Short Term Electricity Price Forecasting with Multistage Optimization Technique of LSSVM-GA

Intan Azmira Wan Abdul Razak[1, 2], Izham Zainal Abidin[1], Yap Keem Siah[1], Aidil Azwin Zainul Abidin[1], Titik Khawa Abdul Rahman[3]

[1] College of Engineering, National Energy University, Selangor, Malaysia

[2] Faculty of Electrical Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia,

[3] Faculty of Engineering Girl Campus, King Abdulaziz University, Jeddah, Saudi Arabia.

intan.azmira@utem.edu.my

***Abstract*—Price prediction has now become an important task in the operation of electrical power system. In short term forecast, electricity price can be predicted for an hour-ahead or day-ahead. An hour-ahead prediction offers the market members with the pre-dispatch prices for the next hour. It is useful for an effective bidding strategy where the quantity of bids can be revised or changed prior to the dispatch hour. However, only a few studies have been conducted in the field of hour-ahead forecasting. This is due to most of the power markets apply two-settlement market structure (day-ahead and real time) or standard market design rather than single-settlement system (real time). Therefore, a multistage optimization for hybrid Least Square Support Vector Machine (LSSVM) and Genetic Algorithm (GA) model is developed in this study to provide an accurate price forecast with optimized parameters and input features. So far, no literature has been found on multistage feature and parameter selections using the methods of LSSVM-GA for hour-ahead price prediction. All the models are examined on the Ontario power market; which is reported as among the most volatile market worldwide. A huge number of features are selected by three stages of optimization to avoid from missing any important features. The developed LSSVM-GA shows higher forecast accuracy with lower complexity than the existing models.**

***Index Terms*—Genetic Algorithms; Hour-Ahead Forecasting; Multistage Optimization; Support Vector Machines.**

I. INTRODUCTION

Price prediction is important to market members in deregulated electricity environment to provide a better bidding strategy. As for day ahead forecast, an hour-ahead prediction is useful for an effective bidding strategy where the quantity of bids can be revised or changed prior to the dispatch hour. In addition, the generation company that has the ability to forecast future prices can optimize the output from the generators. The supply and prices can be reviewed and adjusted based on the production cost to gain an optimum profit. Meanwhile, consumers use the price forecast to manage the consumption, especially during spike occurrences.

However, forecasting electricity price is more challenging compared to predicting the load or demand due to the volatility of price series with unexpected price spikes at any point of series. Some of the factors influencing this volatility are manageable such as load behavior, weather, and fuel price. Nevertheless, some other variables are unpredictable such as bidding strategy and imbalance between supply and demand due to (1) demand under-forecast during the peak hour, (2) failure in transaction of import and export, and (3) energy output forecast error by non-dispatchable generators.

II. LITERATURE REVIEW

Only a few researches have been conducted in the field of hour-ahead price forecasting. This is due to most of the power markets apply two-settlement market structure (day-ahead and real time) or standard market design rather than single-settlement system (real time). A time series model of Multivariate Adaptive Regression Splines (MARS) had been developed and examined on the Ontario power market by [1]. Development of neural network models was also reported by other researchers. In [2], Levenberg-marquardt back propagation algorithm was applied to the Ontario power market while Input–Output Hidden Markov Model (IOHMM) was developed by [3] in the Spanish electricity market. Meanwhile, a hybrid method of recurrent neural networks and excitable dynamics was proposed and tested on the Ontario, New South Wales, Spain, and California power markets [4].

A hybrid of ARMAX, adaptive wavelet neural network (AWNN), and GARCH was applied by [5] to treat linear and nonlinear behaviours of price series on Pennsylvania–New Jersey–Maryland (PJM) market. Meanwhile, an Expectation Maximization technique for maximum likelihood estimation of Recurrent Neural Networks (RNN-EM) was developed by [6]. Multi-layer Perceptron Neural Network trained by Extended Kalman Filter (MLP-EKF) and MLP Neural Network trained by Expectation Maximization algorithm (MLP-EM) was proposed by [7]. On the other hand, researchers of [8] designed an Extended Kalman Filter for Recurrent Neural Network (RNN-EKF).

