 2697–2705.
- [10] Bindiú, R., Chindriú, M., & Pop, G. V. (2009). Day-Ahead Load Forecasting Using Exponential Smoothing. *Scientific Bulletin of the Petru Maior University of Tirgu Mures*, 6(Xxiii).
- [11] Dang-Ha, T. H., Bianchi, F. M., & Olsson, R. (2017). Local short term electricity load forecasting: Automatic approaches. *Proceedings of the International Joint Conference on Neural Networks, 2017-May*, 4267–4274.
- [12] Jónsson, T., Pinson, P., Nielsen, H. A., & Madsen, H. (2014). Exponential smoothing approaches for prediction in real-time electricity markets. *Energies*, 7(6), 3710–3732.
- [13] Kavanagh, K., & Kavanagh, K. (2017). *Short Term Demand Forecasting for the Integrated Electricity Market Short Term Demand Forecasting for the Integrated Single Electricity Market*. 2(1).
- [14] Souza, R. C., & Miranda, C. V. C. De. (2007). Short term load forecasting using double seasonal exponential smoothing and interventions to account for holidays and temperature effects. *TIAIO II – 2º Taller Latino Iberoamericano de Investigación de Operaciones*, 1–8.
- [15] Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, 54(8), 799–805.
- [16] Jason Brownlee, Introduction to Time Series Forecasting with Python, [Online], 2019
- [17] Jason Brownlee, Deep Learning for Time Series Forecasting, [Online], 2019
- [18] Hyndman, R. J. & Athanasopoulos, G., Forecasting: principles and practice, 2019, [Online]. Available: OTexts.org/fpp/
- [19] Hyndman, R., Koehler, A.B., Ord, J.K., Snyder, R.D., Forecasting with Exponential Smoothing: The State Space Approach, *Springer-Verlag Berlin Heidelberg*, 2008

PHÂN TÍCH ĐỘ CHÍNH XÁC DỰ BÁO PHỤ TẢI ĐIỆN CHO KHU VỰC THÀNH PHỐ HỒ CHÍ MINH

Tóm tắt. Dự báo phụ tải ngắn hạn là một trong những thành phần cơ bản trong vận hành hệ thống điện. Các phương pháp dự báo san bằng hàm mũ, mà trong đó đặc biệt là phương pháp Holt-Winters được sử dụng rộng rãi cho dự báo phụ tải ngắn hạn, bởi vì chúng dễ dàng, đơn giản khi sử dụng, cũng như có khả năng thích ứng cao để dự báo cho các khoảng thời gian khác nhau. Bài báo này giới thiệu phương pháp Holt-Winters kết hợp với phương pháp Walk-Forward Validation để dự báo nhu cầu phụ tải cực đại cho khu vực thành phố Hồ Chí Minh, Việt Nam. Dữ liệu được chia thành các tập huấn luyện và kiểm tra trong nhiều trường hợp. Độ chính xác của dự báo MAE và MAPE được sử dụng để phân tích đặc tính dự báo cho các ngày trong tuần.

Từ khóa. Holt- Winters, Dự báo phụ tải ngắn hạn, Walk-Forward Validation, Độ chính xác dự báo.

Ngày nhận bài: 26/11/2019

Ngày chấp nhận đăng: 26/02/2019