

Effect of Demand-Side Management in Electricity Price/Load Forecasting in Smart Grids

H. Shayeghi
Department of Electrical Engineering
University of Mohaghegh Ardabili
Ardabi, Iran

A. Ghasemi
Department of Electrical Engineering
University of Mohaghegh Ardabili
Ardabi, Iran

H. A. Shayanfar*
Department of Electrical Engineering
College of Technical and Engineering
South Tehran Branch
Islamic Azad University
Tehran, Iran

hshayeghi@gmail.com, ghasemi.agm@gmail.com, hashayanfar@gmail.com

Abstract- Electricity price and load forecasting are two important problems for market participants and independent system operators (ISO) in smart grid environments. Most existing papers predict price and load separately, while, the aggregate reaction of consumers can potentially shift the demand curve in the market, resulting in prices that may differ from the initial forecasts. In this regards, demand-side management (DSM) constructs the customers responsible for improving the efficiency, reliability and sustainability of the power system. In this paper, we proposed a new multi-input multi-output (MIMO) system which can consider the interaction between load and price. Therefore, proposed Least Squares Support Vector Machine (LSSVM) to model the nonlinear pattern in price and load. Also, used discrete wavelet transform (DWT) to make valuable subsets. Moreover, proposed feature selection to select best input candidates. Finally, the MIMO-based LSSVM parameters are optimized by artificial bee colony (ABC) algorithm. Simulations carried out NEPOOL region (courtesy ISO New England) electricity market data, and showing that the proposed algorithm has good potential for simultaneous forecasting of electricity price and load in smart grids.

Keywords: Electricity Market, price and load forecast, smart grid, DSM, LSSVM.

I. INTRODUCTION

Growing recognition of the electricity grid modernization to enable novel electricity consumption and generation models has found expression in the infrastructure of the smart grid concept. In fact, the current electricity grid performs rather stable, but, issues such as improving the energy efficiency, the purpose of large-scale renewable energy integration and the reduction in environment emission need a new grid model which is called smart grid [1] as shown in Fig 1. However, it has been already recognized that the Smart Grid is a new electricity network, which highly integrates the advanced sensing and measurement technologies, information and communication technologies (ICTs), analytical and decision-making technologies, automatic control technologies with energy and power technologies and infrastructure of electricity grids [2].

Furthermore, smart grids makes a two-way flow exchange with customers, providing advanced information and options, power export capacity, customers and improved energy efficiency. Note that there are great correlations between these participants in

smart grid. Therefore, all services in smart grid environment have a complexly compared to current grid services. Therefore, there is a strong requirement in the electric power markets for accurate and robust tools that properly forecast electricity price and load signals in simultaneous form, opposite to available methods which only forecast price and load separately [3].

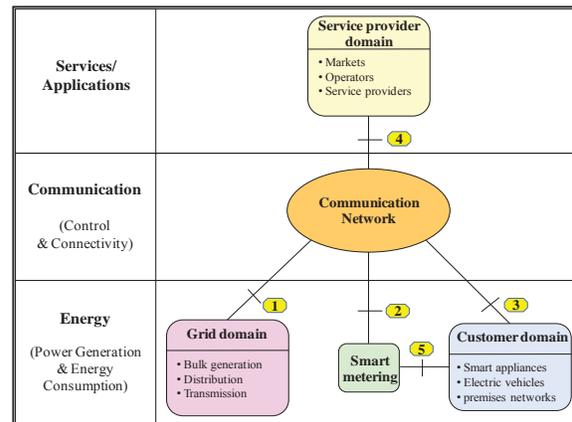


Figure 1. Simplified smart grid domain model

In recent years, some papers have been dedicated to forecasting in electricity markets with this correlation between load and price. In [4] proposed a framework that used MIMO forecasting engine for joint price and demand prediction with data association mining (DAM) algorithms. In [5] a two-stage integrated price and load forecasting framework is developed. At each stage, a hybrid time-series and adaptive wavelet neural network (AWNN) model is used, in which multivariate autoregressive integrated moving average catches the linear relationship of price and load log return series, generalized autoregressive conditional heteroscedastic unveils heteroscedastic character of residuals and AWNN presents non-linear impacts. In [6] proposed a hybrid algorithm based on LSSVM-MIMO+ wavelet packet transform (WPT)+Quasi-Opportunistic Artificial Bee Colony (QOABC) to forecast price and load in smart grid with their correlations. In [7], initial next-day load forecasts are generated using a multi-layer perceptron in the first step. In the second step, it is assumed the next-day electricity prices are known and the variation of prices with respect to the previous day's price is quantified. Fuzzy systems are used to extract

* Corresponding Author. E-Mail: hashayanfar@gmail.com (H. A. Shayanfar)

price-load variation patterns. Finally, the identified patterns are employed to modify and improve initial load forecasts. The load forecasting method in [8] is similar to that of [7] except that RBF neural networks and Adaptive Neural Fuzzy Inference Systems (ANFIS) are employed to generate the load forecasts. A mixed price and demand forecasting method is presented in [9], in which, price and demand are iteratively predicted, and the forecasts are considered in the candidate input set of the subsequent predictors. For forecasting both demand and price, historical price and demand information are considered as input in [9].