A Generalized Regression Neural Network (GRNN) was developed and tested on National Electricity Market of Singapore (NEMS) [9]. Meanwhile, Discrete Cosine Transforms Input Featured Feed-Forward Neural Network (DCT-FFNN) model was tested on Spanish market by [10]. The same authors further improved the prediction by designing classification models using three layered FFNN, Cascade-Forward Neural Network (CFNN) trained by the Levenberg–Marquardt (LM) algorithm, and GRNN models [11].

Most of the existing methods have the ability to predict well during normal condition or without spike occurrences but when the spikes exist, the forecast error become large. To the best of the authors' review, no literature has been found on the application of LSSVM and GA in the electricity price forecast. In addition, the approach of multistage feature and parameter selections using a single optimization technique has not reported yet. Thus, this study developed a forecasting technique to improve hour-ahead electricity price forecasting using multistage optimization for a hybrid model of Least Square Support Vector Machine (LSSVM) and Genetic Algorithm (GA). With a single optimization method of GA, the input features and LSSVM parameters are simultaneously optimized through three-stage optimization approach. This method is proven to give better forecast accuracy as compared to other existing models which can contribute to decision-making and hourly market operation.

III. FUNDAMENTAL OF SVM AND LSSVM

SVM as presented by [12], is a supervised learning model that supports data analysis and pattern recognition for classification and estimation. Assume that an empirical data is set as Equation (1):

$$\left[(x_1, y_1), \ldots, (x_m, y_m)\right] \in X \times \Re; X = \Re^d \tag{1}$$

where X represents the space of the input patterns. For linear functions f, shown by Equation (2):

$$f(x) = \langle w, x \rangle + b \text{ where } w \in X, b \in \Re \tag{2}$$

Support Vector Regression functions to solve for quadratic programs which involve inequality constraint. However, SVM has a high computational problem where the optimization problem is defined as Equation (3):

$$\begin{aligned} &\min \frac{1}{2}\|\omega\|^2 + C\sum_{k=1}^{N}(\xi_k + \xi_k^*) \\ &\text{where } \xi \text{ is slack variable} \\ &\text{subject to } \begin{Bmatrix} y_k - \langle \omega, \phi(x_k) \rangle - b \le \varepsilon + \xi_k \\ \langle \omega, \varphi(x_k) \rangle + b - y_k \le \varepsilon + \xi_k^* \\ \xi_k, \xi_k^* \ge 0 \end{Bmatrix} \end{aligned} \tag{3}$$

While the ε-insensitive loss function is represented as Equation (4):

$$|y - f(x,\omega)|_\varepsilon = \begin{Bmatrix} 0, \text{if } |y - f(x,\omega)| \le \varepsilon \\ |y - f(x,\omega)| - \varepsilon, \text{otherwise} \end{Bmatrix} \tag{4}$$

Cost of error or regularization constant $C > 0$ specifies the trade-off between margin maximization and training error minimization. Finally, the subsequent SVM for nonlinear function estimation becomes as in Equation (5):

$$f(x) = \sum_{k=1}^{N}\left(\alpha_k - \alpha_k^*\right)\langle\varphi(x), \varphi(x_k)\rangle + b \tag{5}$$

