Generalized Support Vector Machines (GSVMs) model for real-world time series forecasting

Mehrnaz Ahmadi
Isfahan University of Technology
Mehdi Khashei (khashei@in.iut.ac.ir)
Isfahan University of Technology https://orcid.org/0000-0002-2607-2665

Research Article

Keywords: Support vector machines (SVMs), Fuzzy support vector machines (FSVMs), Preprocessing-based hybrid models, Time series forecasting, Wind power

Posted Date: April 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-180407/v1

Generalized Support Vector Machines (GSVMs) model for real-world time series forecasting

Mehrnaz Ahmadi, Mehdi Khashei*

Department of Industrial and Systems Engineering, Isfahan University of Technology (IUT), Isfahan 84156-83111, Iran.

Abstract

Support vector machines (SVMs) are one of the most popular and widely-used approaches in modeling. Various kinds of SVM models have been developed in the literature of prediction and classification in order to cover different purposes. Fuzzy and crisp support vector machines are a well-known branch of modeling approaches that are frequently applied for certain and uncertain modeling, respectively. However, each of these models can only be efficiently used in its specified domain and cannot yield appropriate and accurate results if the opposite situations have occurred. While the real-world systems and data sets often contain both certain and uncertain patterns that are complicatedly mixed together and need to be simultaneously modeled. In this paper, a generalized support vector machine (GSVM) is proposed that can simultaneously benefit the unique advantages of certain and uncertain versions of the traditional support vector machines in their own specialized categories. In the proposed model, the underlying data set is first categorized into two classes of certain and uncertain patterns. Then, certain patterns are modeled by a support vector machine, and uncertain patterns are modeled by a fuzzy support vector machine. After that, the function of the relationship, as well as the relative importance of each component, are estimated by another support vector machine, and subsequently, the final forecasts of the proposed model are calculated. Empirical results of wind speed forecasting indicate that the proposed method not only can achieve more accurate results than support vector machines (SVMs) and fuzzy support vector machines (FSVMs) but also can yield better forecasting performance than traditional fuzzy and nonfuzzy single models and traditional preprocessing-based hybrid models of SVMs.

Keywords: *Support vector machines (SVMs), Fuzzy support vector machines (FSVMs), Preprocessing-based hybrid models, Time series forecasting, Wind power.*

Conflict of interest: *The authors declare that they have no conflict of interest.*

1. Introduction

The support vector machines or support vector networks are among the most popular supervised learning machine approaches that have been frequently used for prediction and classification tasks. Numerous researchers have been attracted to use the support vector machines rather than other neural network models, due to some specific advantages of these models. Support vector machines, unlike artificial neural networks, don't need to choose numbers and sizes of hidden layers as well as activation functions to create a high-precision model. Furthermore, the support vector machines often have better accuracy and speed in resolving nonlinear problems [1]. For such reasons, in recent years, support vector machines have been widely used in the field of wind energy forecasting, including single and hybrid models. Mohandes *et al.* [2] have used a support vector machine model to predict wind speed. The time series of 12-year has been collected between 1970 and 1982 in Medina, Saudi Arabia in order to make reliable and accurate results. The results of the support vector machine and multilayer perceptron models have been compared and have been shown that the mean squares error of the SVM is less than the multilayer perceptron. Zhou *et al*. [3] have estimated the short-term wind speed using the least square support vector machine (LSSVM) in North Dakota based on the hourly wind speed data in the year 2002. In the analysis of this study, the data set has been divided into four seasonal data sets and a great effort to provide an accurate adjustment of the LSSVM model by taking three linear core functions, namely Linear, Gaussian, and Polynomial, and Poly functions. The effectiveness of the used model is examined by the mean squared error (MSE), and this comparison demonstrates the ability of the model compared to other models.

* *Corresponding author*.
Tel.: +98-311-3912550-1; Fax: +98-311-3915526
E-mail address: Khashei@cc.iut.ac.ir (M. Khashei).

Jiang *et al.* [4] have developed a hybrid model, including a variational mode decomposition-multi objective salp swarm algorithm and a least square support square machine (LSSVM) model. This model is then used for wind power forecasting by using wind speed data in China for 10 min, 30 min, and 60 min time horizons. By examining the results of the multi-objective optimization models with the single-objective optimization methods, it can be concluded that the multi-objective slap swarm algorithm has much more accurate than other models. Chen *et al.* [5] have used a hybrid model, including the singular spectrum analysis (SSA) and the LSSVM model for wind speed forecasting. In this study, three data sets are collected based on the wind speed data at the National Wind Technology Center (NWTC) of the National Renewable Energy Laboratory (NREL). The experimental results indicate that the proposed model performs better than other conventional methods. Hong *et al.* [6] have developed a high-frequency morphological filter (MHF) and a double similarity search (DSS) algorithm-least squares-support vector machine (LS-SVM) model for short-term wind power and speed forecasting. Results demonstrate that the MHF-DSS model provides more accurate and stable forecasts in comparison to the other methods.

Li *et al.* [7] have used a new multi-objective ant lion algorithm (MOALO) and LSSVM model for ultra-short-term wind speed forecasting. Numerical results show that the proposed model has higher accuracy than other hybrid models. Chen *et al.* [8] have used a hybrid model consisting of the Kalman filter and support vector regression (SVR) for short-term wind speed forecasting. The SVR-KF method is compared with artificial neural networks (ANNs), SVR, autoregressive (AR), and autoregressive integrated with Kalman filter (AR-Kalman) approaches. The forecasting results indicate that the proposed method has much better performance in both one-step-ahead and multi-step-ahead wind speed predictions than the other approaches across all locations. Jiang *et al.* [9] have introduced a new hybrid method (GARCH, LSSVM) for ultra-short-term wind speed forecasting. Results show that the proposed method has more satisfactory performance in both accuracy and stability than others. Guo *et al.* [10] have used a SARIMA- LSSVM model for ultra-short-term wind speed forecasting. To demonstrate the ability of the proposed model in order to model the complex time series, monthly data from January 2001 to December 2006 are used in the Masong and Jiukan Mountains. The numerical results show that the proposed model is efficient.

Yuan *et al*. [11] have used a hybrid model, incorporating an autoregressive fractionally integrated moving average and a least-squares support vector machine for short-term wind power forecasting. The autoregressive fractionally integrated moving average model is used to predict the linear components of the wind power series, and the least-squares support vector machine is applied to predict the nonlinear components of the data. In comparison with other models, simulation results show that the proposed hybrid model has the lowest error values, among other models. Khosravi *et al.* [12] have used the SVR-RBF model for 5 min, 10 min, 30 min, and one-hour wind speed and direction forecasting. In this paper, a large set of wind speed and wind direction data is utilized in order to predict the wind speed accurately and its direction at Bushehr. Numerical results indicate that the proposed SVR-RBF model can achieve the lowest statistical errors and the highest correlation coefficient for all time intervals. Despite the unique benefits of supportive vector machine models in solving nonlinear problems, these models have weaknesses in uncertainty environments with complex patterns. Although fuzzy support vector machines (FSVMs), have also been developed in the literature to solve problems in uncertain environments, no support vector machines model simultaneously models certain and uncertain problems. Articles published in the field of wind power and speed forecasting in recent years based on the support vector machines are shown in Table (1).

Table (1): The SVM based models used/developed in recently published articles to predict wind power and speed.

[Ref.]	Year	Domain	Time-scale	Applied Model(s)	Category
[13]	2019	Wind speed	Ultra short term	SSA, EMD, CNNSVM	Hybrid model
[14]	2018	Wind speed	Ultra short term	CS, V-SVM	Hybrid model
[15]	2018	Wind speed	Ultra short term	RNN, SVM, LSTM	Hybrid model
[16]	2017	Wind power	Medium term	BPNN, RBFNN, LSSVM	Hybrid model
[17]	2017	Wind speed	Ultra short term	SVM, GBM, RF	Hybrid model
[18]	2016	Wind power	Short term	VMD, LSSVM	Hybrid model
[19]	2016	Wind speed	Ultra short term	EMD, SVM	Hybrid model
[20]	2016	Wind power	Medium-term	ELM, SVM	Hybrid model
[21]	2015	Wind speed	Ultra short term	EWT, GSA, LSSVM	Hybrid model
[22]	2015	Wind power	Ultra short term	GSA, LSSVM	Hybrid model
[23]	2015	Wind speed	Short term	PSOGSA, C -LSSVM	Hybrid model
[24]	2015	Wind speed	Ultra short term	PSO, RSVM	Hybrid model
[25]	2015	Wind Power & Speed	Ultra short term	FNN, SVM	Hybrid model
[26]	2015	Wind speed	Short term	ERNN, SVM	Hybrid model
[27]	2014	Wind speed	Short term	WT, GA, SVM	Hybrid model
[28]	2012	Wind Speed	Ultra short term	LE-SVM	Single model

As mentioned, the use of different kinds of support vector machines, especially in the hybrid form, increases the accuracy of forecasts in wind power and wind speed fields. Although support vector machine models have effectively reduced the inconsistencies and weaknesses of artificial neural network models, they may have desired performance, especially in ambiguous patterns [29]. In order to eliminate the shortcomings of the support vector machine models in uncertain environments modeling, the fuzzy support vector machine (FSVM) models have been developed. The fuzzy support vector machines were first introduced by Lin *et al.* in 2002 [30]. Fuzzy support vector machines are one of the most widely-used uncertain models to solve fuzzy nonlinear problems in the subject literature. Although the fuzzy support vector machine models have a high ability to model uncertainty, these models can only model a specific part of the patterns in the data sets (e.g., uncertain patterns), efficiently. While the real data sets often have both certain and uncertain patterns [31]. For this reason, a more comprehensive model is needed that can simultaneously model certain and uncertain patterns.

Therefore, in this paper, a generalized support vector machine is presented, in which the support vector machines and fuzzy support vector machines are combined together. In the proposed model, the raw data set is first decomposed into certain and uncertain patterns by a preprocessing model. Then, certain patterns are modeled by a certain support vector machine, and uncertain patterns are modeled by a fuzzy support vector machine. At the last stage of the proposed model, the function of the relationship, as well as the relative importance of each component, are estimated by another support vector machine, and subsequently, the final forecasts of the proposed model are calculated. The remainder of the paper is organized as follows. In the next section, the methodology and the applied models as components of the proposed model are briefly introduced. In section 3, the explanation of the proposed model is discussed. In section 4, the wind power data sets and evaluation metrics are reviewed. In section 5, the numerical results of using the proposed model for wind power forecasting are reported, and the performance of the proposed model is compared with other models in accuracy. Some conclusions are offered in the last section.

2. Methodology

In this section, the details of the support vector machines (SVMs) are first introduced, then the details of the fuzzy support vector machines (FSVMs) are presented.