In this paper, LSSVM based on MIMO model is proposed which is able to explain linear problems quicker with a more straight forward technique. Moreover, it is necessary to find the optimal LSSVM parameters. In order to achieve this purpose, we used ABC algorithm. Also, to get better the forecast accuracy, the DWT model has been utilized because it is a power tool for noise diminution.

II. PROBLEM EXPRESSION

2.1. DWT

DWT is a powerful tool for noise reduction without destroying the dynamics of the original price series. DWT are discretely translatable and scalable by

$$\psi_{j,k}(x) = a_0^{j/2} \psi(a_0^j x - kb_0) \quad (1)$$

where b_0 and a_0 are referred to translation and scale parameters, respectively, and k and j are two integer numbers. Therefore, the DWT can be defined by

$$DWT(j, k) = \int_{-\infty}^{+\infty} x(t) \psi_{j,k}(x) dx \quad (2)$$

According to these explanations, a DWT consist of two parts; decomposition and reconstruction. In the decomposition phase, the low-pass filter removes the higher frequency components of the signal and high-pass filter picks up the remaining parts. Then, the filtered signals are down-sampled by two and the results are called approximate coefficients and detail coefficients [10]. The reconstruction is just a reversed process of the decomposition and for perfect reconstruction filter banks. Figure 2 depicts the corresponding wavelet decomposition.

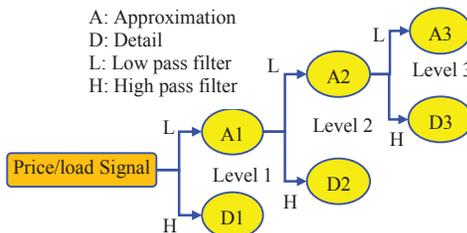


Figure 2. Discrete wavelet decomposition tree

2.2. LSSVM

Given a training set of N data points $\{y_k, x_k\}_{k=1}^N$, where $x_k \in R^n$ is the k -th input pattern and $y_k \in R$ is the k -th output pattern, the classifier can be constructed using the support vector method in the form:

$$y(x) = \text{sign} \left[\sum_{k=1}^N \alpha_k y_k K(x, x_k) + b \right] \quad (3)$$

Where, α_k are called support values and b is a constant. The $K(\cdot, \cdot)$ is the kernel, which is $K(x, x_k) = \exp\{-\|x - x_k\|_2^2 / \sigma^2\}$ (RBF), where κ , θ and σ are constants. For instance, the problem of classifying two classes is defined as:

$$\begin{cases} w^T \phi(x_k) + b \geq +1 & \text{if } y_k = +1 \\ w^T \phi(x_k) + b \leq -1 & \text{if } y_k = -1 \end{cases} \quad (4)$$

This can also be written as:

$$y_k [w^T \phi(x_k) + b] \geq 1, \quad k = 1, \dots, N \quad (5)$$

Where, $\phi(\cdot)$ is a nonlinear function mapping of the input space to a higher dimensional space. LS-SVM classifier:

$$\min_{w, b, e} J_{LS}(w, b, e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^N e_k^2 \quad (6)$$

Subjects to the equality constraints:

$$y_k [w^T \phi(x_k) + b] = 1 - e_k, \quad k = 1, \dots, N \quad (7)$$

The Lagrangian is defined as:

$$L(w, b, e; \alpha) = J_{LS} - \sum_{k=1}^N \alpha_k \{y_k [w^T \phi(x_k) + b] - 1 + e_k\} \quad (8)$$

With Lagrange multipliers $\alpha_k \in R$ (called support values). The conditions for optimality are given by:

$$\begin{cases} \frac{\partial L}{\partial w} = 0 & \rightarrow & w = \sum_{k=1}^N \alpha_k y_k \phi(x_k) \\ \frac{\partial L}{\partial b} = 0 & \rightarrow & \sum_{k=1}^N \alpha_k y_k = 0 \\ \frac{\partial L}{\partial e_k} = 0 & \rightarrow & \alpha_k = \gamma e_k \\ \frac{\partial L}{\partial \alpha_k} = 0 & \rightarrow & y_k [w^T \phi(x_k) + b] - 1 + e_k = 0 \end{cases} \quad (9)$$

For $k = 1, \dots, N$. After elimination of w and e one obtains the solution:

$$\begin{bmatrix} 0 & Y^T \\ Y & ZZ^T + \gamma^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 1_v \end{bmatrix} \quad (10)$$

With:

$$Z = [\phi(x_1)^T y_1; \dots; \phi(x_N)^T y_N], Y = [y_1; \dots; y_N], \quad (11)$$

$$1_v = [1; \dots; 1], e = [e_1; \dots; e_N], \alpha = [\alpha_1; \dots; \alpha_N]$$

Mercer's condition is applied to the matrix $\Omega = ZZ^T$ with:

$$\Omega_{kl} = y_k y_l \phi(x_k)^T \phi(x_l) = y_k y_l K(x_k, x_l) \quad (12)$$

The kernel parameters, i.e. σ for RBF kernel, can be optimally chosen by optimizing an upper bound on the VC dimension [11]. The support values α_k are proportional to the errors at the data points in the LSSVM case, while in the standard SVM case many support values are typically equal to zero.