Referring to Mercer's condition, the inner product $\langle\varphi(x), \varphi(x_k)\rangle$ can be represented by a kernel $K(x, x_k)$, and hence, it can be formulated as Equation (6):

$$f(x) = \sum_{k=1}^{N}\left(\alpha_k - \alpha_k^*\right)K(x, x_k) + b \tag{6}$$

SVM can reduce over-fitting, local minima problems [13], and able to deal with high dimensional input spaces splendidly [14]. Nevertheless, the main disadvantage of SVM is its high computational complexity due to constrained optimization programming. Hence, Least Squares Support Vector Machine (LSSVM) was proposed to diminish the computational burden of SVM, which applies with equality instead of inequality constraints [15]. LSSVM solves a system of linear equations to cater for Quadratic Programming (QP) issue that improves the computational speed [14], [16]. The linear system, namely as Karush- Kuhn-Tucker (KKT), is more straightforward than QP system. LSSVM also maintains the principle of SVM, which possess good generalization capability. LSSVM reduces the Sum Square Errors (SSEs) of training data sets and concurrently diminishing margin error. Meanwhile, in contrast to SVM, LSSVM applies the least squares loss function rather than the ε-insensitive loss function and it can be represented as Equation (7):

$$\begin{aligned} &\min J(\omega, e) = \frac{1}{2}\|\omega\|^2 + \frac{1}{2}\gamma\sum_{k=1}^{N}e_k^2 \\ &\text{subject to } y_k = \begin{Bmatrix} \langle \omega, \varphi(x_k) \rangle + b + e_k, \\ k = 1, \ldots, N \le \varepsilon + \xi_k \end{Bmatrix} \end{aligned} \tag{7}$$

The $e_k \in R$ is error variable and $\gamma \ge 0$ is a regularization constant or penalty parameter which controls the trade-off between the fitting error minimization and smoothness of the estimated function. The Lagrangian is introduced as Equation (8):

$$\begin{aligned} L = \frac{1}{2}\|\omega\|^2 + \frac{1}{2}\gamma\sum_{k=1}^{N}e_k^2 - \\ \sum_{k=1}^{N}\alpha_k\left\{\langle\omega, \varphi(x_k)\rangle + b + e_k - y_k\right\} \end{aligned} \tag{8}$$

The conditions for optimality are shown in Equation (9):

$$\begin{cases} \dfrac{\partial L}{\partial \omega} = 0 \to \omega = \sum_{k=1}^{N}\alpha_k\varphi(x_k) \\ \dfrac{\partial L}{\partial b} = 0 \to \omega = \sum_{k=1}^{N}\alpha_k = 0 \\ \dfrac{\partial L}{\partial e_k} = 0 \to \alpha_k = \gamma e_k, (k = 1, \ldots, N) \\ \dfrac{\partial L}{\partial \alpha_k} = 0 \to \omega = \langle\omega, \varphi(x_k)\rangle + b + e_k - y_k = 0 \end{cases} \tag{9}$$

After exclusion of ω and e, the subsequent linear equation is obtained as Equation (10):

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & \Omega + \dfrac{I}{\gamma} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{10}$$

where $y = [y_1;....; y_N]$, $1_v = [1; ...;1]$, $\alpha = [\alpha_1,\ldots \alpha_N]$. The LSSVM model for regression becomes Equation (11):

$$f(x) = \sum_{k=1}^{N} \alpha_k K(x, x_k) + b \tag{11}$$

IV. Fundamental of Genetic Algorithm

GA that was first introduced by [17] is based on the 'survival of the healthiest' and natural evolution mechanism via reproduction. It can find the optimal solution after some iterative computations. The objective functions are often referred to as fitness functions. Three main operations in GA are selection, crossover, and mutation.

The optimization process is started with a random initial population of chromosomes, followed by fitness evaluation. The next step is the selection of fittest individuals or parents for reproduction, where chromosomes with better fitness values have more potential to yield children during subsequent generation. In order to mimic the natural survival of the fittest progression, the best chromosomes exchange genes via crossover and mutation to create children chromosomes during the reproduction process. With the size of the population is preserved, the highly-fit parent perform crossover with the other parent in the population where parts of two genotypes are swapped. The crossover rate usually ranges from 0.6 to 1.0 [18].

