

Research on the impact of sliding window and differencing procedures on the support vector regression model for load forecasting

Thanh Ngoc Tran¹, Thi Phuc Dang², Binh Minh Lam¹, Anh Tuan Nguyen¹

¹Faculty of Electrical Engineering Technology, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam

²Faculty of Information Technology, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam

Article Info

Article history:

Received Jul 25, 2023

Revised Nov 30, 2023

Accepted Dec 13, 2023

Keywords:

Differencing procedure

Load forecasting

Machine leaning

Sliding window procedure

Support vector regression

ABSTRACT

Load forecasting is a critical aspect of energy management and grid operations. Machine learning techniques as support vector regression (SVR), have been widely used for load forecasting. However, the effectiveness of SVR is highly dependent on its hyperparameters, including the error sensitivity parameter, penalty factor, and kernel function. Furthermore, as the load data follows a time series pattern, the accuracy of the SVR model is influenced by the data's characteristics. In this regard, the present study aims to investigate the impact of combining the sliding window procedure and differencing the input data on the prediction accuracy of the SVR model. The study utilizes daily maximum load data from the Queensland and Victoria states in Australia. The analyses revealed that while the sliding window procedure had a minimal impact on the prediction results, the differencing of the input data significantly improved the accuracy of the predictions.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Thanh Ngoc Tran

Faculty of Electrical Engineering Technology, Industrial University of Ho Chi Minh City

12 Nguyen Van Bao, Ward 4, Go Vap District, Ho Chi Minh City, Vietnam

Email: tranhanhngoc@iuh.edu.vn

1. INTRODUCTION

Electrical load forecasting refers to the process of predicting future electricity load demand within a specific region, system, or grid. The primary purpose of load forecasting is to assist energy providers, utilities, and grid operators in efficiently planning and managing the generation, transmission, and distribution of electricity or energy. Electrical load forecasting plays a vital role in all aspects of the electrical power system, including generation, transmission, distribution, retail sale of electricity [1]–[3]. Load forecasting can be classified into four main categories based on the period of prediction time: Very short-term load forecasting (VSTLF), short term load forecasting (STLF), medium term load forecasting (MTLF), and long-term load forecasting (LTLF). Specifically, VSTLF focuses on predicting the electricity load within a short time horizon, typically ranging from a few minutes to a few hours. Meanwhile, the categories of STLF, MTLF, and LTLF take a longer time to predict the electricity load from a few hours to a few days, a few weeks to a few months, and even several years to decades, respectively [4], [5].

There are numerous techniques and methodologies which can be used for electricity load forecasting, such as multiple regression, exponential smoothing, autoregressive integrated moving average, artificial neural networks, support vector regression, deep learning models, and ensemble methods [6]–[8]. For these approaches, support vector regression (SVR), a variant of support vector machines with several advantages for regression tasks, has gained popularity as a promising solution for electricity load forecasting

in recent years [9], [10]. The prediction precision of the SVR model depends on its hyperparameters of ε (error tolerance), C (penalty parameter), Kernel functions, and Kernel parameters [11], [12]. In addition to the hyperparameters of the SVR model, the preprocessing steps applied to input data, such as sliding window and differencing procedures also have a significant impact on the accuracy of the prediction results. The sliding window procedure is a commonly used technique in combination with SVR for handling time series data, including the case of load forecasting. In the sliding window procedure, the window size is an important factor which must be considered because it determines the number of past observations used as input features for predicting the future load. In other words, the window size should be taken into account when using the sliding window procedure with SVR due to the direct impact on the accuracy of the model's predictions [13]–[16]. Regarding the utilization of sliding window procedure, another important characteristic that cannot be ignored for the daily maximum load data is its repetitive nature. For example, the maximum load on a Monday this week may have a similar value to the maximum load on the previous Mondays, and so on. Therefore, the method of differencing the input data can also be applied to improve the accuracy of the prediction results [17].

In this study, the sliding window procedure and differencing the input data are combined to investigate their impact on the accuracy of load forecasting based on the SVR network. A large number of different combinations of the hyperparameters C , ε , and γ of the SVR network are considered. Additionally, the daily maximum load data taken from the states of Victoria and Queensland, Australia, are used in the study. The obtained forecast errors will be used to analyze the performance of the sliding window procedure and differencing the input data in relation to the accuracy of the SVR model.

The paper is organized as follows. Section 2 provides a critical discussion of support vector regression and introduces the methodology that combines the sliding window procedure and differencing the input data. Section 3 presents the experimental results with analyses and discussions for the findings. Section 4 concludes the paper, summarizing the main findings and contributions of the study.