2.3. Feature selection

The proposed feature selection can be summarized as follows:

Stage I: Mutual information (MI) between $CI(t) \in \{X(t), Y_1(t), Y_2(t), \dots, Y_m(t)\}$ (candidate input) and $x(t)$ (target feature), that is $MI[CI(t), x(t)]$, is calculated based on the binomial distribution method illustrated in [13]. $MI[CI(t), x(t)]$ with high value means $CI(t)$ is a more relevant feature for forecasting $x(t)$. Consequently, the candidate input features of $\{X(t), Y_1(t), Y_2(t), \dots, Y_m(t)\}$ are sorted based on their mutual information with the target feature such that a higher $MI[CI(t), x(t)]$ value results in a higher rank. Afterward, the candidate inputs with $MI[CI(t), x(t)]$ value bigger than TH_1 (relevancy threshold) are retained as the relevant features of the forecast process and the other candidate inputs are filtered out. This stage is irrelevancy filter.

Stage II: Let $S_1 \subset \{X(t), Y_1(t), Y_2(t), \dots, Y_m(t)\}$ denote a subset of candidate inputs chosen in pervious stage. Higher value of MI between two selected candidates $CI_k(t) \in S_1$ and $CI_m(t) \in S_1$, i.e. $MI[CI_k(t), CI_m(t)]$, means more common information between the candidate inputs $CI_k(t)$ and $CI_m(t)$ and so these candidates have a higher level of redundancy. The following redundancy criterion $RC[\cdot]$ calculates the redundancy of each selected feature $CI_k(t) \in S_1$ with the other candidate inputs of S_1 :

$$RC[CI_k(t)] = \max_{CI_m(t) \in S_1} (MI[CI_k(t), CI_m(t)]) \quad (13)$$

We can rank the candidate inputs of S_1 based on the redundancy measure of (13) so that a higher value of $RC[CI_k(t)]$ means $CI_k(t)$ is a more redundant feature or equivalently a less informative candidate input. If $RC[CI_k(t)]$ becomes greater than a redundancy threshold TH_2 , $CI_k(t)$ is considered as a redundant candidate input and so between this candidate and its partner, one feature should be filtered out. For instance, suppose that

$$\arg \max_{CI_m(t) \in S_1} (MI[CI_k(t), CI_m(t)]) = CI_r(t), \quad (14)$$

$$CI_r(t) \in S_1$$

In fact, $CI_r(t)$ has the highest mutual information with $CI_k(t)$ among the features of S_1 . Between $CI_k(t)$ and its partner $CI_r(t)$, one variable should be eliminated. For this purpose, the relevancy factors of these features, i.e. $MI[CI_k(t), x(t)]$ and $MI[CI_r(t), x(t)]$, are considered and the feature with less relevancy factor (less relevant feature or less effective feature for the forecast process) is filtered out. It is possible that more than two features be redundant such that only one of them is eligible to be retained. Therefore, the redundancy filtering process is repeated for all input features of S_1 until no redundancy calculates of (13) become greater than TH_2 .

The redundancy filter, as described above, constitutes the second stage of the proposed two stage feature selection technique. The subsets of feature $S_2 \subset S_1$ that pass the redundancy filter are finally selected candidate inputs by the proposed technique. The candidate features of S_2 are considered as the inputs of the forecast engine. TH_1 and TH_2 are the adjustable parameters of the feature selection technique, which are fine-tuned by the cross-validation method.

2.4. ABC algorithm

In this section, the standard ABC is briefly reviewed [12]. The process of the ABC algorithm is presented as follows:

Step 1. Initialization: generate random population and calculate their fitness values. This population and fitness values called employed bees and nectar amounts, respectively.

Step 2. Move the onlookers: an onlooker bee evaluates the nectar information taken from all employed bees and chooses a food source with a probability related to its nectar amount by "Eq. (15)", this method, known as roulette wheel selection method. The movement of the onlookers follows:

$$p_i = \frac{fit_i}{\sum_{n=1}^{SN} fit_i} \quad (15)$$

Where, p_i and SN are probability of selecting the i^{th} employed bee and number of employed bees, and fit_i is the fitness value of the solution.

$$x_{ij}(t+1) = \theta_{ij}(t) + \phi[\theta_{ij}(t) - \theta_{kj}(t)] \quad (16)$$

Where, $k \in \{1, 2, \dots, BN\}$ and $j \in \{1, 2, \dots, D\}$ are randomly chosen indexes and x_i , t , θ_k and $\phi()$ are the position of the i^{th} onlooker bee, the iteration number, the randomly chosen employed bee and random variable between (-1,1), respectively. D is the number of dimension of optimization problem. BN is number of onlooker bee.