2.1. Support vector machines (SVMs)

A support vector machine is a machine learning algorithm constructed based on the statistical learning theory and the principle of structural risk reduction. Support vector machines were first proposed by Curtis and Vapnik in 1995 [32]. The SVMs have been successfully used for various purposes, such as image retrieval, error detection, text detection, and regression problems. The main idea of this approach is to convert the input space with a nonlinear region into a linear one with large dimensions. In the support vector regression (SVR) models, the support vector machine model is used to approximate the function and regression. Various kernel functions are used in support vector machine models, including polynomial function, radial basis function, sigmoid function, and the linear function. The training data set consisting of input and output pairs is considered $I_i = \{ (x_i, y_i) | i = 1,2,...,N \}$, where $x_i \in R^q$, that q is the dimension of the input vector, $y_i \in R$, the corresponding target value, and N refers to the size of the training data. The regression model is computed according to Eq. (1).

$$y = w^T \theta(x) + b \tag{1}$$

where, w is the weight vector, b is the bias component, and $\theta(x)$ represents a nonlinear mapping function that maps x to a higher dimensional pattern recognition. According to the risk minimization principle to obtain the w, and b the Eq. (1) is minimized by Eq. (2).

$$\begin{aligned} & min\{ \frac{1}{2} W^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i^*) \} \\ & S.T. \qquad y_i - \{ (W^T \theta(x_i)) + b \} \le \psi + \xi_i \qquad i = 1,2,...,N \\ & \qquad\qquad \xi_i, \xi_i^* \ge 0 \qquad i = 1,2,...,N \end{aligned} \tag{2}$$

where, ξ_i, ξ_i^* are the positive lag variables and C is the error parameter. Finally, the support vector regression is introduced by the Lagrangian coefficients (δ_i , δ^*_i) according to Eq. (3).

$$f(x) = \sum_{i=1}^{N} (\delta_i - \delta^*_i) k(x_i, x_j) + b \tag{3}$$

2.2. Fuzzy support vector machine SVM (FSVM)

Since in the traditional support vector machine methods, the training data sets get the same weights, the important data with unique properties are neglected. Therefore, in the fuzzy support vector machine method, by applied the fuzzy membership function (z_i), the training features set (S) are introduced according to Eq. (4) [29].

$$S = \{(x_1, y_1, z_1), \ldots, (x_i, y_i, z_i), (x_N, y_N, z_N)\} \tag{4}$$

That z_i, calculated by:

$$Z_i = \begin{cases} \left[\sum_{k=1}^{2} \sum_{j=1}^{2} \frac{\|x_i - o_j\|^{\frac{2}{N-1}}}{\|x_i - o_k\|^{\frac{2}{N-1}}} \right]^{-1} & , \quad \|x_i - o_k\| \neq 0 \\ 1 & , \quad \|x_i - o_k\| \neq 0 \quad k = j \\ \lambda & , \quad \|x_i - o_k\| \neq 0 \quad k \neq j \end{cases} \tag{5}$$

where, N is the weighted components, and o_j is the center of classes. After selecting the appropriate function kernel and parameter c, an optimization problem is solved according to Eq. (6).

$$\begin{aligned} &\min \quad -\alpha \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j k(x_i, x_j) \\ &s.t. \sum_{i=1}^{N} y_i \alpha_i = 0 \quad, 0 \leq \alpha_i \leq z_i c, \quad i = 1, 2, \ldots, N \end{aligned} \tag{6}$$

After selection the vector $\alpha^* = (\alpha^*_1, \alpha^*_2, \ldots, \alpha^*_N), b^*$ and the construct the decision function is calculated according to Eq. (7)-(8).

$$b^* = y_i - \sum_{i=1}^{N} \alpha_i^* y_i k(x_i - x_j) \tag{7}$$

$$f(x) = \sum_{i=1}^{N} \alpha^*_i y_i k(x_i, x_j) + b^* \tag{8}$$

3. The proposed generalized support vector machines (GSVMs)

Despite all the unique features of certain and uncertain support vector machines, they can only model certain and uncertain parts of the existing patterns in the underlying data. While real-world raw data sets basically have both patterns, simultaneously. In order to separately model certain and uncertain patterns in the subject literature, various models of support vector machines have been developed. Although these support vector machine-based models have yielded acceptable results in terms of performance and accuracy, none of them can simultaneously model the certainty and uncertainty in the data sets. In general, two possible techniques of 1) Simultaneously modeling both patterns and 2) Decomposing patterns, separately modeling, and finally combined, can be considered for lifting this task. In this paper, the second strategy is elected due to less complexity and lower computational time and cost. Therefore, in the

first stage of the proposed model, the raw data is preprocessed by the Kalman filter preprocessor for decomposing certain and uncertain patterns. After that, decomposed certain and uncertain data, along with the lags of the raw data, are used as input of the support vector machine and the fuzzy support vector machine for separately modeling patterns. In this way, certain patterns that often contain lower complexity and less ambiguity are entered to certain support vector machines, and uncertain patterns that often contain higher complexity and ambiguity are entered into the fuzzy support vector machines. In the last stage, the function of the relationship and the relative importance (weight) of these components are estimated by another support vector machine, and later, the final forecasts of the proposed model are calculated. The general process of the proposed method is presented in Fig. (1).

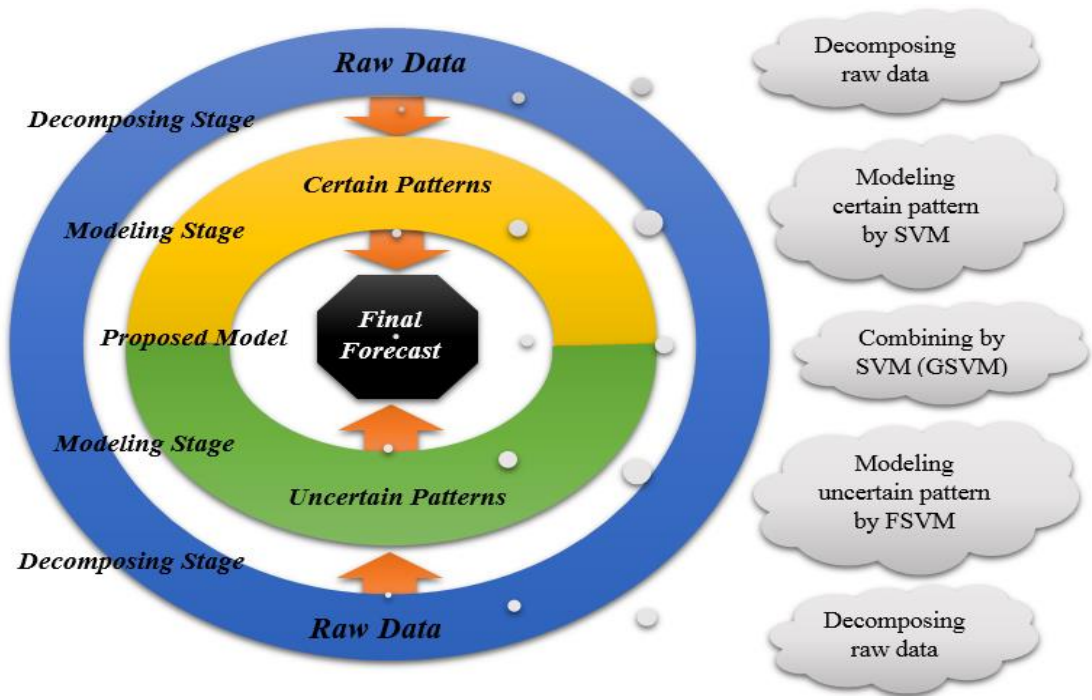


Fig. (1): The structure of the proposed model.

4. Data sets and evaluation metrics

In this section, the ultra-short-term wind power series, which is located in Sotavento Galicia, is chosen as the underlying data set (Fig. 2). This data set is benchmark data in the field of wind energy, which has been frequently used in the literature [33]. This time series is for the first week of May 2015, and totally consists of 168 points. The 85% of data (e.g., 144 observations) is applied as the training sample, and 15% of the remaining data (e.g., 24 observations) is used as the test sample in order to evaluate the performance of the proposed model in comparison with other models. Also, in this paper, three evaluation criteria are used to calculate the accuracy of each model. These criteria are MAE (Mean Absolute Error), MSE (Mean Squared Error), and RMSE (Root Mean Squared Error), which are calculated according to Eq. (9) to Eq. (11).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}(A_i - F_i) \tag{9}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(A_i - F_i\right)^2 \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(A_i - F_i\right)^2} \tag{11}$$

where, A_t is the actual value at time *t*, F_t is the forecasting value at time *t*, and N is the number of data.

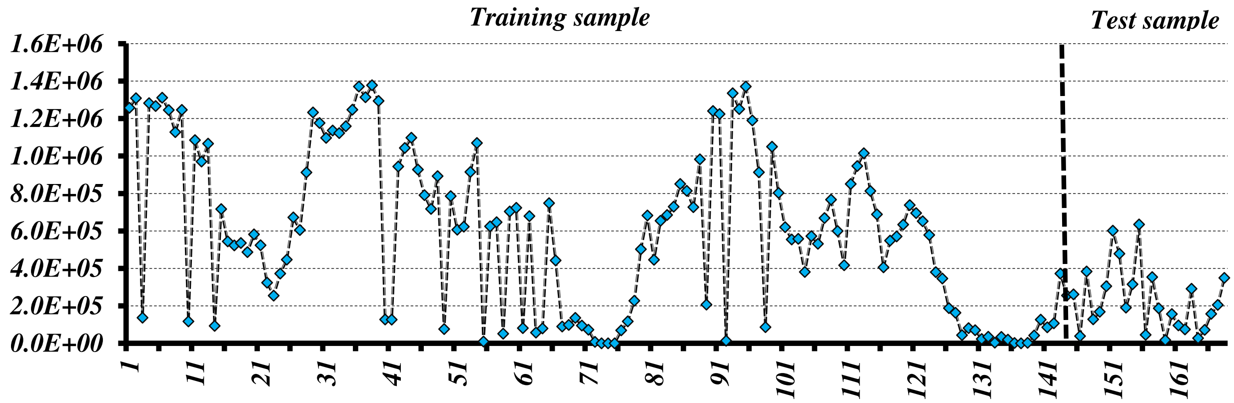


Fig. (2): The wind power data sets.

5. Numerical results of the proposed method for wind power forecasting

In this section, the training data set is used to design different models as well as the proposed model. After designing each model, the performance of each model is calculated using the criteria mentioned above. Ultimately, a comprehensive assessment of the performance of the methods compared with each other in the training data as well as in the data test has been conducted.

5.1. Results of the certain support vector machine (CSVM)

After decomposing of the underlying data, certain patterns in the data, along with the lags of the raw data, are entered into the SVM model. The performance indicators of the designed certain support vector machine (CSVM) model until each hour of the test day are shown in Fig. (3). It can be seen from Fig. (3) that the maximum and minimum values of MAE and MSE values are 5.92×10^4, 2.48×10^5, and 3.51×10^9, 1.35×10^{10}, respectively. The performance of the CSVM model is also presented in Table (2).

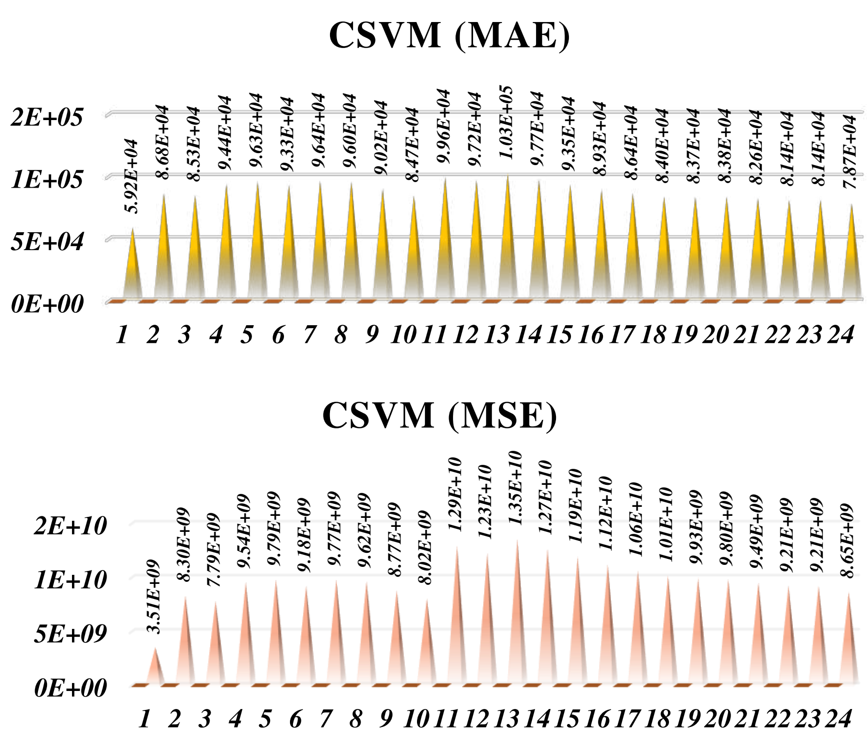


Fig. (3): The performance of the CSVM model for test data.