2. METHOD

2.1. Support vector regression

In SVR, the training data consists of pairs $(x_i, y_i) \in \mathbb{R}^n \times \mathbb{R}$, where $x_i \in \mathbb{R}^n$ and $y_i \in \mathbb{R}$ that represent the feature input vector and the corresponding target value, respectively. The index $i = 1, \dots, m$, is assigned to the number of training instances, while n is referred to the number of features in each instance. In SVR, the input x is mapped to a higher dimensional feature space using the function $\varphi(x)$ which is expressed as (1) [18], [19]:

$$f(x) = \langle \omega, \varphi(x) \rangle + b \quad (1)$$

Here, the symbol $\langle \dots \rangle$ represents the dot product, where ω is the weight vector and b is the bias term. The objective is to estimate the values of ω and b by minimizing the following regularized risk:

$$R = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m L(y_i, f(x_i)) \quad (2)$$

where $L(y_i, f(x_i))$ represents the linear ε -insensitive loss, which is defined as (3).

$$L(y, f(x)) = |y - f(x)|_{\varepsilon} = \max\{0, |y - f(x)| - \varepsilon\} \quad (3)$$

In (2)-(3), C and ε are the parameters of regularization and the error sensitivity, respectively. By substituting (3) for (2) and introducing the positive slack variables ξ_i, ξ_i^* to represent deviations from the ε -zone, the (2) can be reformulated as the objective function (4) subject to the constraints expressed in (5).

$$R = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (4)$$

$$\begin{aligned} y_i - (\omega^T \varphi(x_i) + b) &\leq \varepsilon + \xi_i \\ \omega^T \varphi(x_i) + b - y_i &\leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* &\geq 0; \quad i = 1, 2, \dots, N \end{aligned} \quad (5)$$

By applying the Lagrange multiplier and considering the optimally constrained formulation, (1) can be explicitly represented as (6). In this equation, α_i^* and α_i represent the Lagrange multipliers, and $K(x_i, x)$ denotes the Kernel function, which is defined as the dot product between $\varphi(x_i)^T$ and $\varphi(x)$:

$$f(x) = \sum_{i=1}^N (\alpha_i^* - \alpha_i) K(x_i, x) + b \quad (6)$$

One commonly used kernel function is the radial basis function (RBF), which is often employed with a Gaussian RBF kernel. In this paper, the authors have utilized the Gaussian RBF kernel for their study.

$$K(x, y) = e^{-\gamma \|x-y\|^2} \quad (7)$$

A deep learning model, including SVR, typically consists of two types of parameters [20]–[25]. The first type comprises the model parameters that are learned during the training process. The second type consists of hyperparameters, which are set before training and govern the behavior of the model. Based on the detailed analysis of the SVR model described above, there are three hyperparameters that significantly impact its performance. These include the ε parameter which defines the constraints of the function $f(x)$; the C parameter used to determine the balance between the regularization term and the empirical error; and the Kernel parameter. These hyperparameters have a strong influence on the accuracy of the SVR model.

2.2. The sliding window procedure and differencing the input data

The maximum daily load of electricity is a time series data that can be described in a following mathematical model:

$$y(t) = \{y_1, y_2, \dots, y_N, \dots\} \quad (8)$$

where $1, 2, \dots, N, \dots$ represents the order of observed variables. To apply the load forecasting model with SVR, a common method by utilizing the sliding window procedure to generate input and target data for both training and prediction is used. The sliding window procedure operates based on two parameters of the window size and stride. The window size shows the length of the window that is applied for the time series data. The stride parameter determines the shift of the window, which is referred to the distance between two consecutive windows, normally, with a default value of 1. Table 1 illustrates the operation process of the sliding window procedure with a window size of N and a stride of 1.

Table 1. The Sliding window procedure

Input	Target
y_1, y_2, \dots, y_N	y_{N+1}
y_2, y_3, \dots, y_{N+1}	y_{N+2}
...	...
$y_h, y_{h+1}, \dots, y_{N+h-1}$	y_{N+h}

Based on the operation process of the sliding window procedure, we can see that the window size parameter can potentially impact the accuracy of the SVR model. This is because it determines the amount of information included in the input data set. Therefore, in this study, the window size parameter will be investigated, and specifically, the value of window size = N will be examined to assess its impact on the SVR model's performance.

Regarding the repetitive nature due to the possible repetition of the maximum load on a specific day of a week with those on the same day of the previous weeks as mentioned above, the differencing procedure can also be applied to the input data to enhance the accuracy of the prediction results. The differencing of time series data is a technique used to transform the data by taking the differences between consecutive observations. It is commonly employed to remove trends and seasonality from the data, making it more stationary and suitable for analysis and modeling. The differencing process involves subtracting the previous observation from the current observation to calculate the difference. This can be done once (first differencing) or multiple times (higher-order differencing) depending on the characteristics of the data. The first differencing is performed by subtracting each observation from its preceding observation, as shown in (9). Meanwhile, the seasonal differencing can be utilized by taking the difference between an observation and the corresponding observation from the previous seasonal period, as shown in (10), where d represents the seasonal period which could be characterized for the length of a season or a specific time interval.