Step 3. Move the scouts: when selected a food source, all the employed bees associated with it abandon the food source, and become scout. The scouts are moved by:

$$\theta_{ij} = \theta_{ij, \min} + r(\theta_{ij, \max} - \theta_{ij, \min}) \quad (17)$$

Where, r denotes a random factor between 0 and 1. $\theta_{ij, \max}$ and $\theta_{ij, \min}$ are lower and upper boundary of x_i , respectively.

Step 4. Update the best food source found so far: Memorize the best food source found so far.

Step 5. Termination checking: checking termination criteria satisfied, if it is satisfied then stop algorithm otherwise go to step 2.

2.5. DSM model

Power system distribution networks are designed for peak loads. For optimum utilization of network capacity, utilities employ DSM with objective of minimum possible peak load. DSM ensures maximum load factor and thus maximizing total profit of utilities. For better insight of different DSM techniques, load distribution over a day is represented by a normal distribution curve as in Fig 3.

A. Valley Filling: The loads during off peak hours, i.e., region III, are increased to achieve flatter profile.

$$[\uparrow p_3] \Leftrightarrow [load\ addition] \quad (18)$$

B. Load Shifting: The shift able loads during peak hours are shifted to off-peak hours, resulting lower peak of curve and a flatter profile i.e. region 1 are shifted to region 2 and 3.

$$[\uparrow p_3 \cap \uparrow p_2] \Leftrightarrow [\downarrow p_1] \quad (19)$$

C. Peak Clipping: The load from peak hours (region 1) is reduced like scheduled power cuts.

$$[\downarrow p_1] \Leftrightarrow [\text{load removal}] \quad (20)$$

D. Energy Conservation: This applied when reduction in load is required all over the load curve.

$$[\downarrow p_1 \cap \downarrow p_3 \cap \downarrow p_2] \Leftrightarrow [\text{load removal}] \quad (21)$$

E. Load Building: It employed when increased energy consumption is required due to surplus production.

$$[\uparrow p_1 \cap \uparrow p_3 \cap \uparrow p_2] \Leftrightarrow [\text{load addition}] \quad (22)$$

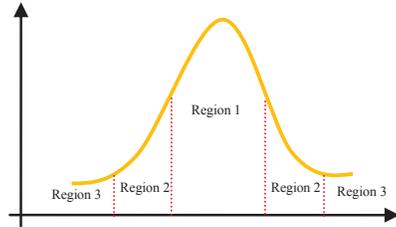


Figure 3. Representation of load profile; P_1 : Probability of load lying in region 1, P_2 : Probability of load lying in region 2, P_3 : Probability of load lying in region 3.

III. EFFECT OF DSM IN FORECASTING

Currently let's see what happens if a component of DSM has the ability to response to the price variations. This state has been shown in Fig 4. In this case, the demand curve has two parts [4]. The first one is that which cannot response to the electricity price variations (price taking part). This part is shown as a vertical line. The second part is related to the price responsive part of demand and has a minus slope in Fig 4. In fact, this part represents the DSM of the system. As can be seen in Fig. 4, proper DSM in the system can protect the demand-side from price spikes. Therefore the price taking part of the demand will also benefits from the DSM of the system. As mentioned before, by emerging the smart grids the response of consumers to the price signals of the electricity market will not be weak and the spot electricity price and demand will mutually affect each other.

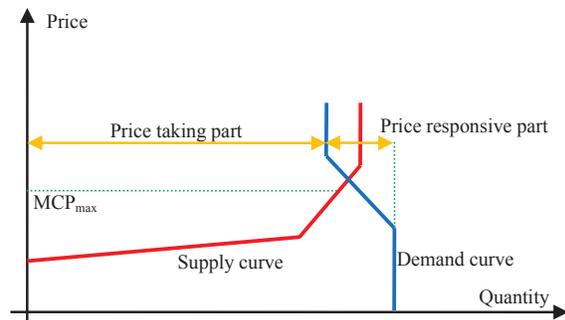


Figure 4. The effect of DSM on price spikes in the electricity markets.

Therefore in this new electricity grid, load forecasting and price forecasting cannot be implemented independently. In fact, the penetration level of smart grid technologies like advanced metering infrastructure (AMI) systems, smart monitors and smart controllers and

also the arrangement of different kinds of consumers (i.e. the ratio of industrial, commercial, and residential consumers) will affect the overall response of the demand-side to the price signals. Although, in long-term these factors will become fixed and the overall response of the demand-side will be stable, but they have great effect on DR in the transition period. The structure of the forecaster is also likely to be changed due to the dynamic relation between load and price. Fig 5, schematically shows possible inputs and outputs of the future load and price forecasting systems.