After crossover, mutation is performed for any parent chromosome to maintain the variety of the solution candidates by bringing small and random changes into them. Mutations are accomplished randomly by changing a "1" bit into a "0" bit or a "0" bit into a "1" bit. In contrast to crossover, mutation is an unusual process, but by introducing new genetic material to the evolutionary progress, possibly thus avoiding chromosomes from being trapped in local minima. The mutation rate is usually 0.001 [19] or less than 0.1 [18].

The flowchart of GA operation is also illustrated in Figure 1 in Section VII. There are four core elements that influence the performance of GAs; population size, number of generations, crossover rate, and mutation rate. Chances of obtaining global optimum can be increased by having a larger size of population (i.e. hundreds of chromosomes) and generations (thousands), but considerably increasing the computational time [18].

V. The Ontario Power Market

In Ontario, electricity power market is conducted by Independent Electricity System Operator (IESO) which controls power system operation, forecasting short term demand and supply of electricity, and managing the real time spot market electricity price. The Ontario electricity market is a single settlement market, which applies real-time system while the day-ahead system is under progress. The Dispatch Scheduling and Pricing Software (DSPS) is used to provide schedules, prices for energy and operating reserves, and dispatch decisions. Operating reserve is generation capacity where the IESO can call upon on short notice to remain equilibrium between supply and demand during sudden load surge or generator outage.

Due to the single settlement real-time power market, Ontario was reported as one of the most volatile market in the world [20] and hence gives a big challenge for electricity price forecaster. Proper selection of features influences the efficiency and accuracy of forecasting. The important features for electricity price forecasting are analyzed and being selected in the next section.

VI. Selection of Input Features for Forecasting

Correlation analysis is performed to observe the correlations between price and other features as tabulated in Table 1. The analysis uses only the data that is publicly accessible at http://www.ieso.ca/. The notation (h-1) indicates an hour before the forecasting day. The data is selected for January 1- December 31, 2004. Pre-dispatch prices are predicted the price for one, two and three hours ahead. Total market demand (TMD) is the total energy provided by the IESO by combining all output from generators and all scheduled imports to the province. It is also equivalent to the summation of all load supplied from the market, exports from the province and all line losses exist on the IESO-grid.

Meanwhile, Ontario demand is the total energy supplied from the IESO for supplying load within Ontario. The IESO determines Ontario demand by deducting exports from the TMD capacity. It is also equivalent to the summation of all loads within Ontario that is supplied from the market and all line losses occurred on the IESO-grid. The uplift charge is applied to all customers in the wholesale market. This fund is used by the IESO to pay for such items like operating reserve and energy losses on the IESO-grid.

Regarding the correlation coefficient, there is a study shows that high correlation can be considered for correlation coefficient with a range [0.5,1], while medium correlation is referred for correlation coefficient within the range of [0.3, 0.49] [21]. From Table 1, previous HOEP, demand, TMD, and pre-dispatch prices show high correlations with the target HOEP. Hence, HOEP, demand, and 1-hour pre-dispatch price are the selected features for further analysis. However, TMD can be negligible as it has very high correlation with demand data and the TMD contribution can be represented by demand effect. Meanwhile, future HOEP has the highest correlation with past demand than other features.

The correlation of the next hour HOEP with past HOEP and demand were observed for up to past 22 days (528 hours) to demonstrate daily and weekly effects while preventing from missing any important features [9], [22]–[26]. However, only HOEP and demand for past 15 days will be accounted during feature selection to reduce the computational burden. Meanwhile, when the distance of HOEP with the past HOEP and demand becomes further, the correlation between the features and target HOEP becomes lower (HOEP $_{(h\text{-}24)}$ = 0.58, HOEP $_{(h\text{-}168)}$ = 0.48, HOEP $_{(h\text{-}336)}$ = 0.44, HOEP $_{(h\text{-}504)}$ = 0.41, demand $_{(h\text{-}24)}$ = 0.6, demand $_{(h\text{-}168)}$ = 0.6, demand $_{(h\text{-}336)}$ = 0.55, demand $_{(h\text{-}504)}$ = 0.53). Therefore, total features used are [(15 days x 24 hours price) + (15 days x 24 hours demand) + 1-hour pre-dispatch price = 721].