Table (2): The performance of the CSVM model in the test day.

Time (h)	Evaluation Characteristics		Time (h)	Evaluation Characteristics	
	MAE	MSE		MAE	MSE
1:00	5.92×10^4	3.51×10^9	**13:00**	1.03×10^4	1.35×10^{10}
2:00	8.68×10^4	8.30×10^9	**14:00**	9.77×10^4	1.27×10^{10}
3:00	8.53×10^4	7.79×10^9	**15:00**	9.35×10^4	1.19×10^{10}
4:00	9.44×10^4	9.54×10^9	**16:00**	8.93×10^4	1.12×10^{10}
5:00	9.63×10^4	9.79×10^9	**17:00**	8.64×10^4	1.06×10^{10}
6:00	9.33×10^4	9.18×10^9	**18:00**	8.40×10^4	1.01×10^{10}
7:00	9.64×10^4	9.77×10^9	**19:00**	8.37×10^4	9.93×10^9
8:00	9.60×10^4	9.62×10^9	**20:00**	8.38×10^4	9.80×10^9
9:00	9.02×10^4	8.77×10^9	**21:00**	8.26×10^4	9.49×10^9
10:00	8.47×10^4	8.02×10^9	**22:00**	8.14×10^4	9.21×10^9
11:00	9.96×10^4	1.29×10^{10}	**23:00**	8.14×10^4	9.21×10^9
12:00	9.72×10^4	1.23×10^{10}	**24:00**	$\mathbf{7.87\times10^4}$	$\mathbf{8.65\times10^9}$

Results of Table (2) indicate that the CSVM model can totally achieve 7.87×10^4, 8.65×10^9, and 9.30×10^4 in MAE, MSE, and RMSE in the whole test day, respectively. The error values of the CSVM model for training and test data sets are also shown in Fig. (4). It can be visually concluded from these error values that the support vector machine employed for modeling certain patterns can appropriately do it and can yield satisfactory results. It can demonstrate that the data generation process of the underlying data contains certain patterns.

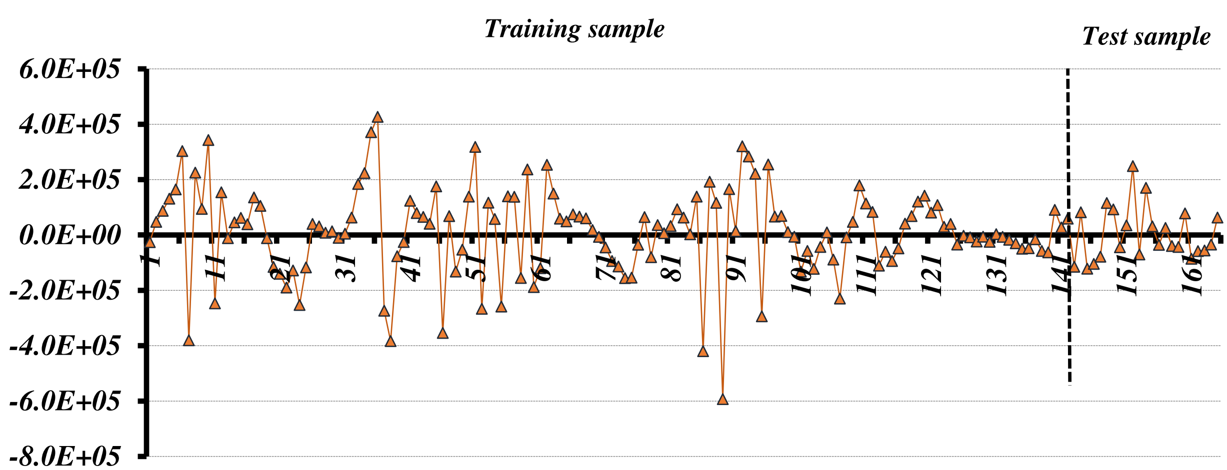
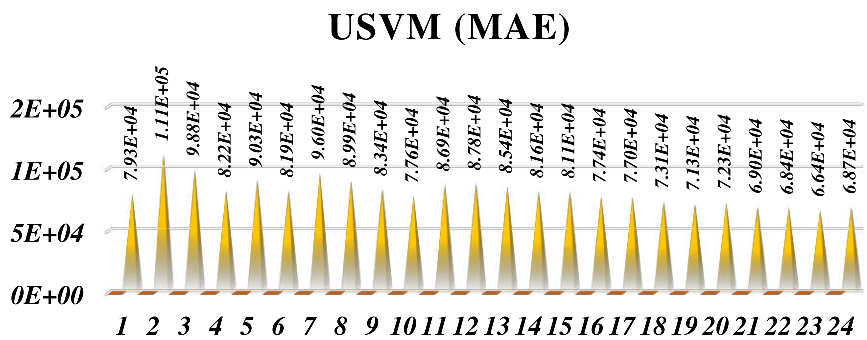


Fig. (4): Errors of the CSVM model for training and test data sets.

5.2. Results of the fuzzy support vector machine (USVM)

In a similar fashion, the uncertain patterns in the data, along with the lags of the raw data, are entered into the fuzzy support vector machine model. The performance indicators of the designed uncertain support vector machine (USVM) model until each hour of the test day are shown in Fig. (5). Results show that the maximum and minimum values of MAE and MSE values are 6.64×10^4, 1.11×10^5, and 6.26×10^9, 1.34×10^{10}, respectively. The performance of the USVM model is also reported in Table (3).



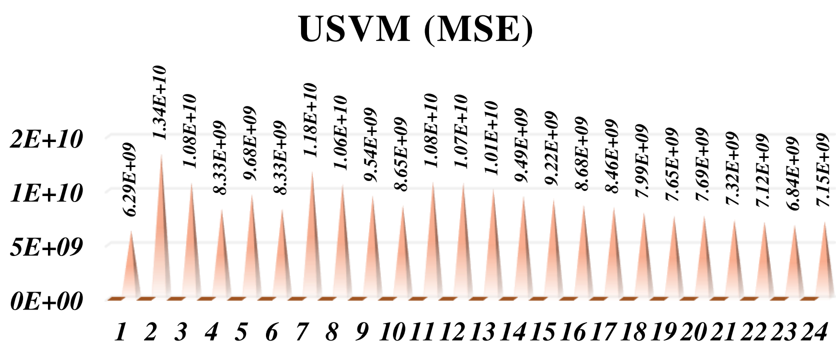


Fig. (5): The performance of the USVM model for test data.

Table (3): The performance of the USVM model in the test day.

Time (h)	Evaluation Characteristics		Time (h)	Evaluation Characteristics	
	MAE	MSE		MAE	MSE
1:00	7.93×10^4	6.29×10^9	**13:00**	8.54×10^4	1.01×10^{10}
2:00	1.11×10^5	1.34×10^{10}	**14:00**	8.16×10^4	9.49×10^9
3:00	9.88×10^4	1.08×10^{10}	**15:00**	8.11×10^4	9.22×10^9
4:00	8.22×10^4	8.33×10^9	**16:00**	7.74×10^4	8.68×10^9
5:00	9.03×10^4	9.68×10^9	**17:00**	7.70×10^4	8.46×10^9
6:00	8.19×10^4	8.33×10^9	**18:00**	7.31×10^4	7.99×10^9
7:00	9.60×10^4	1.18×10^{10}	**19:00**	7.13×10^4	7.65×10^9
8:00	8.99×10^4	1.06×10^{10}	**20:00**	7.23×10^4	7.69×10^9
9:00	8.34×10^4	9.54×10^9	**21:00**	6.90×10^4	7.32×10^9
10:00	7.76×10^4	8.65×10^9	**22:00**	6.84×10^4	7.12×10^9
11:00	8.69×10^4	1.08×10^{10}	**23:00**	6.64×10^4	6.84×10^9
12:00	8.78×10^4	1.07×10^{10}	**24:00**	$\mathbf{6.87\times10^4}$	$\mathbf{7.15\times10^9}$

Results of Table (3) indicate that the USVM model can achieve 6.87×10^4, $7.15\times10^{9,}$ and 8.64×10^4 in MAE, MSE, and RMSE in the whole test day, respectively. The error values of the USVM model for training and test data sets are also shown in Fig. (6). It can be visually determined from these error values that the fuzzy support vector machine can acceptability model uncertain patterns in the data. It can demonstrate that the data generation process of the underlying data may include uncertain patterns in addition to certain ones.

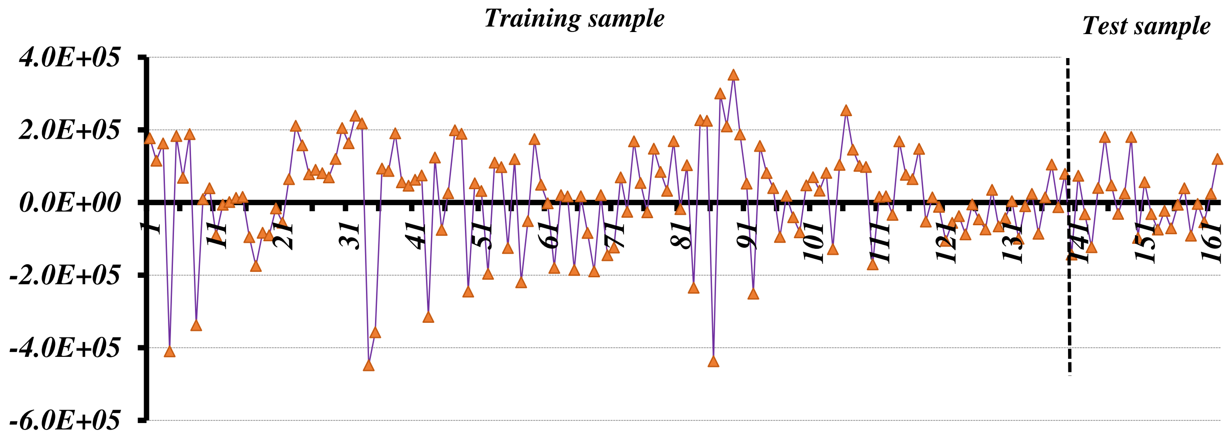


Fig. (6): Errors of the USVM model for training and test data sets.

5.3. Results of the proposed model (GSVM)

After modeling certain and uncertain patterns in the underlying data, based on the procedure of the proposed model, the relationship between these two types of patterns, as well as their relative importance to produce the actual values must be functionally estimated. Thus, in the last step of the proposed model, obtained results of the CSVM and USVM, along with the lags of the raw data, are entered into a support vector machine, and the final forecasts of the proposed method are calculated. These forecasting results using evaluation metrics are presented in Fig. (7).