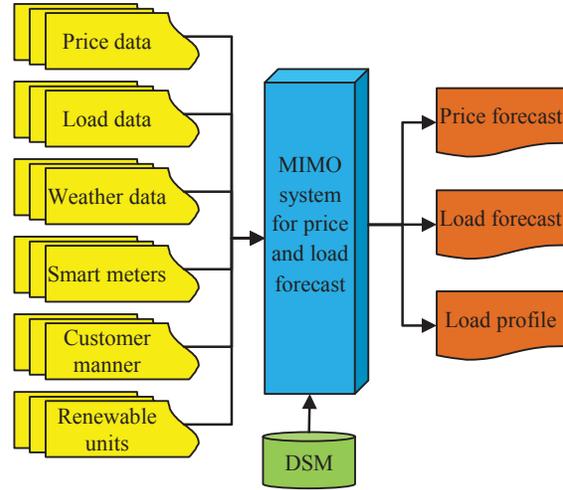


Figure 5. Structure of price and load forecast in smart grid with DSM.

IV. PROPOSED PRICE AND LOAD FORECASTING BASED MIMO CHANNEL

In rest of this section, the proposed forecast is describe,

Stage 1: Firstly, the historical data for price and load are sorted as input data. Now, with DWT system, these signals individually are decomposed in detail (price; D_p and load; D_L) and approximate (price; A_p and load; A_L) subsets. Then, proposed feature selection is used to select best data with more relevancies and least redundancy.

Stage 2: When the Multi-Input Multi-Output (MIMO) is launched with $\{x_{p1}, x_{p2}, \dots, x_{pm}\}$ and $\{x_{l1}, x_{l2}, \dots, x_{lm}\}$ vectors then the training process will be starting. This performance about learning is shown in Fig 6 with more detail.

Stage 3: This stage used inverse DWT transform to guess the hourly prices for d^{th} day by means of the estimates d^{th} day of the constitutive subsets. In other words, this model is used in turn in order to reconstruct the estimate signal for price and load, i.e.,

$$W^{-1}(\{a_h^{est}, b_h^{est}, c_h^{est}, d_h^{est}; h = T + 1, \dots, T + 24\}) = P_h^{W, est} \quad (23)$$

The objective function is calculated by the following equation,

$$Obj = \frac{1}{N} \left(\sum_{i=1}^N \left(\frac{P_{act_i} - P_{for_i}}{P_{act_i}} + \frac{L_{act_i} - L_{for_i}}{L_{act_i}} \right) \right) \quad (24)$$

Where, P_{act} and P_{for} are the forecast and actual values of price signal, L_{act} and L_{for} are the forecast and actual values of load signal, respectively.

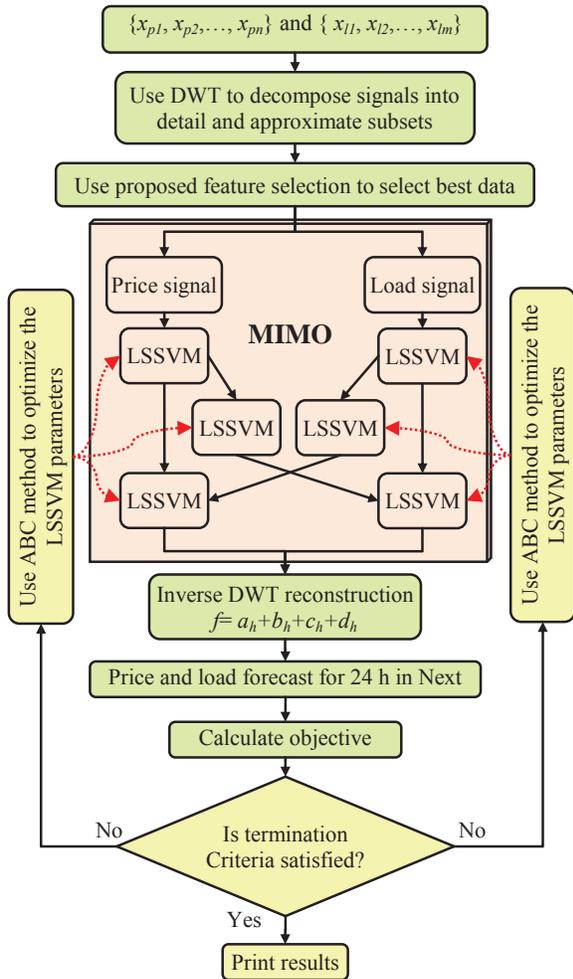


Figure 6. Schematic diagram of the proposed hybrid forecasting methodology

Stage 4: The aim of this stage is error minimizing, in fact, we done this process by adjusting the LSSVM parameters using forecasted and actual values. If the maximum number of training process ($iter_{Max}$) is reached, then finish. If not, go to *stage 2*.