$$dy(y) = y(t) - y(t - 1) \quad (9)$$

$$dy(y) = y(t) - y(t - d) \quad (10)$$

Therefore, the parameter N in the sliding window and the parameter d in the differencing procedures can affect the prediction accuracy of the SVR model. Figure 1 presents the integration of the sliding window and differencing procedures applied to the SVR model, including the following main processes:

- Firstly, the data $([y_1, y_2, \dots, y_n])$ is split into training data $(y_1, y_2, \dots, y_{n-h})$ and testing data (y_{n-h+1}, \dots, y_n) , the testing and training data have lengths of h and $n-h$, respectively.
- The training data undergoes differencing in the input data through the diff process, and the working principle of the diff process is described by (9)-(10).
- The split 1 process separates the training data into the input and output sets for the training process, denoted as X_{train} and Y_{train} , respectively. On the other hand, the split 2 process separates the data into the input and output sets for testing, denoted as X_{test} and Y_{test} . The operating principles of the split 1 and split 2 processes are based on the operating principles of the sliding window procedure described in Table 1.

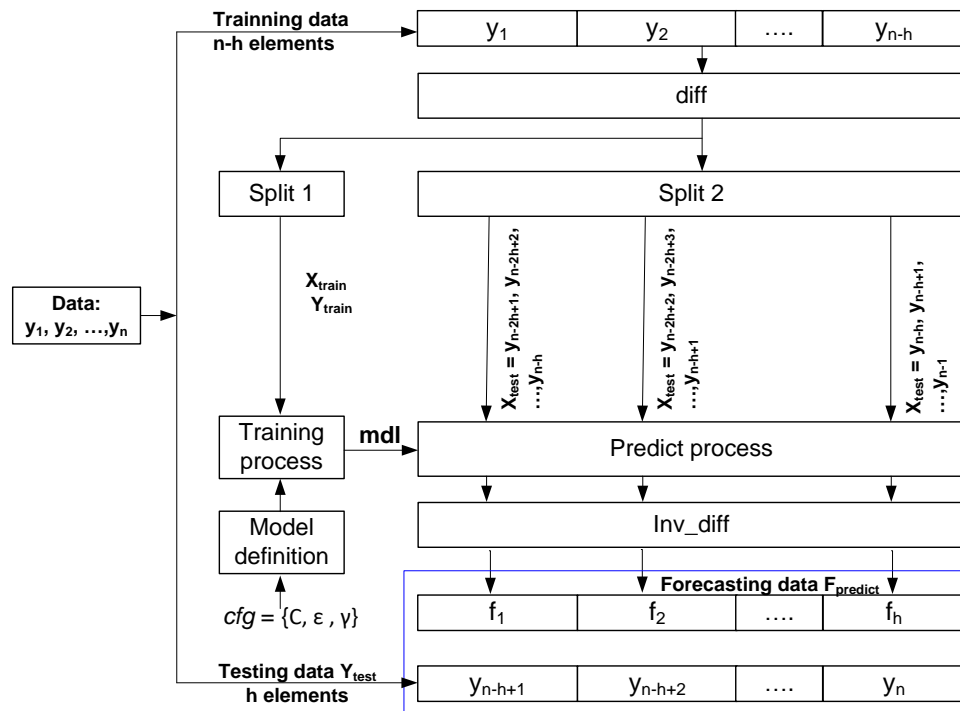


Figure 1. Methodology for the combination of the sliding window and differencing procedures

The X_{train} and Y_{train} data will be used for training the SVR network, and its trained model will be represented as mdl . This mdl model will be used as an input for the prediction process, along with input X_{test} , and the output will be the predicted values $F_{predict}$. The mean absolute percentage error (MAPE) is proposed in this study as the measure of the difference between the predicted values $F_{predict}$ and the actual values Y_{test} as defined by (11) [26], [27]. Analyzing the MAPE values with different values of N in the sliding window procedure and the value of d in the differencing procedure allows to evaluate their impact on the SVR network.

$$MAPE = \frac{1}{h} \sum_{i=1}^h \left| \frac{Y_{test} - F_{predict}}{Y_{test}} \right| \tag{11}$$

3. RESULTS AND DISCUSSION

3.1. Experimental settings

In this study, the maximum load data exported from the Victoria (VI) and Queensland (QL) grids in Australia in a period from June 6, 2010 to May 31, 2014 is used. The load profiles of these two datasets are depicted in Figure 2 with Victoria dataset in Figure 2(a) and Queensland dataset in Figure 2(b). To conduct the analysis, the datasets will be divided into two subsets: the training set and the test set. The test set will comprise the last 364 days of data, while the remaining data will be allocated for training purposes. Simultaneously, a large number of different combinations of the hyperparameters C , ϵ , and γ of the SVR

network are considered in this study to enhance the reliability of the research results. The range, values, and the number of surveys for these hyperparameters are presented in Table 2. The total number of combinations for the hyperparameters C , ε , and γ is 700 cases.