Table 1
Correlation Coefficient of Some Features with Target HOEP

Input	Target	Correlation	
		HOEP	Demand
HOEP / demand $_{(h\text{-}1)}$		0.80	0.67
HOEP / demand $_{(h\text{-}2)}$		0.65	0.59
HOEP / demand $_{(h\text{-}3)}$		0.52	0.49
HOEP / demand $_{(h\text{-}24)}$		0.58	0.60
HOEP / demand $_{(h\text{-}25)}$		0.53	0.56
HOEP / demand $_{(h\text{-}26)}$		0.45	0.48
HOEP / demand $_{(h\text{-}48)}$		0.46	0.50
HOEP / demand $_{(h\text{-}49)}$		0.42	0.47
HOEP / demand $_{(h\text{-}168)}$		0.48	0.60
HOEP / demand $_{(h\text{-}169)}$		0.45	0.56
HOEP / demand $_{(h\text{-}336)}$	HOEP $_{(h)}$	0.44	0.55
HOEP / demand $_{(h\text{-}360)}$		0.36	0.49
HOEP / demand $_{(h\text{-}504)}$		0.41	0.53
HOEP / demand $_{(h\text{-}528)}$		0.34	0.46
1-hour pre-dispatch price		0.71	
2-hour pre-dispatch price		0.70	
3-hour pre-dispatch price		0.68	
Total Market Demand $_{(h\text{-}1)}$		0.65	
Imports $_{(h\text{-}1)}$		0.15	
Exports $_{(h\text{-}1)}$		-0.31	
Uplift Charge $_{(h\text{-}1)}$		0.09	
Total demand $_{(h)}$	Demand $_{(h)}$	0.97	

VII. The Proposed Hybrid Model

A hybrid model of LSSVM-GA is developed with three-stage optimization of feature and parameter. During the first stage, all 721 features are applied and the GA selects a certain number of significant features to be fed into the LSSVM. At the same time, GA optimizes the LSSVM parameters; gamma (γ) and sigma (σ). During the second stage of optimization, GA optimizes the features and parameters that have been selected from the first stage of optimization. These processes are repeated for the next optimization stages until no improvement is observed in the fitness value or Mean Absolute Percentage Error (MAPE). MAPE is formulated as in Equation (12):

$$\mathrm{MAPE} = \frac{100}{N} \times \sum_{t=1}^{N} \frac{\left| P_{actual_t} - P_{forecast_t} \right|}{P_{actual_t}} \tag{12}$$

P_{actual} and P_{forecast} are actual and forecasted HOEP at hour t, respectively, while N is the number of hours in a day. Meanwhile, Mean Absolute Error (MAE) is also calculated as in Equation (13):

$$\mathrm{MAE} = \frac{1}{N} \times \sum_{t=1}^{N} \left| P_{actual_t} - P_{forecast_t} \right| \tag{13}$$

Figure 1 illustrates the flowchart of hybrid LSSVM-GA during each stage of optimization.

