

Month ahead average daily electricity price profile forecasting based on a hybrid nonlinear regression and SVM model: an ERCOT case study



Ziming MA¹, Haiwang ZHONG¹, Le XIE², Qing XIA¹,
Chongqing KANG¹

Abstract With the deregulation of the electric power industry, electricity price forecasting plays an increasingly important role in electricity markets, especially for retailers and investment decision making. Month ahead average daily electricity price profile forecasting is proposed for the first time in this paper. A hybrid nonlinear regression and support vector machine (SVM) model is proposed. Off-peak hours, peak hours in peak months and peak hours in off-peak months are distinguished and different methods are designed to improve the forecast accuracy. A nonlinear regression model with deviation compensation is proposed to forecast the prices of off-peak hours and peak hours in off-peak months. SVM is adopted to forecast the prices of peak hours in peak months. Case studies based on data from ERCOT validate the effectiveness of the proposed hybrid method.

Keywords Electricity price forecasting, Month ahead average daily electricity price profile, Nonlinear regression model, Support vector machine (SVM), Electric reliability council of Texas (ERCOT)

1 Introduction

1.1 Motivation

During the past two decades, deregulation of the electric power industry has occurred in many regions. Retailors and large customers purchase electricity from wholesale markets and most of the small-and-medium-sized customers buy their electricity from retail markets. Electricity price forecasting plays an increasingly important role in generating, retailing, and planning.

Electricity price forecasting can be categorized into short-term and mid-term forecasting. Short-term price forecasting can help market participants determine the day-ahead bid/offer for the wholesale market to maximize the benefit and manage the risks [1]. Mid-term monthly average electricity price forecasting (e.g., month ahead) is essential in guiding the participants over mid- and long-term time scales. In retail markets, the month-to-month variable rate plan is prevalent. The retail rate varies from month to month based on the market price and retailers usually issue the retail rate for next month one month ahead of time. In this sense, mid-term price forecasting will be helpful for retailers. In addition, mid-long-term price forecasting can provide price signals for generation expansion.

However, the disadvantages of monthly average electricity price forecasting are obvious.

CrossCheck date: 25 January 2018

Received: 30 June 2017 / Accepted: 25 January 2018 / Published online: 26 February 2018

© The Author(s) 2018. This article is an open access publication

✉ Qing XIA
qingxia@tsinghua.edu.cn

Ziming MA
mzm_thu@foxmail.com

Haiwang ZHONG
zhonghw@mail.tsinghua.edu.cn

Le XIE
le.xie@tamu.edu

Chongqing KANG
cqkang@tsinghua.edu.cn

¹ Department of Electrical Engineering, Tsinghua University, Beijing 100084, China

² Department of Electrical/Computer Engineering, Texas A&M University, College Station, TX 77843, USA



- 1) For retailers, although monthly average electricity price forecasting can help them set the rate plan, the contribution of monthly average price to the retailers is very limited. The reason is as follows. Monthly average price is the load weighted price of system load profile. If the load profiles of the customers of the retailer are different from the system load profile, the average cost per KWh of the retailer may be different from the monthly average electricity price due to the significant price difference between peak and off-peak hours. Therefore, if a month ahead average daily electricity price profile can be forecasted, retailers can set the month-to-month variable rate plan based on the load profiles of the customers. The rate can even be customized for each consumer based on his/her load profile.
- 2) For GenCos, electricity prices in different periods are essential for investment decision making. Electricity prices in different periods are taken into consideration in microgrid planning [2, 3]. The microgrid can benefit significantly from high prices in peak hours if the generation cost is lower than the wholesale market price [4]. Energy storage systems are cost-effective if the price differences between peak hours and off-peak hours are significant [5]. However, the monthly average price cannot provide enough information for microgrid planning.

Therefore, month ahead average daily electricity price profile forecasting is very essential.

1.2 Literature review and contributions

A variety of methods and ideas have been tried for electricity price forecasting with varying degrees of success [6]. The Global Energy Forecasting Competition (GEFCOM) has been held annually by the IEEE Working Group on Energy Forecasting since 2012. The theme of GEFCOM2014 is electricity price forecasting [7]. Electricity price forecasting can be categorized into short-term and medium-and-long term time scales.

Most research focuses on short-term electricity price forecasting. Numerous methods have been proposed in the past. There are mainly three types of methodologies, including the game theory model, simulation model and time series model [8]. In [9], a dynamic platform is proposed to foster the development of data-mining enhanced multi-agent systems. It is utilized to generate power load and settlement price prediction in day-ahead market in Greece. Time-series models are also popular in short-term price forecasting, which includes parsimonious stochastic models, regression or causal models and artificial intelligence (AI) models [8]. In [10], a new functional

forecasting method is proposed, which attempts to generalize the standard seasonal ARMAX time-series model to the L^2 Hilbert space. A neural network (NN) model is presented in [11]. In this model, different influential factors are the feedback. Historical prices from the financial market, weekly price/load information, historical loads and day type are chosen as the explanatory factors. A support vector machine (SVM) model considering the impacts of oil and natural gas prices is proposed in [12] to forecast daily electricity price. In [13], a genetic optimal regression of relevance vector machines (SRM) model is proposed. The final prediction model is the optimal linear assembly of several individual SRM models with different kernel functions.

With the continuous deepening of research, hybrid models that combine several methods have been developed. Hybrid models usually outperform individual forecasting models. A short-term forecasting of the electricity price with data-driven algorithms is studied in [14]. A stacked denoising autoencoder (SDA) model, a class of deep neural networks (DNN), and its extended version are utilized to forecast the daily electricity price profile. In [15], a hybrid architecture combining the advantages of autoregressive integrated moving average (ARIMA) models and the local learning techniques is proposed. A hybrid model is proposed in [16]. In this model, an artificial neural network (ANN), an adaptive neuro-fuzzy inference system and an autoregressive moving average (ARMA) are utilized to generate three independent price forecasts. A new data fusion algorithm is then proposed to combine them. A hybrid approach to construct prediction intervals of marginal clearing prices (MCPs) with a two-stage formulation is proposed in [17].