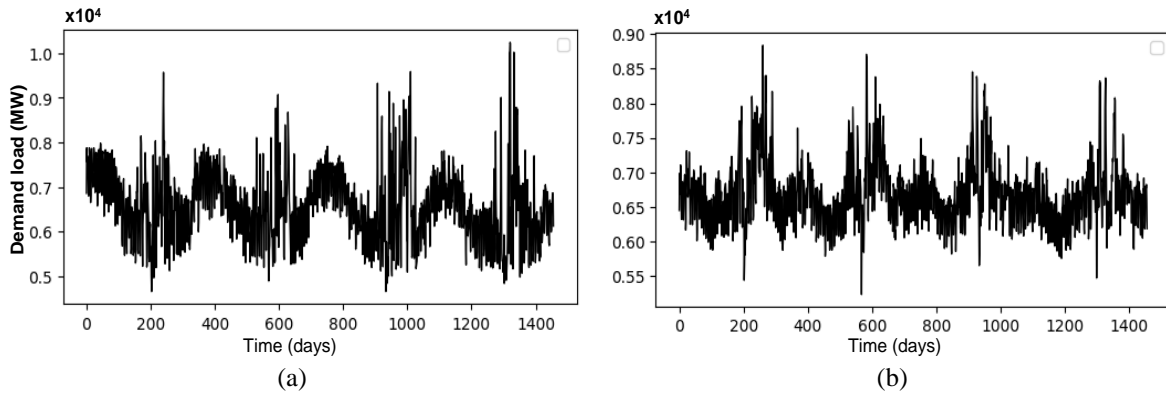


Figure 2. The maximum load data (a) Victoria and (b) Queensland

Table 2. The characteristic of hyperparameters

Hyperparameter	Values of elements	Number of elements
C	[1e-2, 4e-2, 7e-2, 1e-1, 4e-1, 7e-1, 1e0, 4e0, 7e0, 1e1]	A=10
ε	[1e-3, 4e-3, 7e-3, 1e-2, 4e-2, 7e-2, 1e-1]	B=7
γ	[1e-2, 4e-2, 7e-2, 1e-1, 4e-1, 7e-1, 1e0, 4e0, 7e0, 1e1]	C=10

3.2. Experimental results

The boxplot charts in Figures 3 and 4 present the statistical results of the forecasting errors for the VI and QL datasets, respectively. Each figure corresponds to a specific combination of N and d values. Tables 3 and 4 provide detailed statistical information regarding the respective error values shown in Figure 3 and Figure 4. Specifically, the first row values listed in Table 3 ($N=7$, $d=0$, mean=6.9, std=1.2, ..., 100th=11.1) correspond to the statistical measurements of the first column ($N=7$, $d=0$) in Figure 3, and so on. The similar correspondence is also made for Table 4 and Figure 4. Based on the results shown in Figure 3 and Figure 4 along with the statistical values of prediction errors provided in Tables 3 and 4, we can analyze the influence of the parameter N in the sliding window procedure and the parameter d in the differencing procedure on the SVR network.

Analyzing the influence of the parameter N . Specifically, for the case of $d=0$ and the VI dataset, we can observe that the different N values such as $N=7$, $N=14$, or $N=28$ yield similar prediction errors. The mean values of the errors are approximately 6.9 (for $N=7$), 6.7 (for $N=14$), and 6.9 (for $N=28$), respectively. The standard deviation values are approximately 1.1-1.2 for all three cases. The minimum values are 5.8, 5.6, and 5.6. The 25th percentile values are 6.1, 5.8, and 6. The 50th percentile values are 6.3, 6.2, and 6.6. The 75th percentile values are 7.4, 7.2, and 7.5. The maximum values are 11.1, 11.0, and 10.8. Similar results are obtained for the cases of $d=1$ and $d=7$, as well as when using the QL dataset. Therefore, with different N values, the statistical properties of the prediction errors remain almost the same. This indicates that the input N value has an insignificant impact on the prediction results in this study.

Analyzing the influence of the parameter d . Let consider $N=7$ for the VI dataset. When $d=1$, it shows the lowest statistical values among $d=0$ and $d=7$ in terms of prediction errors. Specifically, the mean value is 6.2 (for $d=1$) i.e. lower than 6.9 (for $d=0$) and 6.8 (for $d=7$), the standard deviation is 0.5 compared to 1.2 and 0.5, the minimum value is 5.5 compared to 5.8 and 6.2, the 25th percentile value is 5.9 compared to 6.1 and 6.4, the 50th percentile value is 6.1 compared to 6.3 and 6.6, the 75th percentile value is 6.4 compared to 7.4 and 7, and the maximum value is 7.5 compared to 11.1 and 8.2. Similar results are obtained for $N=14$ and 28, where $d=1$ also exhibits the lowest statistical values. Therefore, using differencing with $d=1$ as input can give the higher accuracy of the prediction results in comparison with the case of using the original data $d=0$ or $d=7$ in this study.