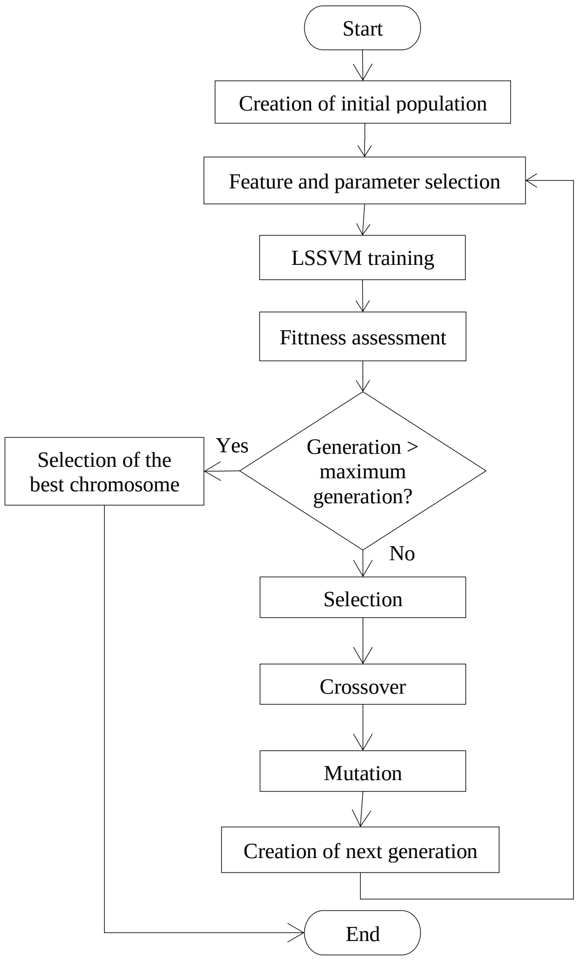

Figure 1: Flowchart of hybrid LSSVM-GA

As a comparison with previous researchers, six forecast models are developed to represent the whole year of 2004. Each model is trained with ten weeks' data prior to the forecasting week as shown in Table 2.

Table 2
Training and Testing Period of 2004

		Training (10 weeks)	Testing (a week)
Spring low point	Week 1	Feb 16 -Apr 25	Apr 26-May 2
	Week 2	Feb 23 -May 2	May 3-9
Summer peak demand	Week 3	May 17 - July 25	July 26 - Aug.1
	Week 4	May 24 - Aug 1	Aug 2 - 8
High demand winter	Week 5	Oct 4 - Dec 12	Dec 13 - 19
	Week 6	Oct 11 - Dec 19	Dec 20 - 26

VIII. RESULT AND DISCUSSION

Table 3 shows the improvement in MAPE when performing the third stage of optimization. It can be noted that the average MAPE is reduced after each stage of optimization except on the fourth stage in which the error increases. Further optimization process may remove some of the important features. The numbers of population and generation are case dependent and usually, the simulation progress is stopped when convergence is reached. Similarly, the features that have been optimized by the GA are case dependent, which is different for each training data set.

Table 3
MAPE for LSSVM-GA model

	W1	W2	W3	W4	W5	W6	Average
Stage 1	8.83	9.70	5.91	9.99	8.85	9.60	8.81
Stage 2	7.67	8.27	5.47	8.06	7.47	8.80	7.62
Stage 3	**7.55**	**7.45**	**5.55**	**7.88**	**7.21**	**8.77**	**7.40**
Stage 4	7.60	7.61	5.77	8.09	7.69	9.07	7.64

Meanwhile, the lowest MAPE is shown during W3 while the highest MAPE occurs during W6. Furthermore, the developed models of LSSVM-GA were compared with other existing models as tabulated in Table 4. There are few existing methods have been implemented for an hour-ahead electricity price forecast in Ontario. Based on the observation from Table 4, the hybrid model of LSSVM-GA outperforms other existing models in terms of accuracy and simplicity. For example, Recurrent Neural Network (RNN) with excitable dynamics model [4] has a more complicated structure which developed to deal with spiky and non-spiky price region. The Fitz-Hugh Nagumo (FHN) system handles the spike portion by the help of RNN model which control the parameters and time scales of FHN. Meanwhile, Feedforward Neural Network (FFNN) is developed to predict the residue errors of RNN-FHN when predicting the stable or non-spiky region. In addition, the output of FFNN is fed to RNN-FHN model to improve the forecasting. Furthermore, Evolutionary Strategies (ES) is incorporated to train the feedforward and feedback weights of the RNN whereas the FFNN is trained by the backpropagation algorithm.

IX. CONCLUSION

In the area of hour-ahead electricity price forecasting, the accuracy of the prediction is the main issue. Market participants use the forecast to review and change the bids prior to the dispatch hour. For accomplishing this goal, selection of input features and parameter is very important during the model development. Until recently, no study has investigated the approach of LSSVM, GA, as well as multistage optimization technique in hour-ahead price prediction.