V. SIMULATION RESULTS AND DISCUSSION

The model are trained on hourly data from the New England Power Pool (NEPOOL) region (courtesy ISO New England) from 2004 to 2007 and tested on out-of-sample data from 2008 [14]. The most widely used criterion to measure forecasting error is the Mean Absolute Percentage Error (MAPE), the Forecast Mean Square Error (FMSE), and the Error Standard Deviation (ESD). This accuracy is calculated as a function of the actual prices that occur. The daily MAPE can be given by:

$$MAPE_{day/week} = \frac{1}{N} \sum_{i=1}^N \frac{|P_{iACT} - P_{iFOR}|}{P_{AVE-ACT}} \quad (25)$$

$$P_{AVE-ACT} = \frac{1}{N} \sum_{i=1}^N P_{iACT} \quad (26)$$

Where, P_{iFOR} and P_{iACT} are referred to forecasted and actual values, and $P_{AVE-ACT}$ is the average value of P_{iACT} as given in Eq. (26) to avoid the adverse effects of prices close to zero. Furthermore, the FMSE is the square root of the average of 24 (daily) squared differences between the actual prices and the forecasted ones;

$$FMSE_{day/week} = \sqrt{\frac{1}{N} \sum_{i=1}^N (P_{iACT} - P_{iFOR})^2} \quad (27)$$

The ESD index, one of the important performance criteria, is given by:

$$ESD_{day/week} = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_i - E_i^{Ave})^2}, \quad (28)$$

$$E_i = P_{iACT} - P_{iFOR}, E_i^{Ave} = \frac{1}{N} \sum_{i=1}^N E_i$$

The hourly electricity price and load signals of the NYISO electricity market are shown in Figs. 7 and 8, respectively. Also, to the reader's convenience, Fig. 9 shows the relation between load and price signals. Figure 10 shows the weekly forecast load signal, and, similarly, Fig 11 is referred to price signal in this period.

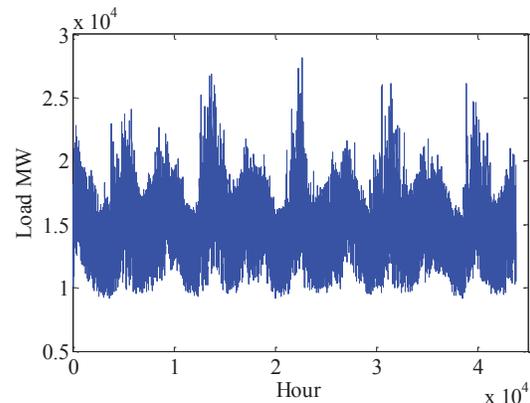


Figure 7. Load signal in the NEPOOL electricity market.

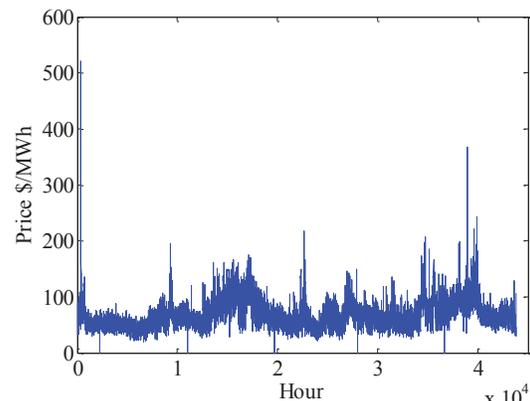


Figure 8. Price signal in the NEPOOL electricity market.

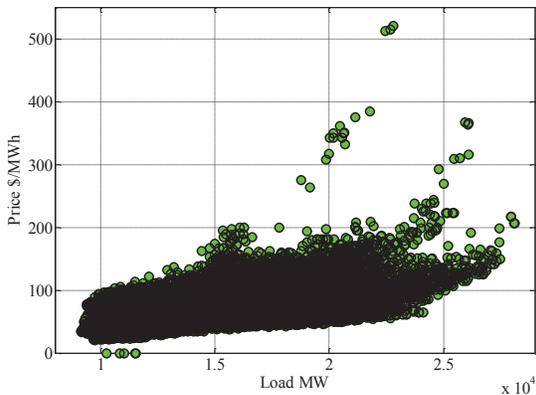


Figure 9. Relation between load price signals in the NEPOOL electricity market.

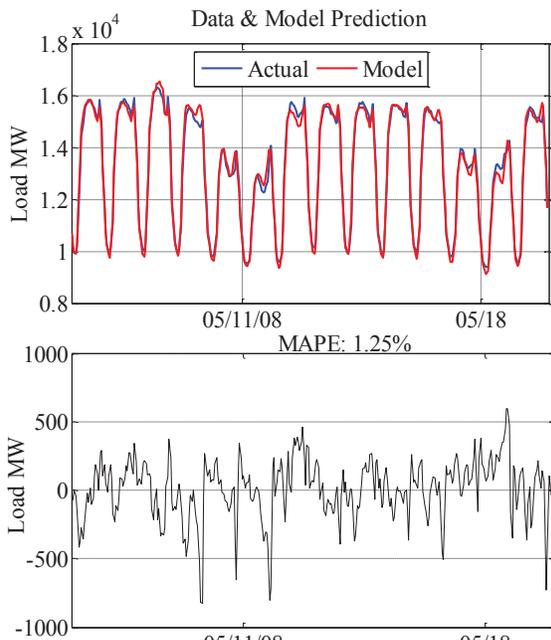


Figure 10. Weekly forecasted electricity load result for the NEPOOL electricity market.