Figure 5 presents graphical illustrations of the prediction results for different combinations of N and d , corresponding to specific values of hyperparameters. Figure 5(a) corresponds to the VI dataset with hyperparameters $C=7e-2$, $\varepsilon=7e-3$, $\gamma=7e-2$. Figure 5(b) corresponds to the QL dataset with hyperparameters

$C=7e-1$, $\varepsilon=7e-2$, $\gamma=7e-1$. Observing that the prediction curves (depicted in dashed line) for the cases of $d=1$ closely align with the actual data obtained for the cases of $d=0$ and $d=7$. This observation aligns with the analyses mentioned above.

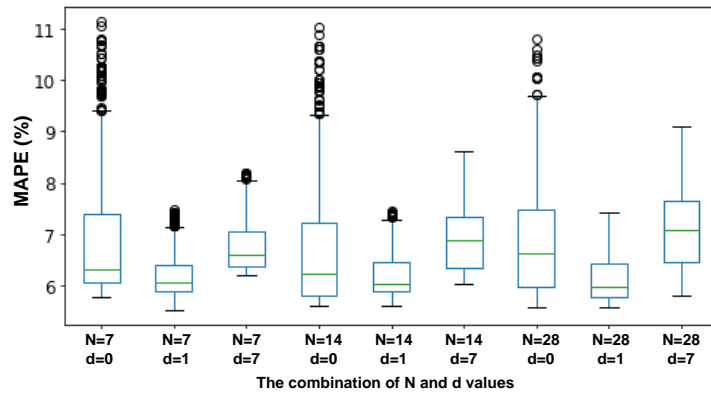


Figure 3. The boxplot of error rate for Victoria case

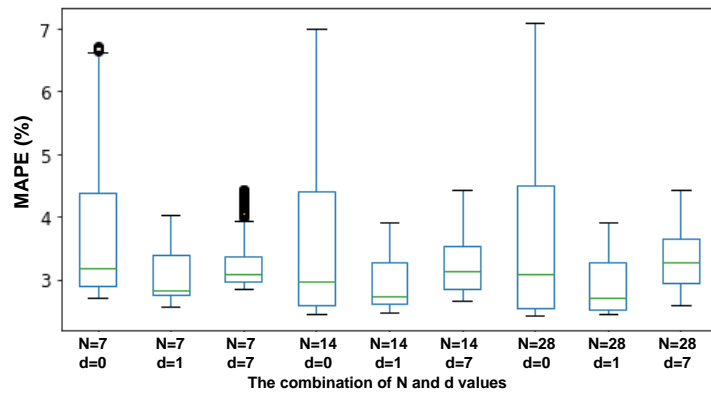


Figure 4. The boxplot of error rate for Queensland case

Table 3. The statistical characteristic of error rate MAPE (%) for VI case

N	d	mean	std	min	25 th	50 th	75 th	100 th
7	0	6.9	1.2	5.8	6.1	6.3	7.4	11.1
7	1	6.2	0.5	5.5	5.9	6.1	6.4	7.5
7	7	6.8	0.5	6.2	6.4	6.6	7	8.2
14	0	6.7	1.2	5.6	5.8	6.2	7.2	11
14	1	6.2	0.4	5.6	5.9	6	6.4	7.5
14	7	6.9	0.6	6	6.4	6.9	7.3	8.6
28	0	6.9	1.1	5.6	6	6.6	7.5	10.8
28	1	6.1	0.4	5.6	5.8	6	6.4	7.4
28	7	7.1	0.7	5.8	6.5	7.1	7.6	9.1

Table 4. The statistical characteristic of error rate MAPE (%) for QL case

N	d	mean	std	min	25 th	50 th	75 th	100 th
7	0	3.8	1.1	2.7	2.9	3.2	4.4	6.7
7	1	3	0.4	2.6	2.7	2.8	3.4	4
7	7	3.3	0.4	2.9	3	3.1	3.4	4.4
14	0	3.6	1.3	2.4	2.6	3	4.4	7
14	1	2.9	0.4	2.5	2.6	2.7	3.3	3.9
14	7	3.3	0.5	2.7	2.8	3.1	3.5	4.4
28	0	3.6	1.3	2.4	2.5	3.1	4.5	7.1
28	1	2.9	0.4	2.4	2.5	2.7	3.3	3.9
28	7	3.3	0.5	2.6	2.9	3.3	3.6	4.4