Hence, a hybrid model of LSSVVM-GA for hour-ahead electricity forecast was developed in this study where only one output is produced at one time. By using the most recent features, GA optimizes the input features and LSSVM parameters simultaneously. This approach may reduce the significant features over the optimization stages while refining the LSSVM parameter values. The developed models of LSSVM-GA outperform other existing models for the same market and test periods. These contributions may help market participants to bid effectively, maintaining efficient hourly operation, and eventually increasing company's profit.

ACKNOWLEDGMENT

We would like to dedicate our appreciation to Ministry of Higher Education Malaysia, Universiti Teknikal Malaysia, Melaka (UTeM) and National Energy University (UNITEN) for providing financial and moral support throughout conducting this study.

Table 4
WMAPE of HOEP Forecast for the Ontario Electricity Market

Ref.	Year	Method	Test week 1	2	3	4	5	6	Average MAPE
		LSSVM+GA	7.55	7.45	5.55	7.88	7.21	8.77	**7.40**
[4]	2013	recurrent NN + excitable dynamics	10.76		9.12		11.61		10.45
[6]	2011	recurrent NN + Expectation Maximization algorithm (RNN-EM)	15.09	15.16	10.52	10.21	15.78	15.71	13.72
		RNN + Extended Kalman Filter (RNN-EKF)	16.01	16.54	11.89	11.96	16.59	16.45	14.91
		MLP+EKF	16.83	16.74	12.64	15.25	16.77	16.96	15.87
		MLP+EM	15.48	15.39	11.87	12.07	16.78	16.73	14.72
[1]	2006	MARS (case 1)	13.3	12.9	9.4	14.4	12.9	15.5	13.07
		MARS (case 2)	12.5	12.3	8.6	11.7	11.8	13.9	11.80
		IESO	23.78	25.26	10.41	16.22	22.06	23.51	20.21

REFERENCES

[1] H. Zareipour, K. Bhattacharya, and C. a. Canizares, "Forecasting the hourly Ontario energy price by multivariate adaptive regression splines," in *2006 IEEE Power Engineering Society General Meeting*, 2006, pp. 1–7.

[2] K. B. Sahay, "One hour ahead price forecast of Ontario electricity market by using ANN," in *2015 International Conference on Energy Economics and Environment (ICEEE)*, 2015, pp. 1–6.

[3] A. Mateo, A. Muñoz, and J. García-González, "Modeling and Forecasting Electricity Prices with Input/Output Hidden Markov Models," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 13–24, 2005.

[4] V. Sharma and D. Srinivasan, "A hybrid intelligent model based on recurrent neural networks and excitable dynamics for price prediction in deregulated electricity market," *Eng. Appl. Artif. Intell.*, vol. 26, no. 5–6, pp. 1562–1574, 2013.

[5] L. Wu and M. Shahidehpour, "A Hybrid Model for Day-Ahead Price Forecasting," *IEEE Trans. Power Syst*, vol. 25, no. 3, pp. 1519–1530, 2010.

[6] D. Mirikitani and N. Nikolaev, "Nonlinear maximum likelihood estimation of electricity spot prices using recurrent neural networks," *Neural Comput. Appl.*, vol. 20, no. 1, pp. 79–89, Feb. 2011.

[7] J. F. G. de Freitas, M. Niranjan, and A. H. Gee, "Dynamic learning with the EM algorithm for neural networks," *J. VLSI Signal Process.*, vol. 26, no. 1/2, pp. 119–131, 2000.

[8] C. M and B. L, "Simple recurrent network trained by RTRL and extended Kalman filter algorithms," *Neural Netw. World*, vol. 13, no. 3, pp. 223–234, 2003.