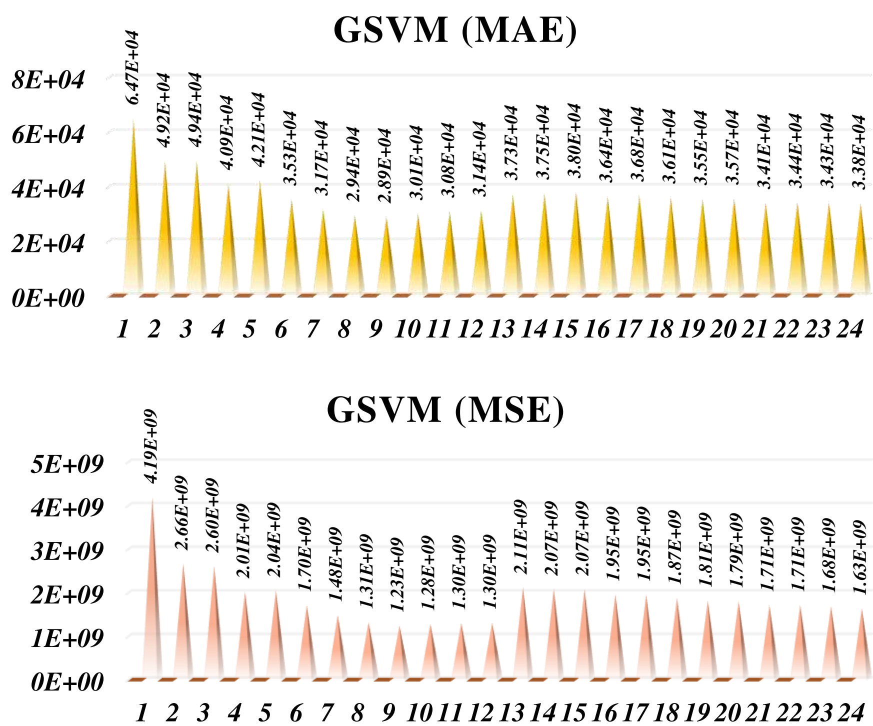


Fig. (7): The performance of the proposed model.

The actual and fitted values of the proposed model for train and test samples are shown in Fig. (8). The evaluation metrics for the proposed model for test day are reported in Table (4). Results of MAE, MSE, and RMSE for this model are 3.38×10^4, 1.63×10^9, and 4.04×10^4, respectively. Also, the error values of the proposed model for the train and test data sets are also shown in Fig. (9).

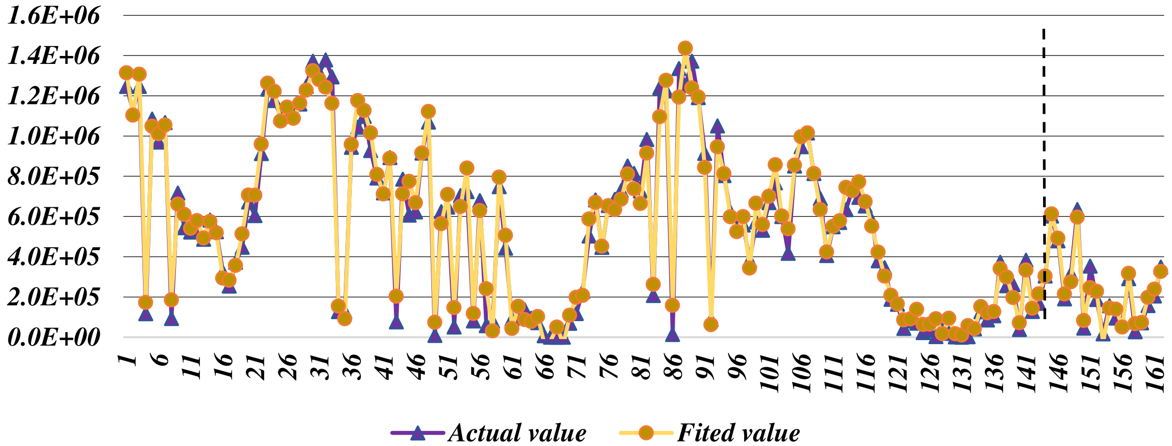


Fig. (8): Actual and fitted values of the proposed model for train and test samples.

Table (4): Results of the proposed model on a test day.

Time (h)	Evaluation Characteristics		Time (h)	Evaluation Characteristics	
	MAE	MSE		MAE	MSE
1:00	6.47×10^4	4.19×10^9	**13:00**	3.73×10^4	2.11×10^9
2:00	4.92×10^4	2.66×10^9	**14:00**	3.75×10^4	2.07×10^9
3:00	4.94×10^4	2.60×10^9	**15:00**	3.80×10^4	2.07×10^9
4:00	4.09×10^4	2.01×10^9	**16:00**	3.64×10^4	1.95×10^9
5:00	4.21×10^4	2.04×10^9	**17:00**	3.68×10^4	1.95×10^9
6:00	3.53×10^4	1.70×10^9	**18:00**	3.61×10^4	1.87×10^9
7:00	3.17×10^4	1.48×10^9	**19:00**	3.55×10^4	1.81×10^9
8:00	2.94×10^4	1.31×10^9	**20:00**	3.57×10^4	1.79×10^9
9:00	2.89×10^4	1.23×10^9	**21:00**	3.41×10^4	1.71×10^9
10:00	3.01×10^4	1.28×10^9	**22:00**	3.44×10^4	1.71×10^9
11:00	3.08×10^4	1.30×10^9	**23:00**	3.43×10^4	1.68×10^9
12:00	3.14×10^4	1.30×10^9	**24:00**	$\mathbf{3.38\times10^4}$	$\mathbf{1.63\times10^9}$

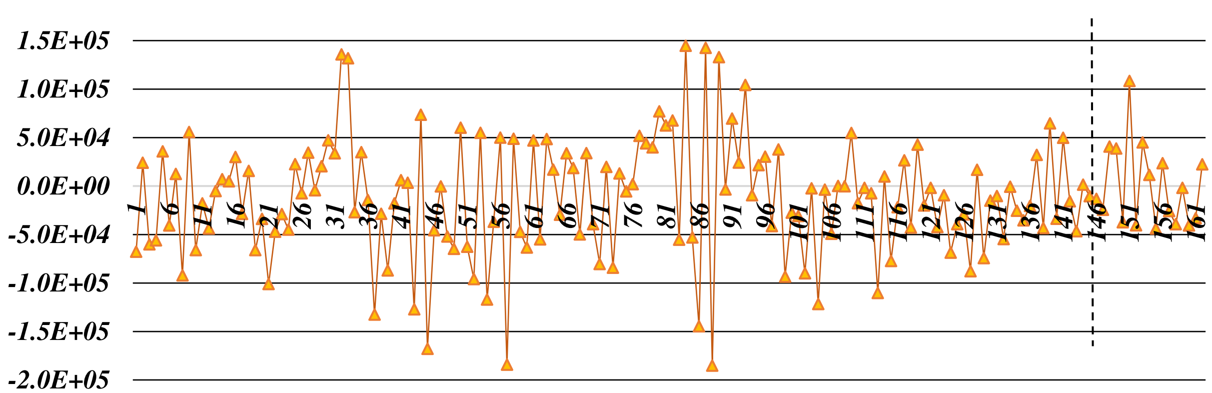


Fig. (9): Errors of the proposed model for train and test data sets.

5.4. Comparison with other models

In this section, the performance of the proposed model in the test and train data sets are compared with other forecasting models. These models involve the autoregressive integrated moving average (ARIMA), the general regression neural networks (GRNNs), the radial basis functions (RBFs), the multilayer perceptrons (MLPs), the support vector machines (SVMs), the fuzzy autoregressive integrated moving average (FARIMA), the fuzzy multilayer perceptrons (FMLPs), the fuzzy support vector machines (FSVMs), the Kalman preprocessing based hybrid model of ARIMA (KARIMA), the Kalman preprocessing based hybrid model of MLP (KMLP), the Kalman preprocessing based hybrid model of SVM (KSVM), the Kalman preprocessing based hybrid model of FARIMA (KFARIMA), the Kalman preprocessing based hybrid model of FMLP (KFMLP), the Kalman preprocessing based hybrid model of FSVM (KFSVM). The performance of the models above and improvement percentages of the proposed model against them in the train and test data in MAE and RMSE criteria are summarized in Table (5-6), respectively. Numerical results show that the proposed model can improve 58.10% and 57.05% the performance of the CSVM model in MAE and 60.06% and 56.56% in RMSE in train and test data sets, respectively. Moreover, the proposed model can improve 56.61% and 50.80% the performance of the USVM model in MAE and 56.78% and 52.25% in RMSE in train and test data sets, respectively.

Table (5): Forecasting performance of different models in training and test data sets.

Model	MAE		RMSE	
	Train	Test	Train	Test
ARIMA	2.11×10^5	1.47×10^5	2.99×10^5	1.75×10^5
GRNN	1.93×10^5	1.41×10^5	2.67×10^5	1.62×10^5
RBF	1.76×10^5	1.31×10^5	2.46×10^5	1.54×10^5
MLP	1.58×10^5	1.10×10^5	2.24×10^5	1.31×10^5
SVM	1.50×10^5	1.05×10^5	2.13×10^5	1.25×10^5
CSVM	1.16×10^5	7.87×10^4	1.58×10^5	9.30×10^4
USVM	1.12×10^5	6.87×10^4	1.46×10^5	8.46×10^4
FARIMA	1.85×10^5	1.13×10^5	2.40×10^5	1.39×10^5
FMLP	1.40×10^5	8.54×10^4	1.82×10^5	1.05×10^5
FSVM	1.37×10^5	8.37×10^4	1.78×10^5	1.03×10^5
KARIMA	1.44×10^5	1.06×10^5	2.01×10^5	1.34×10^5
KMLP	7.07×10^4	4.88×10^4	9.19×10^4	5.99×10^4
KSVM	6.27×10^4	4.33×10^4	8.12×10^4	5.27×10^4
KFARIMA	1.39×10^5	1.02×10^5	1.95×10^5	1.30×10^5
KFMLP	6.64×10^4	4.59×10^4	8.64×10^4	5.63×10^4
KFSVM	5.85×10^4	4.06×10^4	7.43×10^4	5.21×10^4
Proposed	$\mathbf{4.86\times10^4}$	$\mathbf{3.38\times10^4}$	$\mathbf{6.31\times10^4}$	$\mathbf{4.04\times10^4}$

Table (6): Improvement percentages of the proposed model against other models.

Model	MAE		RMSE	
	Train	Test	Train	Test
ARIMA	76.97%	77.01%	78.90%	76.91%
GRNN	74.82%	76.03%	76.37%	75.06%
RBF	72.39%	74.20%	74.35%	73.77%
MLP	69.24%	69.27%	71.83%	69.16%
SVM	67.60%	67.81%	70.38%	67.68%
CSVM	58.10%	57.05%	60.06%	56.56%
USVM	56.61%	50.80%	56.78%	52.25%
FARIMA	73.73%	70.09%	73.71%	70.94%
FMLP	65.29%	60.42%	65.33%	61.52%
FSVM	64.53%	59.62%	64.55%	60.78%
KARIMA	66.25%	68.11%	68.61%	69.85%
KMLP	31.26%	30.74%	31.34%	32.55%
KSVM	22.49%	21.94%	22.29%	23.34%
KFARIMA	65.04%	66.86%	67.64%	68.92%
KFMLP	26.81%	26.36%	26.97%	28.24%
KFSVM	16.92%	16.75%	15.07%	22.46%

In addition, empirical results indicate that the proposed model can improve 67.60% and 67.81% the performance of the SVM model in MAE and 70.38% and 67.68% in RMSE in train and test data sets, respectively. Furthermore, the proposed model can improve 64.53% and 59.62% the performance of the FSVM model in MAE and 64.55% and 60.78% in RMSE in train and test data sets, respectively. Finally, in order to a more comprehensive comparison of the proposed model with other forecasting models, the above-mentioned models are categorized into different classes as follows. The improvement percentages of the proposed model in comparison with these classes of models are given in Table (7).