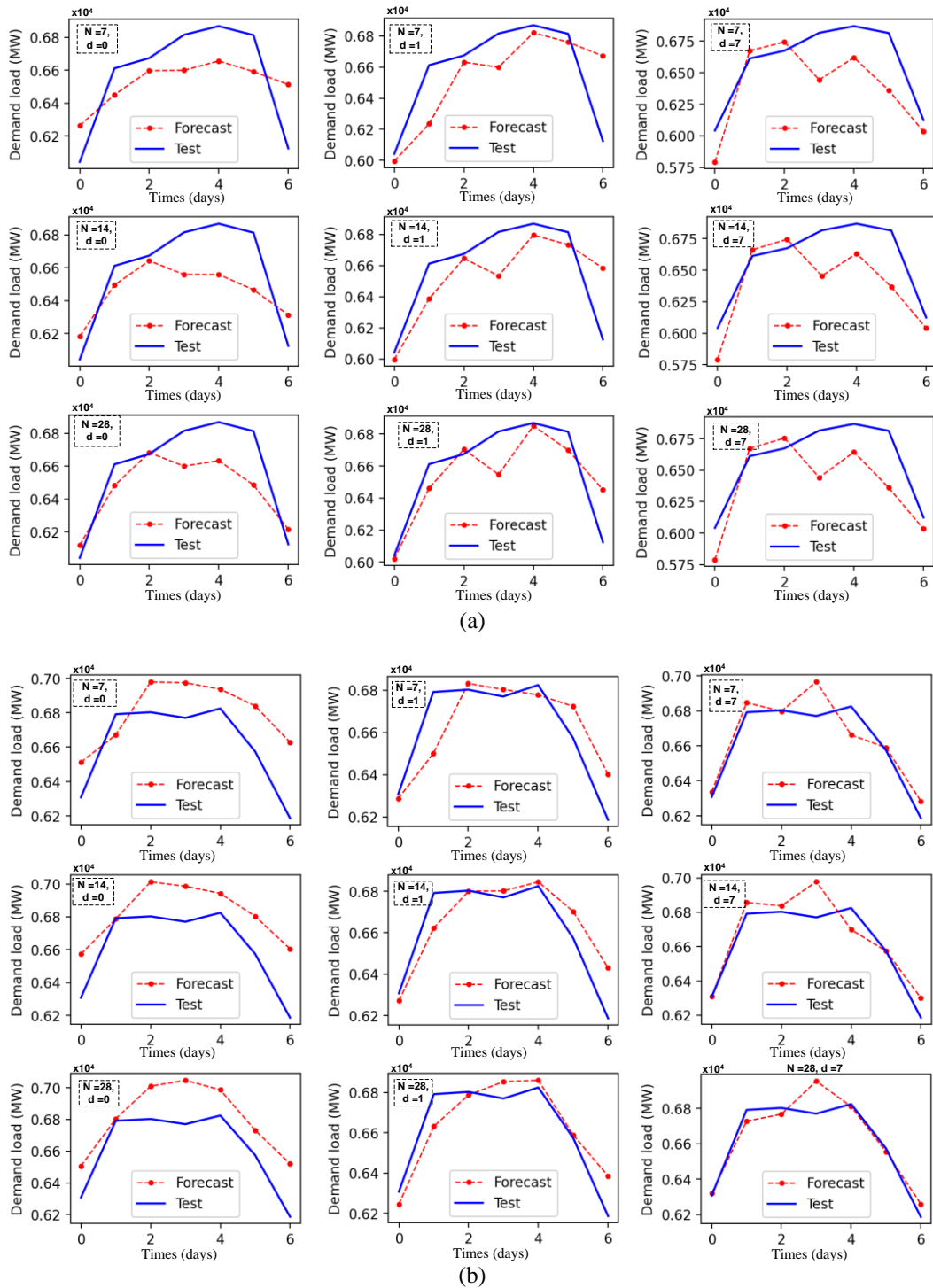


Figure 5. The plot of prediction and test values (a) VI and (b) QL

4. CONCLUSION

The study provided valuable insights into the impact of the N and d parameters on the SVR model and offers potential approaches for improving the accuracy of load forecasting and other predictive applications. The obtained results revealed that the impact on the prediction errors can be ignored in the case of changing the values of N . However, altering the values of d gives a strong influence. Specifically, in this study, using $d=1$ yields the highest accuracy in the prediction results in comparison with the original data ($d=0$) or the case with $d=7$. These findings are highly meaningful not only for the load forecasting problem of substations, but also for other forecasting applications.




ACKNOWLEDGEMENTS

This work is supported by Industrial University of Ho Chi Minh City under Grant No. 22/2D03.