The minute differences between the target value and predicted value may be due to measurement errors or special occasion day or severe weather conditions in present year. Variations due to seasonal changes are taken care by the inputs x_p and x_l viz. load and price of previous day and average load and price of previous week, as seasonal changes are gradual. Load and price on same day of previous week caters for weekend loads and prices. Effect of previous year's variations is avoided by using normalized data; however one limitation of model is that it would not be able to do prediction for holiday, rainy day or severe weather conditions. The network performance curve, between MSE and epochs is shown in Fig 12. The best performance achieved for validation is at epoch 67 with MSE of 3.28×10^{-3} . The regression plot between network response and target is shown in Fig. 13. The R (correlation coefficient value) values for training,

validation and testing are 0.9943, 0.9903 and 0.9939 respectively. The overall R value is 0.9942 resulting in very close prediction.

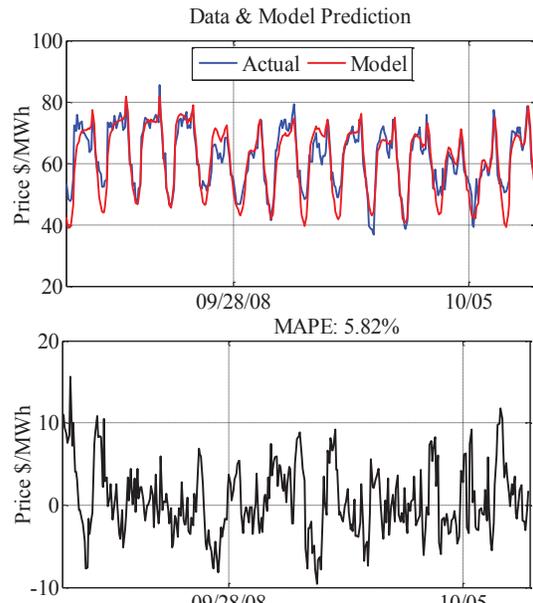


Figure 11. Weekly forecasted electricity price result for the NEPOOL electricity market.

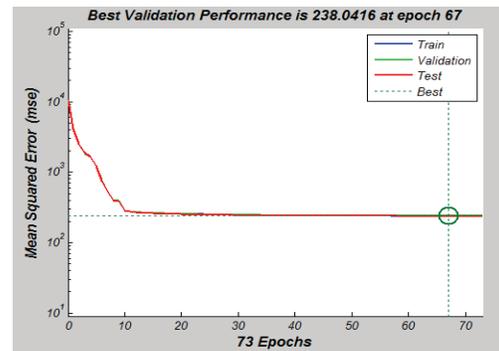


Figure 12. Performance curve for LSSVM training.

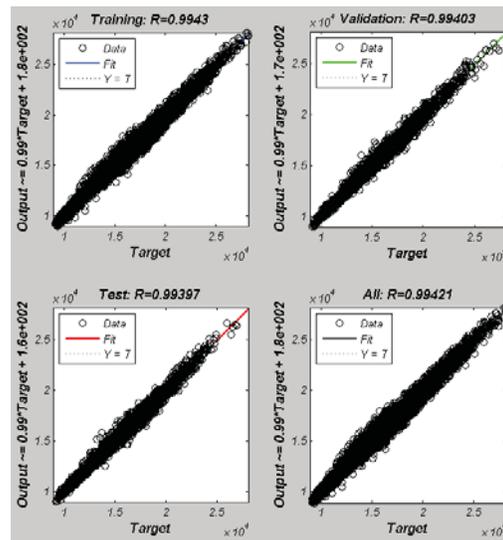


Figure 13. Regression plot.

VI. CONCLUSIONS

A hybrid short-term load and price forecast with a new feature selection framework is proposed in this paper to enhance the forecast accuracy in smart grid. The correlation between load and price signals is modeled with a LSSVM learning algorithm. In other words, a hybrid forecasting framework is proposed which takes into account the bidirectional price-demand relationships when forecasting electricity market price and demand. Hybrid models combine the capabilities of different modeling approaches and are reported to improve forecasting performance in previous applications. The New England hourly price and load data are used to evaluate the performance of the feature selection algorithm. Simulation results demonstrate the efficiency of the selected input sub-series to improve the forecast accuracy. The results also show that the forecast errors for the special days are significantly decreased by including all the selected sub-series in the input set. The selected sub-series are used as the inputs to the individual LSSVMs and the outputs are combined using the determined weighting coefficients to provide the load and price forecast.