1) *SCLS*: Statistical Certain Linear Single model(s) (e.g., ARIMA),

2) *ICNS*: Intelligent Certain Nonlinear Single model(s) (e.g., GRNN, RBF, MLP, and SVM),

3) *SULS*: Statistical Uncertain Linear Single model(s) (e.g., FARIMA),

4) *IUNS*: Intelligent Uncertain Nonlinear Single model(s) (e.g., FMLP, and FSVM),

5) *PCSH*: Preprocessing Certain Statistical Hybrid model(s) (e.g., KARIMA),

6) *PCIH*: Preprocessing Certain Intelligent Hybrid model(s) (e.g., KMLP and KSVM),

7) *PUSH*: Preprocessing Uncertain Statistical Hybrid model(s) (e.g., KFARIMA),

8) *PUIH*: Preprocessing Uncertain Intelligent Hybrid model(s) (e.g., KFMLP and KFSVM),

9) *CLNS*: Certain Linear Nonlinear Single model(s) (e.g., ARIMA, GRNN, RBF, MLP, and SVM),

10) *ULNS*: Uncertain Linear Nonlinear Single model(s) (e.g., FARIMA, FMLP, and FSVM),

11) *PCH*: Preprocessing Certain Hybrid model(s) (e.g., KARIMA, KMLP, and KSVM), and

12) *PUH*: Preprocessing Uncertain Hybrid model(s) (e.g., KFARIMA, KFMLP, and KFSVM).

The numerical results of Table (7) show that the proposed model can achieve more accurate and more reliable results than all mentioned classes of models in MAE and RMSE in both training and test samples. The proposed model can averagely improve these models 56.58% in MAE, and 57.73% in RMSE in train and test data sets, respectively. The results show that the most improvement is related to the statistical certain linear single model (e.g., ARIMA), that is 76.99%, and 77.91% in MAE, and RMSE in train and test data sets, respectively. Also, the result of Table (7) indicates that certain linear nonlinear single models (e.g., ARIMA, GRNN, RBF, MLP, and SVM), and statistical uncertain linear single model (e.g., FARIMA) are in second and third place in terms of accuracy, respectively. The proposed model can averagely improve CLNS, and SULS models, 72.53%, 73.45%, and 71.91%, 72.33% in MAE and RMSE in train and test data sets, respectively.

Table (7): Improvement percentages of the proposed model against other classes of models.

Classes	**MAE**		**RMSE**	
	Train	**Test**	**Train**	**Test**
1- SCLS models	76.97%	77.01%	78.90%	76.91%
2- ICNS models	71.01%	71.83%	73.23%	71.42%
3- SULS models	73.73%	70.09%	73.71%	70.94%
4- IUNS models	64.91%	60.02%	64.94%	61.15%
5- PCSH models	66.25%	68.11%	68.61%	69.85%
6- *PCIH models*	26.88%	26.34%	26.82%	27.95%
7- *PUSH models*	65.04%	66.86%	67.64%	68.92%
8- *PUIH models*	21.87%	21.56%	21.02%	25.35%
9- *CLNS models*	72.20%	72.86%	74.37%	72.52%
10- *ULNS models*	67.85%	63.38%	67.86%	64.41%
11- *PCH models*	40.00%	40.26%	40.75%	41.91%
12- *PUH models*	36.26%	36.66%	36.56%	39.87%
AVERAGE	**56.91%**	**56.25%**	**57.87%**	**57.60%**

Also, the result of Table (7) indicates that the lowest values of the improvement are related to the preprocessing uncertain intelligent hybrid models (e.g., KFMLP and KFSVM). The proposed model can averagely improve the predictive performance of the 1PUIH models, 21.71%, and 23.19% in MAE and RMSE in train and test data sets, respectively. Finally, in order to show the accuracy of the proposed model, the error of the proposed model compared to its components and base models is shown in Fig. (10-11). As can be seen, the proposed model has obtained more accurate and reliable results than SVM, FSVM, CSVM, and USVM models.

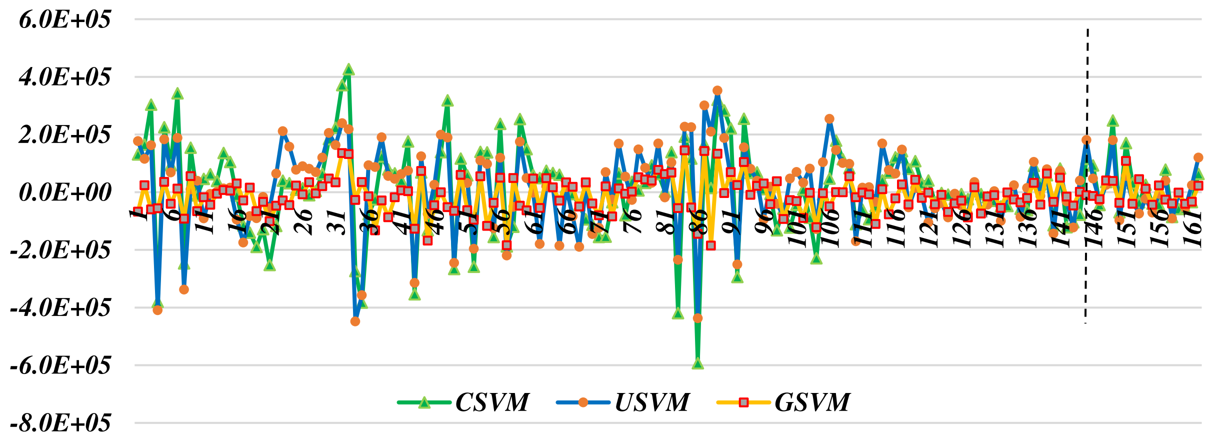


Fig. (10): The proposed model errors against its components in the test and train samples.

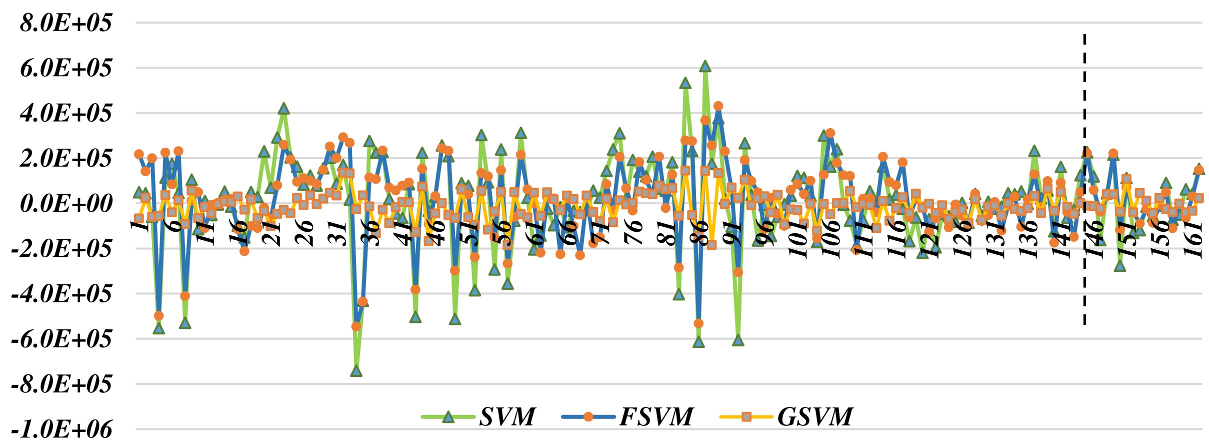


Fig. (11): The proposed model errors against base models in the test and train samples.

The statistical characteristics of Fig. (10-11) show that the proposed generated support vector machine (GSVM) model errors in the training sample are: Mean: -1.36×10^4, Min: -1.85×10^5, Max: 1.45×10^5, Standard Deviation: 6.16×10^4. While these characteristics for the support vector machine (SVM) are: Mean: 1.39×10^4, Min: -7.41×10^5, Max: 6.08×10^5, Standard Deviation: 2.14×10^5; for fuzzy support vector machine (FSVM) are: Mean: 1.16×10^4, Min: -6.45×10^5, Max: 5.29×10^5, Standard Deviation: 1.86×10^5, for certain support vector machine (CSVM) are: Mean: 1.27×10^4, Min: -5.93×10^5, Max: 4.27×10^5, Standard Deviation: 1.58×10^5, and for uncertain support vector machine (USVM) are: Mean: 1.57×10^4, Min: -4.48×10^5, Max: 3.52×10^5, Standard Deviation: 1.45×10^5 , respectively. Moreover, the statistical characteristics of the proposed model errors in the test sample are: Mean: 5.72×10^1, Min: -4.65×10^4, Max: 1.08×10^5, Standard Deviation: 4.04×10^4. While these characteristics for the support vector machine are: Mean: 1.21×10^4, Min: -2.76×10^5, Max: 2.25×10^5, Standard Deviation: 1.24×10^5; and for and fuzzy support vector machine are: Mean: 1.10×10^4, Min: -2.41×10^5, Max: 1.95×10^5, Standard Deviation: 1.08×10^5, for certain support vector machine (CSVM) are: Mean: 4.93×10^3, Min: -1.22×10^5, Max: 2.49×10^5, Standard Deviation: 9.49×10^4, and for uncertain support vector

machine (USVM) are: Mean: 3.59×10^3, Min: -1.23×10^5, Max: 1.81×10^5, Standard Deviation: 8.34×10^4, respectively. These results clearly indicate that the proposed model can yield more accurate and more reliable results than both base models (e.g., support vector machines (SVMs), and fuzzy support vector machines (FSVMs)), and its components models (e.g., certain support vector machines (CVMs), and uncertain support vector machines (UCVMs)).

6. Conclusion

literature shows that it is not possible to provide a compatible model globally to predict wind energy because of variable patterns that vary based on a variety of factors. The support vector machine is a nonlinear model that is a widely-used technique to predict and classification of various data, especially wind speed and wind power. Despite the unique advantages of support vector machines in solving nonlinear problems, these models have a fundamental weakness in the uncertainty modeling in ambiguous environments, accurately. Fuzzy support vector machines have been proposed to eliminate the weakness of traditional SVM models and to appropriately model the uncertain patterns. Although each of the support vector machines and the fuzzy support vector machines separately has a high capacity in certainty and uncertainty modeling, the real-world data sets often have certain and uncertain patterns simultaneously. Therefore, none of them can properly be used, and a new version of the support vector machines is needed that can simultaneously model the certainty and uncertainty in the data. For these reasons, in this paper, a Generalized Support Vector Machine (GSVM) model is proposed to predict certain and uncertain patterns simultaneously. Therefore, after decomposing the input data into certain and uncertain patterns, certain patterns are modeled by a support vector machine, and uncertain patterns are modeled by a fuzzy support vector machine. After that, the weight of each model is calculated by another support vector machine. Numerical results show that the proposed model can achieve more accurate results than its components, base models, and other single and hybrid models.

Authorship contributions: Mehrnaz Ahmadi and Mehdi Khashei equally contributed to the design and implementation of the research, to the analysis of the results, and to the writing of the manuscript.

Reference

[1] Ling-Ling Li, Xue Zhao, Ming-Lang Tseng, Raymond R. Tan, "Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm", Journal of Cleaner Production, 242, 118447, 2020.

[2] Mohandes, M.A., Halawani, T.O., Rehman, S., Hussain, A.A., "Support vector machines for wind speed prediction", Renewal Energy, 29, 939-947, 2004.

[3] Zhou, J., Shi, J., Li, G., "Fine tuning support vector machines for short-term wind speed forecasting", Energy Conversion and Management, 52, 1990-1998, 2011.

[4] Ping Jiang, Ranran Li, Hongmin Li, "Multi–objective algorithm for the design of prediction intervals for wind power forecasting model", Applied Mathematical Modelling, 67, 101–122, 2019.

[5] Yanhua Chen, Zhaoshuang Heb, Zhihao Shang, Caihong Li, Lian Li, Mingliang Xua, "A novel combined model based on echo state network for multi–step ahead wind speed forecasting: A case study of NREL", Energy Conversion and Management, 179, 13–29, 2019.

[6] DY. Hong, T.Y. Ji, M.S. Li, Q.H. Wu, "Ultra–short–term forecast of wind speed and wind power based on morphological high frequency filter and double similarity search algorithm", Electrical Power and Energy Systems, 104, 868–879, 2019.