REFERENCES

- [1] J. W. Taylor and P. E. McSharry, "Short-term load forecasting methods: An evaluation based on European data," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2213–2219, Nov. 2007, doi: 10.1109/TPWRS.2007.907583.
- [2] N. K. Thokala, A. Bapna, and M. G. Chandra, "A deployable electrical load forecasting solution for commercial buildings," in *2018 IEEE International Conference on Industrial Technology (ICIT)*, Feb. 2018, pp. 1101–1106, doi: 10.1109/ICIT.2018.8352332.
- [3] V. Gupta, "An overview of different types of load forecasting methods and the factors affecting the load forecasting," *International Journal for Research in Applied Science and Engineering Technology*, pp. 729–733, Apr. 2017, doi: 10.22214/ijraset.2017.4132.
- [4] I. K. Nti, M. Teimeh, O. Nyarko-Boateng, and A. F. Adekoya, "Electricity load forecasting: a systematic review," *Journal of Electrical Systems and Information Technology*, vol. 7, no. 1, Sep. 2020, doi: 10.1186/s43067-020-00021-8.
- [5] A. K. Singh, Ibraheem, S. Khatoon, M. Muazzam, and D. K. Chaturvedi, "Load forecasting techniques and methodologies: a review," in *2012 2nd International Conference on Power, Control and Embedded Systems*, Dec. 2012, pp. 1–10, doi: 10.1109/ICPES.2012.6508132.
- [6] A. Al Mamun, M. Sohel, N. Mohammad, M. S. Haque Sunny, D. R. Dipta, and E. Hossain, "A comprehensive review of the load forecasting techniques using single and hybrid predictive models," *IEEE Access*, vol. 8, pp. 134911–134939, 2020, doi: 10.1109/ACCESS.2020.3010702.
- [7] S. Kumar, L. Hussain, S. Banarjee, and M. Reza, "Energy load forecasting using deep learning approach-LSTM and GRU in spark cluster," *2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT)*, Kolkata, India, 2018, pp. 1–4, doi: 10.1109/EAIT.2018.8470406.
- [8] N. Ahmad, Y. Ghadi, M. Adnan, and M. Ali, "Load forecasting techniques for power system: research challenges and survey," *IEEE Access*, vol. 10, pp. 71054–71090, 2022, doi: 10.1109/ACCESS.2022.3187839.
- [9] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, Aug. 2004, doi: 10.1023/B:STCO.0000035301.49549.88.
- [10] Y. Chen *et al.*, "Short-term electrical load forecasting using the support vector regression (SVR) model to calculate the demand response baseline for office buildings," *Applied Energy*, vol. 195, pp. 659–670, Jun. 2017, doi: 10.1016/j.apenergy.2017.03.034.
- [11] M. Karasuyama and R. Nakano, "Optimizing SVR hyperparameters via fast cross-validation using AOSVR," in *2007 International Joint Conference on Neural Networks*, Aug. 2007, pp. 1186–1191, doi: 10.1109/IJCNN.2007.4371126.
- [12] A. Zakrani, A. Najm, and A. Marzak, "Support vector regression based on grid-search method for agile software effort prediction," in *Colloquium in Information Science and Technology, CIST*, Oct. 2018, pp. 492–497, doi: 10.1109/CIST.2018.8596370.
- [13] H.-L. Sun, R.-X. Jia, and Y.-F. Xue, "Customer demand forecasting based on SVR using moving time window method," in *2009 16th International Conference on Industrial Engineering and Engineering Management*, Oct. 2009, pp. 104–107, doi: 10.1109/ICIEEM.2009.5344626.
- [14] I. A. Khan, A. Akber, and Y. Xu, "Sliding window regression based short-term load forecasting of a multi-area power system," *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*, Edmonton, AB, Canada, 2019, pp. 1–5, doi: 10.1109/CCECE.2019.8861915.
- [15] S. Park, S. Jung, S. Jung, S. Rho, and E. Hwang, "Sliding window-based LightGBM model for electric load forecasting using anomaly repair," *Journal of Supercomputing*, vol. 77, no. 11, pp. 12857–12878, Apr. 2021, doi: 10.1007/s11227-021-03787-4.
- [16] Z. Xia, S. Xue, L. Wu, J. Sun, Y. Chen, and R. Zhang, "ForeXGBoost: passenger car sales prediction based on XGBoost," *Distributed and Parallel Databases*, vol. 38, no. 3, pp. 713–738, May 2020, doi: 10.1007/s10619-020-07294-y.
- [17] T. T. Ngoc, L. Van Dai, and C. M. Thuyen, "Support vector regression based on grid search method of hyperparameters for load forecasting," *Acta Polytechnica Hungarica*, vol. 18, no. 2, pp. 143–158, 2021, doi: 10.12700/APH.18.2.2021.2.8.
- [18] M. Bozic and M. Stojanovic, "Application of SVM methods for mid-term load forecasting," *Serbian Journal of Electrical Engineering*, vol. 8, no. 1, pp. 73–83, 2011, doi: 10.2298/sjee1101073b.
- [19] S. Lin, S. Zhang, J. Qiao, H. Liu, and G. Yu, "A parameter choosing method of SVR for time series prediction," in *2008 The 9th International Conference for Young Computer Scientists*, Nov. 2008, pp. 130–135, doi: 10.1109/ICYCS.2008.393.
- [20] S. Lessmann, R. Stahlbock, and S. F. Crone, "Optimizing hyperparameters of support vector machines by genetic algorithms," in *Proceedings of the 2005 International Conference on Artificial Intelligence*, 2005, vol. 1, pp. 74–80.
- [21] K. Smets, B. Verdonk, and E. M. Jordaan, "Evaluation of performance measures for SVR hyperparameter selection," in *2007 International Joint Conference on Neural Networks*, Aug. 2007, pp. 637–642, doi: 10.1109/IJCNN.2007.4371031.
- [22] Q. Huang, J. Mao, and Y. Liu, "An improved grid search algorithm of SVR parameters optimization," in *2012 IEEE 14th International Conference on Communication Technology*, Nov. 2012, pp. 1022–1026, doi: 10.1109/ICCT.2012.6511415.
- [23] J. T. Jeng, "Hybrid approach of selecting hyperparameters of support vector machine for regression," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 36, no. 3, pp. 699–709, Jun. 2006, doi: 10.1109/TSMCB.2005.861067.
- [24] D. Anguita, S. Ridella, F. Riviello, and R. Zunino, "Hyperparameter design criteria for support vector classifiers," *Neurocomputing*, vol. 55, no. 1–2, pp. 109–134, Sep. 2003, doi: 10.1016/S0925-2312(03)00430-2.
- [25] M. Jiang, S. Jiang, L. Zhu, Y. Wang, W. Huang, and H. Zhang, "Study on parameter optimization for support vector regression in solving the inverse ECG problem," *Computational and Mathematical Methods in Medicine*, vol. 2013, pp. 1–9, 2013, doi: 10.1155/2013/158056.
- [26] S. Kaewchada, S. Ruang-On, U. Kuhapong, and K. Songsri-In, "Random forest model for forecasting vegetable prices: a case study in Nakhon Si Thammarat Province, Thailand," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 5, pp. 5265–5272, Oct. 2023, doi: 10.11591/ijece.v13i5.pp5265-5272.
- [27] Mukhtar *et al.*, "Hybrid model in machine learning–robust regression applied for sustainability agriculture and food security," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 4, pp. 4457–4468, Aug. 2022, doi: 10.11591/ijece.v12i4.pp4457-4468.