REFERENCES

- [1] P. Siano, "Demand response and smart grids-A survey," *Renewable and Sustainable Energy Reviews*, 2014;30:461-478.
- [2] Z. Fan, P. Kulkarni, S. Gormus, C. Efthymiou, G. Kalogridis, M. Sooriyabandara, et al, "Smart grid communications: overview of research challenges, solutions, and standardization activities," *IEEE Commun Surv Tutor*, 2013;15 (1):21-38.
- [3] J. Aghaei, M.I. Alizadeh, "Demand response in smart electricity grids equipped with renewable energy sources: a review," *Renewable Sustainable Energy Rev*, 2013;18:64-72.
- [4] M. Amir, Z. Hamidreza, D.R. William, "Electricity Price and Demand Forecasting in Smart Grids," *IEEE Trans. smart grid*, 2012;3(2):664-674.
- [5] L. Wu, M. Shahidehpour, "Hybrid model for integrated day-ahead electricity price and load forecasting in smart grid," *IET Generation, Transmission & Distribution*, 2014;8(12):1937-1950.
- [6] H. Shayeghi, A. Ghasemi, M. Moradzadeh, M. Nooshyar, "Simultaneous day-ahead forecasting of electricity price and load in smart grids," *Energy Conversion and Management*, 2015;95:371-384.
- [7] A. Khotanzad, E. Zhou, and H. Elragal, "A neuro-fuzzy approach to short-term load forecasting in a price-sensitive environment," *IEEE Trans. Power Syst.*, 2002;17(4):1273-1282.
- [8] Z. Yun, Z. Quan, S. Caixin, L. Shaolan, L. Yuming, and S. Yang, "RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment," *IEEE Trans. Power Syst.*, 2008;23(3):853-858.
- [9] N. Amjady, A. Daraeepour, "Mixed price and load forecasting of electricity markets by a new iterative prediction method," *Electr. Power Syst. Res.*, 2009; 79(9):1329-1336.
- [10] H. Shayeghi, A. Ghasemi, "Day-ahead electricity prices forecasting by a modified CGSA technique and hybrid WT in LSSVM based scheme," *Energy Conversion and Management*, 2013;74:482-491.
- [11] V.N. Vapnik, "The nature of statistical learning theory," New York: Springer; 1995.
- [12] H. Shayeghi, A. Ghasemi, "A modified artificial bee colony based on chaos theory for solving non-convex emission/economic dispatch," *Energy Conversion and Management*, 2014;79:344-54.
- [13] N. Amjady, F. Keynia, Day-ahead Price Forecasting of Electricity Markets by Mutual Information Technique and Cascaded Neuro-Evolutionary Algorithm, *IEEE Trans. Power Syst.*, 2009;24(1):306-318.
- [14] <http://www.iso-ne.com/>.



Hossein Shayeghi received the B.S. and M.S.E. degrees in Electrical and Control Engineering in 1996 and 1998, respectively. He received his Ph.D. degree in Electrical Engineering from Iran University of Science and Technology, Tehran, Iran in 2006. Currently, he is a Professor in Technical Engineering Department of University of Mohaghegh Ardabili, Ardabil, Iran. His research interests are in the application of robust control, artificial intelligence and heuristic optimization methods to power system control design, operation and planning and power system restructuring. He has authored and co-authored of five books in Electrical Engineering area all in Farsi, two book chapters in international publishers and more than 180 papers in international journals and conference proceedings. Also, he collaborates with several international journals as reviewer boards and works as editorial committee of three international journals. He has served on several other committees and panels in governmental, industrial, and technical conferences. He was selected as distinguished researcher of the University of Mohaghegh Ardabili several times. In 2007 and 2010 he was also elected as distinguished researcher in engineering field in Ardabil province of Iran. Also, he is a member of Iranian Association of Electrical and Electronic Engineers (IAEEE) and IEEE. Currently, he is head of Ardabil Technology Incubation Center (ATIC) at University of Mohaghegh Ardabili since 2008.



Ali Ghasemi received the B.Sc. and M.Sc. (Honors with first class) degree in electrical engineering from Isfahan University of Technology (IUT), Isfahan, and University of Mohaghegh Ardabili (UMA), Ardabil, Iran, in 2009 and 2011, respectively. Currently he is pursuing the Ph.D. degree in the electrical engineering and computer science of UMA. His research interests are application of forecast methods, operation adaptive and robust control of power systems, Planning, Power System Restructuring and applications of heuristic techniques. He has authored of a book in Electrical Engineering area in Farsi, and more than 120 papers in reputable international journals and conference proceedings. Also, he collaborates as editorial committee and reviewer of 20 international journals. He is a member of the Iranian Association of Electrical and Electronic Engineers (IAEEE). In 2012 and 2013, He received the award of the 4th *Electric Power Generation Conference* and *UMA* for his M.Sc. thesis. He is the recipient honor M.Sc and Ph.D student Award of *UMA*, 2014. He received the 2013 best young researcher award of the *Young Researcher and Elite Club*.



Heidarali Shayanfar received the B.S. and M.S.E. degrees in Electrical Engineering in 1973 and 1979, respectively. He received his Ph. D. degree in Electrical Engineering from Michigan State University, U.S.A., in 1981. Currently, he is a Full Professor in Electrical Engineering Department of Iran University of Science and Technology, Tehran, Iran. His research interests are in the Application of Artificial Intelligence to Power System Control Design, Dynamic Load Modeling, Power System Observability Studies, Voltage Collapse, Congestion Management in a Restructured Power System, Reliability Improvement in Distribution Systems, Smart Grids and Reactive Pricing in Deregulated Power Systems. He has published more than 497 technical papers in the International Journals and Conferences proceedings. He is a member of Iranian Association of Electrical and Electronic Engineers and IEEE.