[7] Ranran Li, Yu Jin, "A wind speed interval prediction system based on multi–objective optimization for machine learning method", Applied Energy, 228, 2207–2220, 2018.

[8] Kuilin Chen, Jie Yu, "Short–term wind speed prediction using an unscented Kalman filter based state–space support vector regression approach", Applied Energy, 113, 690–705, 2014.

[9] Yan Jiang, Guoqing Huang, Xinyan Peng, Yongle Li, Qingshan Yang, "A novel wind speed prediction method: Hybrid of correlation–aided DWT, LSSVM and GARCH", Journal of Wind Engineering & Industrial Aerodynamics, 174, 28–38, 2018.

[10] Zhenhai Guo, Jing Zhao, Wenyu Zhang, Jianzhou Wang, "A corrected hybrid approach for wind speed prediction in Hexi Corridor of China", Energy, 36, 1668–1679, 2011.

[11] Yuan, X., Tan, Q., Lei, X., Yuan, Y., Wu, X., "Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine", Energy, 129, 122-137, 2017.

[12] A. Khosravi, R.N.N. Koury, L. Machado, J.J.G. Pabon, "Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro–fuzzy inference system", Sustainable Energy Technologies and Assessments, 25, 146–160, 2018.

[13] Xiwei Mi, Hui Liu, Yanfei Li, "Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine", Energy Conversion and Management, 180, 196–205, 2019.

[14] Ping Jiang, Yun Wang, Jianzhou Wang, "Short–term wind speed forecasting using a hybrid model", Energy, 119, 561–577, 2018.

[15] Chuanjin Yu, Yongle Li, Yulong Bao, Haojun Tang, Guanghao Zhai, "A novel framework for wind speed prediction based on recurrent neural networks and support vector machine", Energy Conversion and Management, 178, 137–145, 2018.

[16] Jinqiang Liu, Xiaoru Wang, Yun Lu, "A novel hybrid methodology for short–term wind power forecasting based on adaptive neuro–fuzzy inference system", Renewable Energy, 103, 620–629, 2017.

[17] Cong Feng, Mingjian Cui, Bri–Mathias Hodge, Jie Zhang, "Short–term wind speed forecasting by spectral analysis from long–term observations with missing values", Applied Energy, 190, 1245–1257, 2017.

[18] Yachao Zhang, Kaipei Liu, Liang Qin, Xueli an, "Deterministic and probabilistic interval prediction for short–term wind power generation based on variational mode decomposition and machine learning methods", Energy Conversion and Management, 112, 208–219, 2016.

[19] Chi Zhang, Haikun Wei, Junsheng Zhao, Tianhong Liu, Tingting Zhu, Kanjian Zhang, "Improved week–ahead predictions of wind speed using simple linear models with wavelet decomposition", Renewable Energy, 96, 727–737, 2016.

[20] Zhengtang Liang, Jun Liang, Chengfu Wang, Xiaoming Dong, Xiaofeng Miao, "Short–term wind power combined forecasting based on error forecast correction", Energy Conversion and Management, 119, 215–226, 2016.

[21] Jianming Hu, Jianzhou Wang, Kailiang Ma, "A hybrid technique for short–term wind speed prediction", 81, energy, 563–574, 2015.

[22] Xiaohui Yuan, Chen Chen, Yanbin Yuan, Yuehua Huang, Qingxiong Tan, "Short–term wind power prediction based on LSSVM–GSA model", Energy Conversion and Management, 101, 393–401, 2015.

[23] Yun Wang, Jianzhou Wang, Xiang Wei, "A hybrid wind speed forecasting model based on phase space reconstruction theory and Markov model: A case study of wind farms in northwest China", Energy, 91, 556–572, 2015.

[24] Xiaobing Kong, XiangjieLiu, RuifengShi, Kwang Y. Lee, "Wind speed prediction using reduced support vector machines with feature selection", Neurocomputing, 169, 449–456, 2015.

[25] F.Tagliaferri, I.M.Viola, R.G.J.Flay, "Wind direction forecasting with artificial neural networks and support Vector machines", Ocean Engineering, 97, 65–73, 2015.

[26] Jianzhou Wang, Shanshan Qin, Qingping Zhou, Haiyan Jiang, "Medium–term wind speeds forecasting utilizing hybrid models for three different sites in Xinjiang, China", Renewable Energy, 76, 91–101, 2015.

[27] Da Liu, Dongxiao Niu, Hui Wang, Leilei Fan, "Multi–step forecasting for wind speed using a modified EMD–based artificial neural network model", Renewable Energy, 62, 592–597, 2014.

[28] Yubing Yang, Yun Zhao, "Prevailing Wind Direction Forecasting for Natural Ventilation Adjustment in Greenhouses Based on LE–SVM", Renewable Energy, 16, 252–258, 2012.

[29] Zendehboudi, A.," Implementation of GA-LSSVM modelling approach for estimating the performance of solid desiccant wheels", Energy Conversion and Management, 127, 245-255, 2016.

[30] Chun-Fu Lin, Sheng-De Wang, "Fuzzy support vector machines", IEEE, 13, 464-471, 2002.

[31] C. Mencar, W. Pedrycz" Granular counting of uncertain data", Fuzzy Sets and Systems, 387, 108-126, 2020.

[32] Cortes, C., Vapnik, V., "Support vector network", Machine Learning, 20, 273-297, 1995.

[33] http://sotaventogalicia.com/en/real-time-data/historical.

Figures

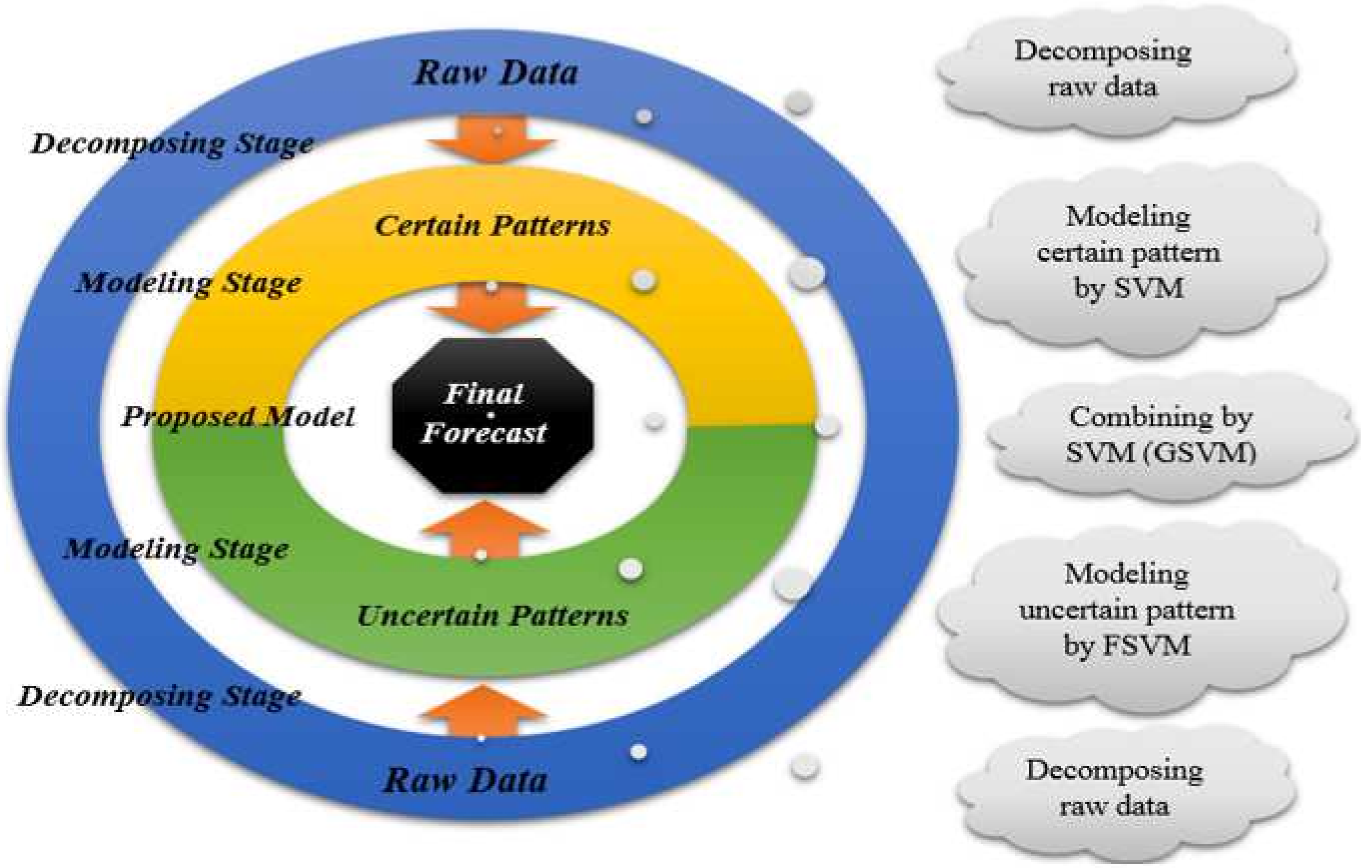


Figure 1

The structure of the proposed model.

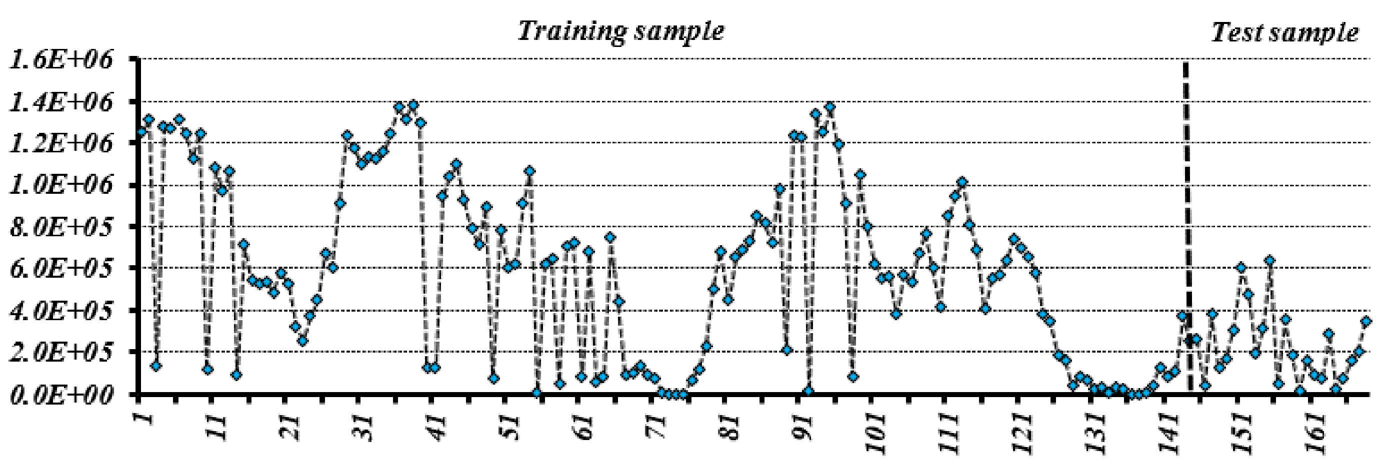


Figure 2

The wind power data sets.