BIOGRAPHIES OF AUTHORS

Thanh Ngoc Tran    received the B.Sc. and M.Sc. degrees from the Ho Chi Minh University of Technology, Vietnam in 2001 and 2006, respectively. He received his Ph.D. degree in information and electricity engineering from Hunan University, Changsha, China in 2012. Currently, he is a lecturer with the Faculty of Electrical Engineering Technology, Industry University of Ho Chi Minh City, Vietnam. His main research interests include HVDC transmission systems, active filters, power quality and load forecasting. He can be contacted at email: tranthanngoc@iuh.edu.vn.






Thi Phuc Dang    received the specialist degree from the Lomonosov Moscow State University, Moscow, Russia in 2008. In 2018 she received Ph.D. degree in system analysis, control and information processing from Peoples' Friendship University of Russia, Moscow, Russia. Since 2018, she is a lecturer of Faculty of information Technology, Industry University of Ho Chi Minh City, Vietnam. Her research interests include machine learning, computer vision, NLP, deep learning, operator network. She can be contacted at email: phucdt@iuh.edu.vn.



Binh Minh Lam    received the B.E. degree in electrical-electronics engineering from BachKhoa University, Vietnam, in 2002, and the M.S. degree in telecommunication and information systems from University of Essex, U.K, in 2004. She is currently a lecturer of the Faculty of Electrical Engineering Technology of the Industrial University of Ho Chi Minh City, Vietnam. She is also pursuing a Ph.D. degree at Department of Electronic Engineering, Pukyong National University, Busan, Korea. Currently she is a lecturer with the Faculty of Electrical Engineering Technology, Industry University of Ho Chi Minh City, Vietnam. Her current research interests include high-frequency radio frequency identification, far field communication, and low-power sensor networking. She can be contacted at email: lambinhminh@iuh.edu.vn.



Anh Tuan Nguyen    received the B.Sc. degrees from the Industrial University of Ho Chi Minh City, Vietnam in 2012 and M.Sc. degrees from the HCMC University of Technology and Education, Vietnam in 2015, respectively. Currently, he is a lecturer with the Faculty of Electrical Engineering Technology, Industry University of Ho Chi Minh City, Vietnam. His main research interests include HVDC transmission systems, active filters, power quality and load forecasting, Scada, PLC, automatic system. He can be contacted at email: nguyenanhtuan@iuh.edu.vn.