[9] S. Anbazhagan, "Day-Ahead Price Forecasting in Asia ' S First Liberalized Electricity Market Using Artificial Neural Networks," in *Second International Conference on Sustainable Energy and Intelligent System (SEISCON 2011)*, 2011, vol. 4, no. 4, pp. 476–485.

[10] S. Anbazhagan and N. Kumarappan, "Day-ahead deregulated electricity market price forecasting using neural network input featured by DCT," *Energy Convers. Manag.*, vol. 78, pp. 711–719, 2014.

[11] S. Anbazhagan and N. Kumarappan, "A neural network approach to day-ahead deregulated electricity market prices classification," *Electr. Power Syst. Res.*, vol. 86, pp. 140–150, 2012.

[12] V. N. Vapnik, *Statistical Learning Theory*. new York: Wiley, 1998.

[13] G. Xie, S. Wang, Y. Zhao, and K. K. Lai, "Hybrid approaches based on LSSVR model for container throughput forecasting: A comparative study," *Appl. Soft Comput.*, vol. 13, no. 5, pp. 2232–2241, May 2013.

[14] H. Wang and D. Hu, "Comparison of SVM and LS-SVM for Regression," in *2005 International Conference on Neural Networks and Brain*, 2005, no. 5, pp. 279–283.

[15] J. A. K. Suykens and J. Vandewalle, "Least Squares Support Vector Machine Classifiers," *Neural Process. Lett.*, vol. 9, no. 3, pp. 293–300, 1999.

[16] S. Li and L. Dai, "Classification of gasoline brand and origin by Raman spectroscopy and a novel R-weighted LSSVM algorithm," *Fuel*, vol. 96, pp. 146–152, Jun. 2012.

[17] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. 1975.

[18] E. Elbeltagi, T. Hegazy, and D. Grierson, "Comparison among five evolutionary-based optimization algorithms," *Adv. Eng. Informatics*, vol. 19, no. 1, pp. 43–53, Jan. 2005.

[19] D. Zhijie, Li; Xiangdong, Comparative Research on Particle Swarm Optimization and Genetic AlgorithmLiu; Xiangdon, "Comparative Research on Genetic Algorithm, Particle Swarm Optimization and Hybrid GA-PSO," in *Computer and Information Science*, 2010, vol. 3, pp. 120–127.

[20] H. Zareipour, K. Bhattacharya, and C. a. Cañizares, "Electricity market price volatility: The case of Ontario," *Energy Policy*, vol. 35, no. 9, pp. 4739–4748, Sep. 2007.

[21] B. Dursun, F. Aydin, M. Zontul, and S. Sener, "Modeling and estimating of load demand of electricity generated from hydroelectric power plants in Turkey using machine learning methods," *Adv. Electr. Comput. Eng.*, vol. 14, no. 1, pp. 121–132, 2014.

[22] N. Amjady, a. Daraeepour, and F. Keynia, "Day-ahead electricity price forecasting by modified relief algorithm and hybrid neural network," *IET Gener. Transm. Distrib.*, vol. 4, no. 3, p. 432, 2010.

[23] N. Kumarappan and S. Anbazhagan, "Classification of Day-Ahead Prices in Asia's First Liberalized Electricity Market Using GRNN," in *IET Chennai 3rd International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2012)*, 2012, pp. 207–211.

[24] H. Shayeghi and A. Ghasemi, "Day-ahead electricity prices forecasting by a modified CGSA technique and hybrid WT in LSSVM based scheme," *Energy Convers. Manag.*, vol. 74, pp. 482–491, Oct. 2013.

[25] V. Koban, I. Zlatar, M. Pantos, and M. Omladic, "A remark on forecasting spikes in electricity prices," in *2015 12th International Conference on the European Energy Market (EEM)*, 2015, pp. 1–5.

[26] H. Shayeghi, a. Ghasemi, M. Moradzadeh, and M. Nooshyar, "Simultaneous day-ahead forecasting of electricity price and load in smart grids," *Energy Convers. Manag.*, vol. 95, pp. 371–384, 2015.