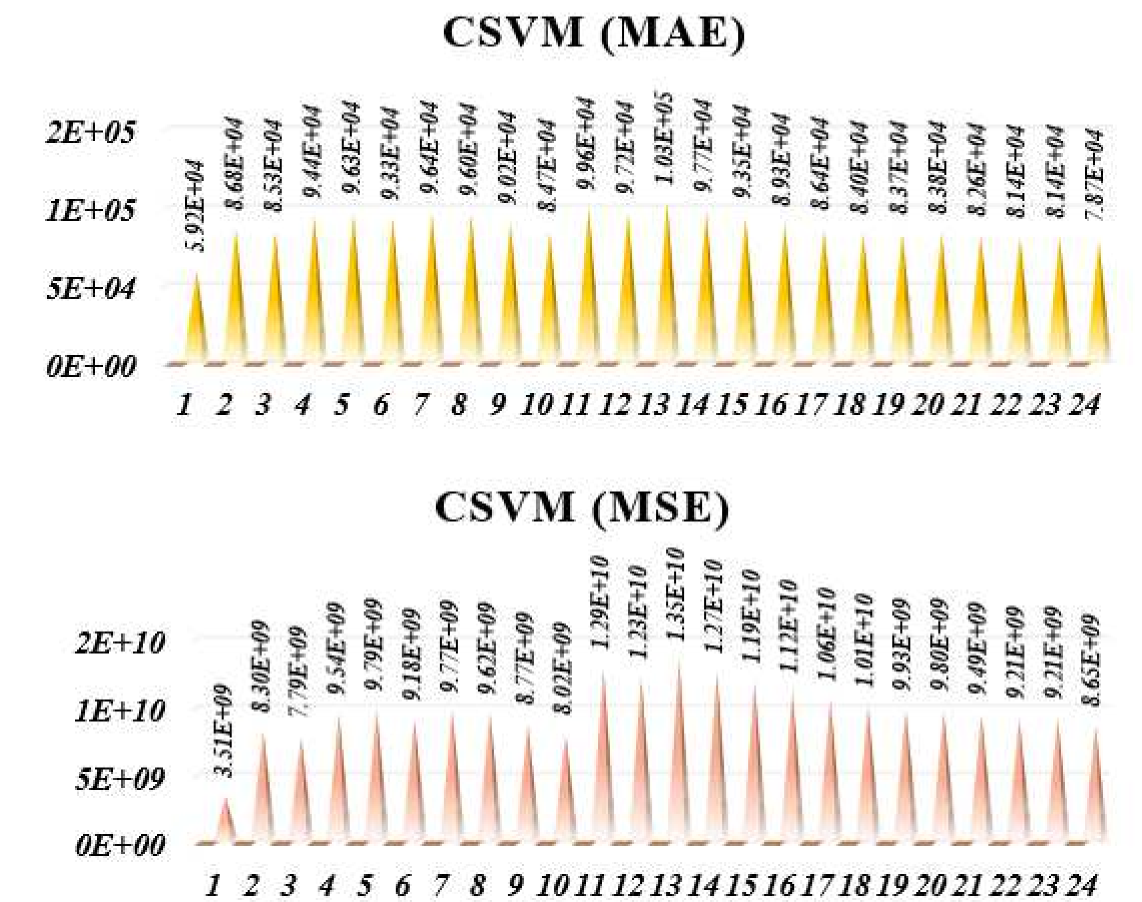


Figure 3

The performance of the CSVM model for test data.

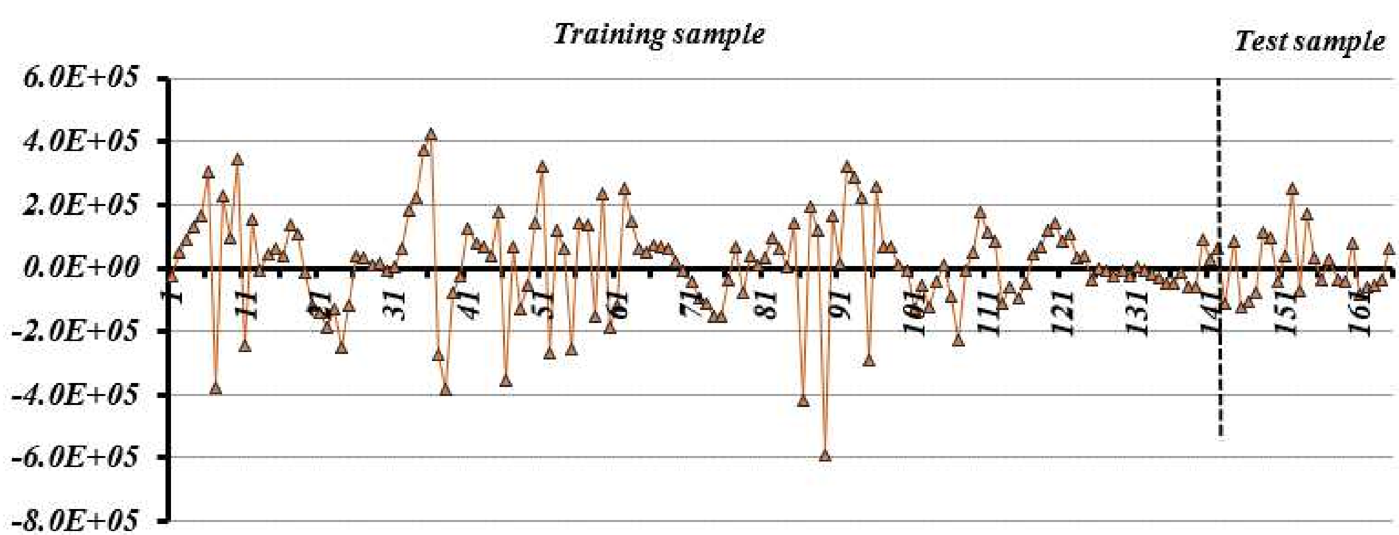


Figure 4

Errors of the CSVM model for training and test data sets.

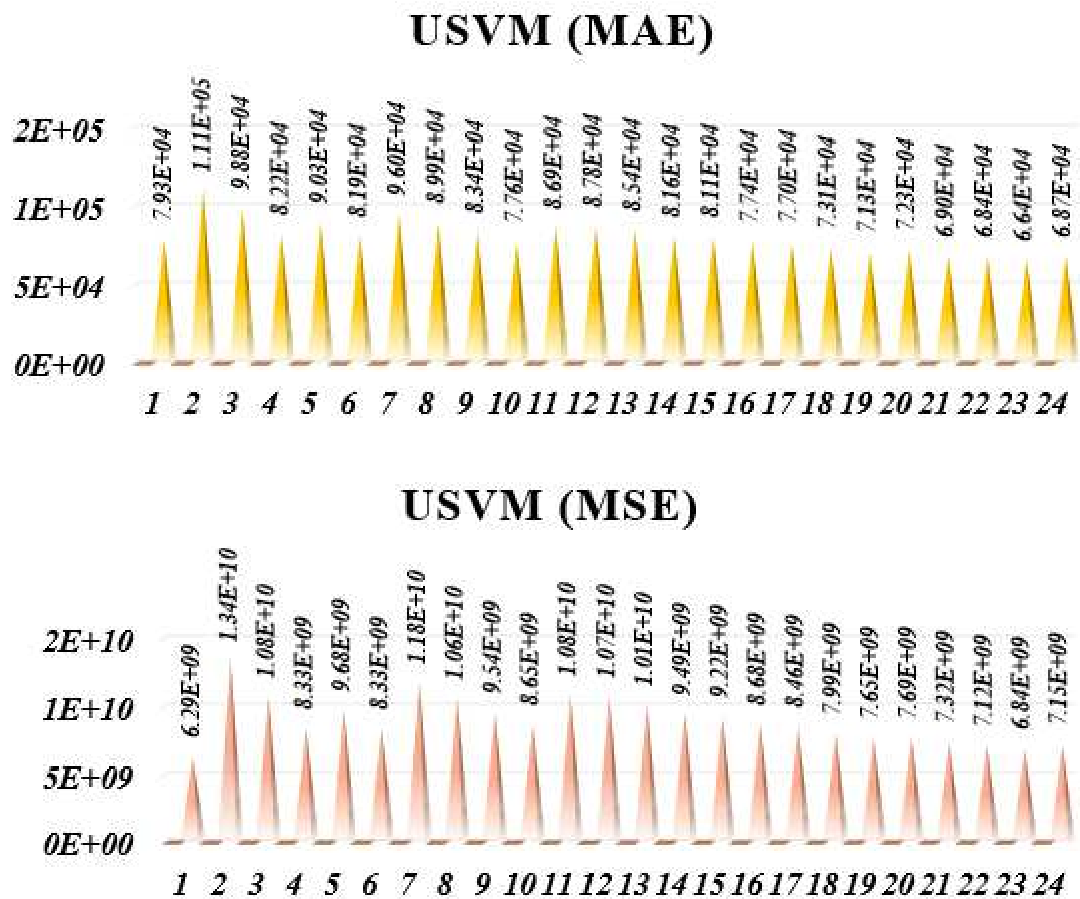


Figure 5

The performance of the USVM model for test data.

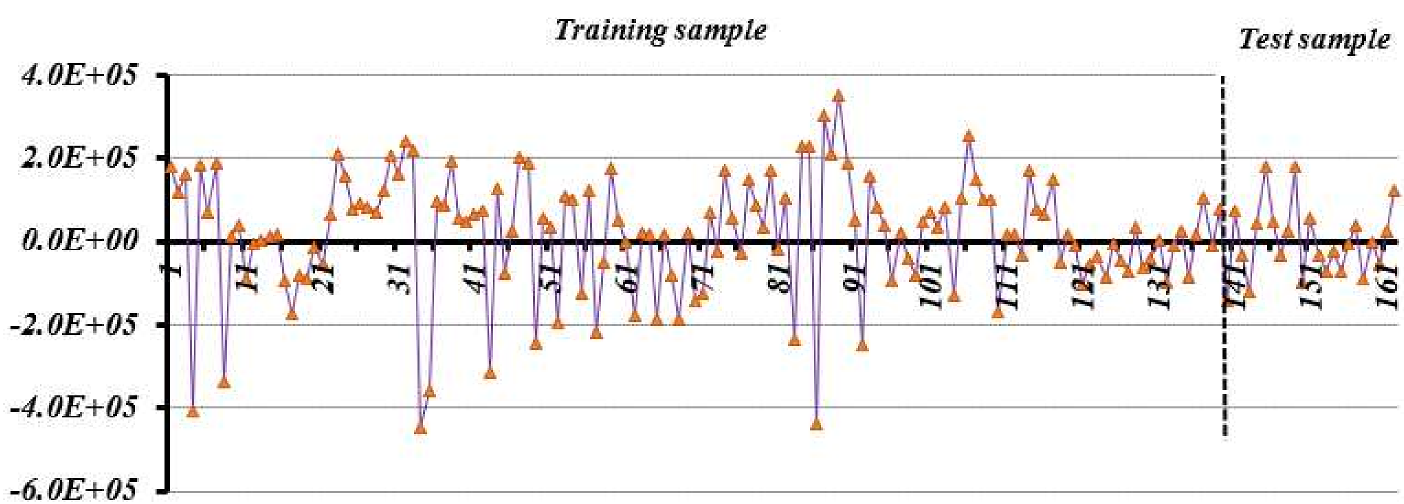


Figure 6

Errors of the USVM model for training and test data sets.

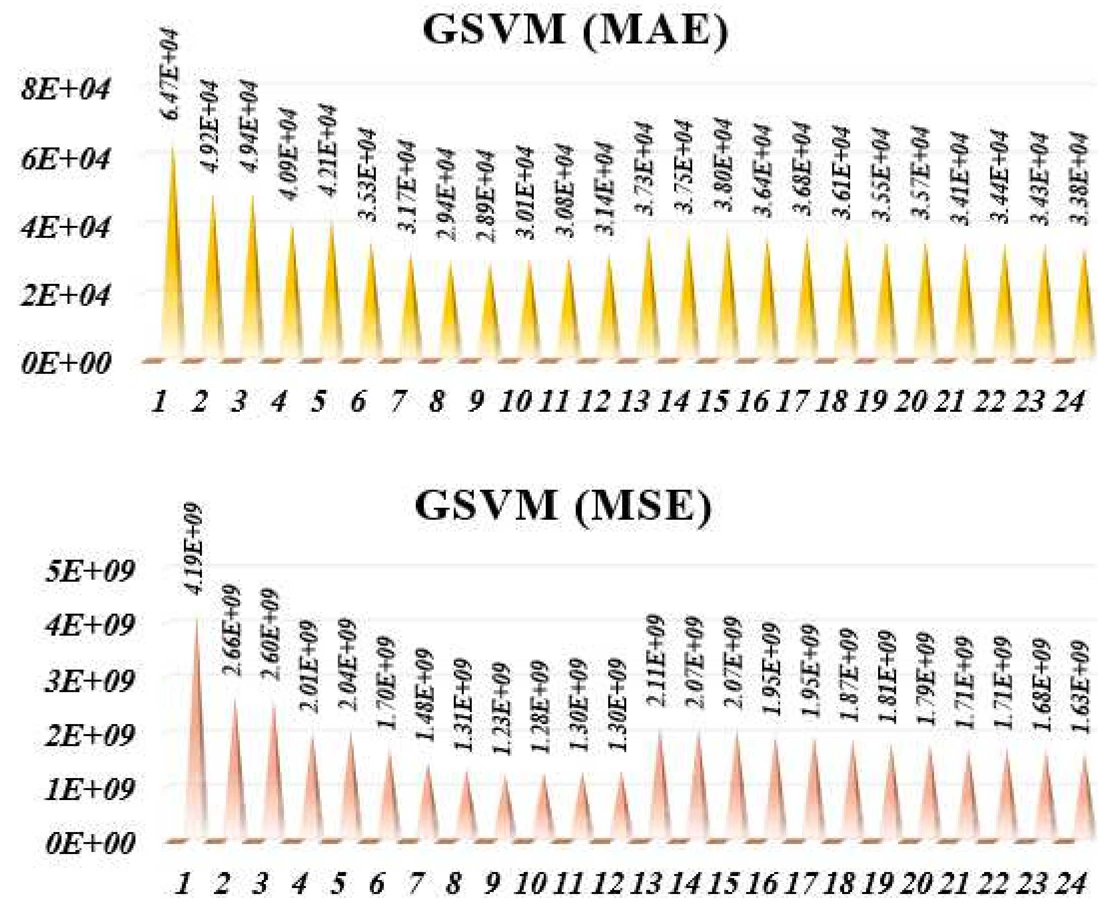


Figure 7

The performance of the proposed model.

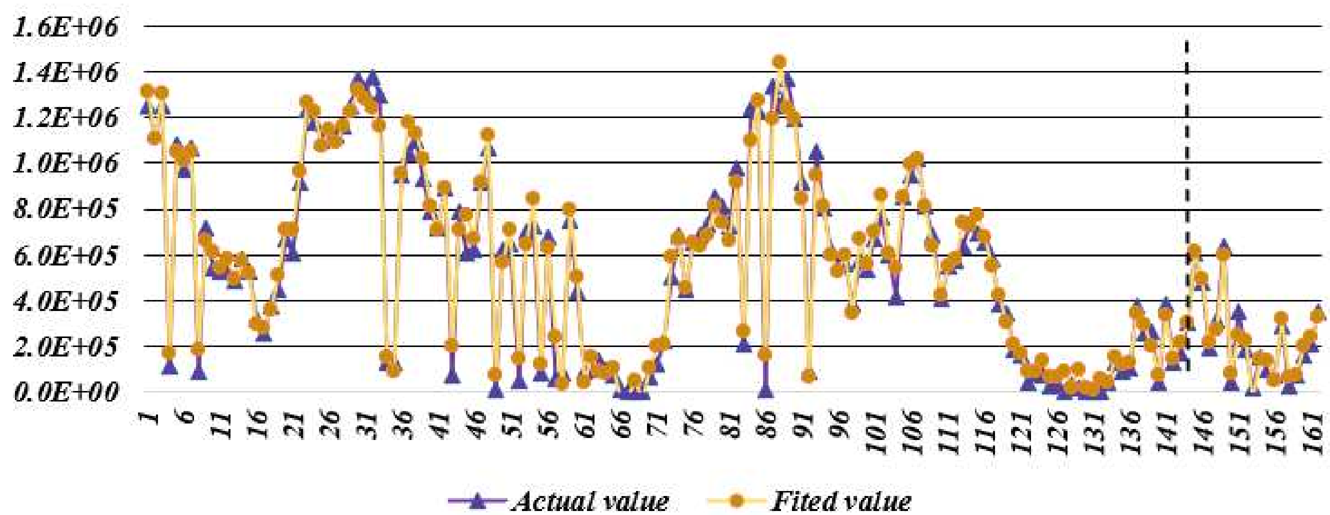


Figure 8

Actual and fitted values of the proposed model for train and test samples.

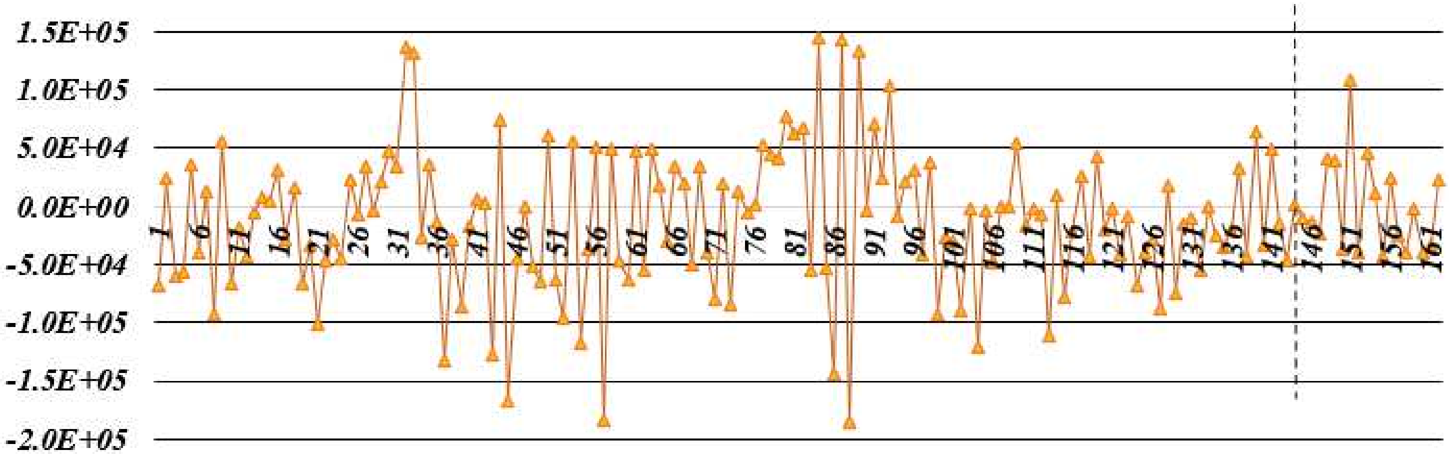


Figure 9

Errors of the proposed model for train and test data sets.

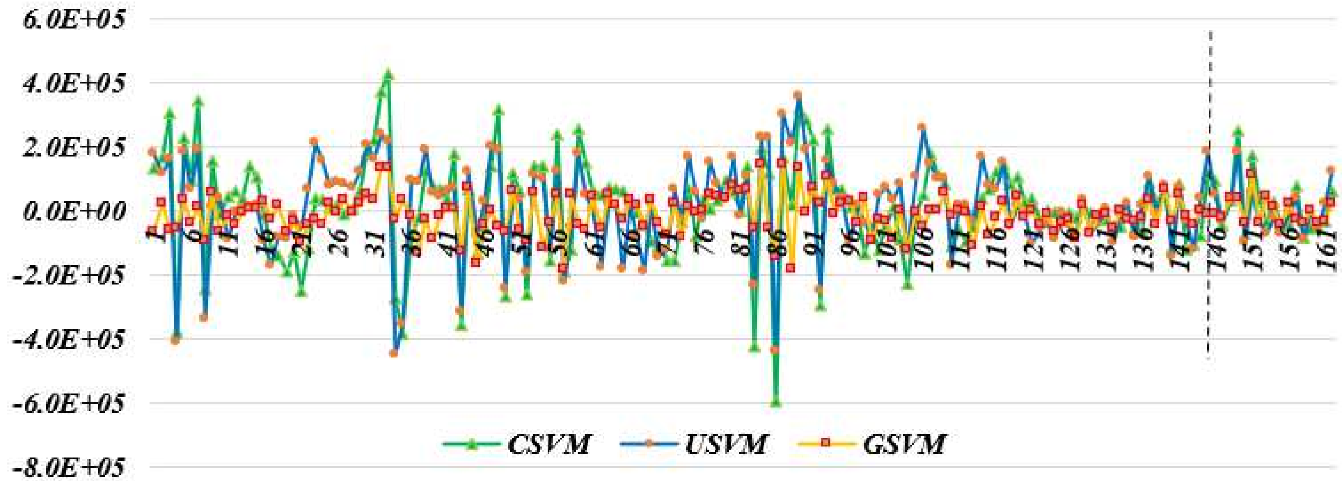


Figure 10

The proposed model errors against its components in the test and train samples.

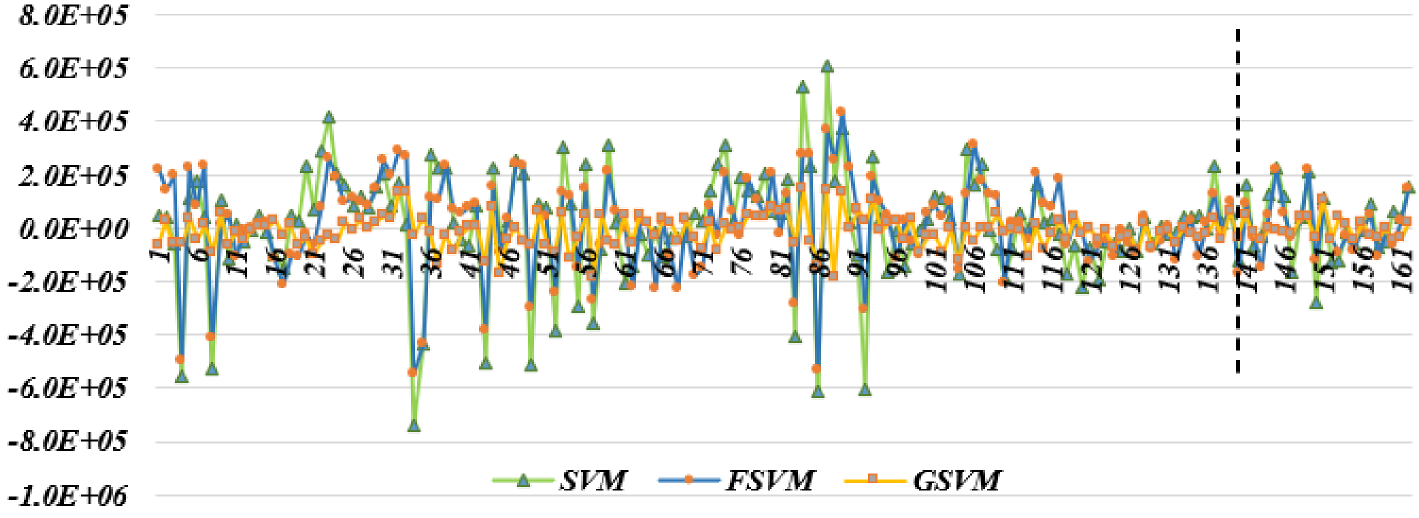


Figure 11

The proposed model errors against base models in the test and train samples.