Although short-term electricity price forecasting has been well studied, only a few studies focus on mid-term price forecasting. Mid-term electricity price forecasting is much more complicated, whereas short-term price forecasting can take advantage of trends from the immediate past [18]. Data-driven approaches, which take impact factors as inputs, are prevalent in mid-term price forecasting. In [19], several methods with some economic data as inputs are utilized to forecast the monthly average price, and the best mean absolute percentage error (MAPE) in these methods is 12.97%. An SVM model considering calendar day, fuel prices, electric loads, weather and import/export power is proposed in [20], and the MAPE is 8.04%. A data-driven approach with two regression-based linear forecasting models is proposed in [21], and the MAPE is 9.67%.

Reference [22] forecasts the mid-term UK baseload electricity prices by forecasting the prices of each day and averaging the forecasts afterwards. As is known, block transaction is adopted in the UK spot market. The UK

baseload electricity price has a lower price fluctuation. Therefore, it is less sensitive to load variation and fuel price fluctuation. However, the day-ahead price in Electric reliability council of Texas (ERCOT) varies from time to time and day to day. It is highly sensitive to load variations and fuel price fluctuations. Month ahead daily load forecasting and daily fuel price forecasting is not accurate enough, which may result in low accuracy of hourly day-ahead price forecasting if we forecast the daily electricity prices of ERCOT and average the forecasts afterward. The month ahead monthly load forecasting accuracy is approximately 98%. The month ahead fuel price forecasting can learn from the future price, which usually has good results [23]. Hence, the advantage of forecasting the average daily electricity price profile directly instead of forecasting the hourly electricity prices and average the forecasts afterwards is clear due to the higher accuracy of monthly load forecasting and average fuel price forecasting.

To the best of our knowledge, up until now, month ahead average daily electricity price profile forecasting has not yet been investigated, but it merits in-depth study. The main contributions of this paper are:

- 1) Month ahead average daily day-ahead electricity price profile forecasting is proposed in this paper. A hybrid nonlinear regression and SVM model is proposed for month ahead average daily electricity price profile forecasting.
- 2) A nonlinear price regression model with deviation compensation is proposed to improve the forecast accuracy.
- 3) Off-peak hours, peak hours in peak months and peak hours in off-peak months are distinguished and different methods are adopted to further improve prediction accuracy.

1.3 Differences between the proposed average daily price profile forecasting and hourly price forecasting and averaging afterwards

The conventional mid-term average price forecasting is to forecast the hourly price for the next month and average the price afterwards. The distinctions between the proposed method and the conventional ones are as follows.

- 1) The forecast object

The conventional method forecasts the hourly prices of each day for next month, and then averages the forecast prices of each period, e.g., averages the forecast prices at 10:00 a.m. every day. In contrast, the proposed method directly forecasts the average price of a period, e.g., the average price at 10:00 a.m. for the next month.

- 2) The training set

For the conventional method, the training set includes the historical hourly price data for each day, e.g., 24-hour prices for one day. In contrast, the proposed method training set includes historical average daily price profiles, e.g., the average price at 10:00 a.m. for a month.

- 3) The historical impact factor data

For the conventional method, the historical impact factor data include the historical load profile of each day and the historical fuel price of each day. For the proposed method, the historical impact factor data include the historical average load profile of each month, historical average fuel price of each month, etc.

- 4) The forecasting accuracy of input factors

For the conventional method, the forecasting accuracy of the month ahead daily load profile and fuel price of each day is relatively low. However, the electricity price is strongly sensitive to the load and fuel prices. Forecasting errors of these input factors may result in non-negligible errors of the forecasted electricity price. For the proposed method, the forecasting accuracy of month ahead average daily load profile and month ahead average fuel price is relatively high. In other words, the forecasting errors of the input factors are small, and the forecasting error of the proposed method mainly originates from the modeling error.

Therefore, month ahead average daily price profile forecasting is proposed in this paper to improve forecast accuracy.

2 ERCOT electricity market and data sources

2.1 Overview of ERCOT electricity market

ERCOT was formed in 1970 and then became the central operating coordinator for Texas. With the deregulation of the electric power industry, ERCOT became an ISO in 1996 [24]. ERCOT developed the zonal wholesale market and then transformed to the nodal market in 2010. ERCOT manages the flow of electric power to 24 million Texas customers, covering approximately 90 percent of the state's electric load [25]. The total operational capacity is 78543 MW. The capacity fuel types in percentages are shown in Table 1.

ERCOT runs both day-ahead and real-time markets. The two-sided auction is adopted. Demand is allowed to bid in spot markets. In the day-ahead market, producers submit their offers and consumers submit their bids to ERCOT. Energy is co-optimized with ancillary services (AS) and



Table 1 Capacity fuel types in percentages in ERCOT

Fuel	Percentage	Fuel	Percentage
Natural gas	62.33	Coal	23.51
Nuclear	5.91	Wind	3.71
Wind-coastal	1.92	Solar	1.36
Hydro	0.52	Other	0.74

certain congestion revenue rights (CRR). The locational marginal price (LMP) in ERCOT has two components, which are the energy and congestion component.

The retail market in Texas is deregulated and well-developed. As of September 2014, 114 retailers were actively doing business in ERCOT [26]. In terms of rate structure, there are mainly three types of plans offered in the market, which are the fixed rate plan, variable rate plan (month-to-month) and indexed plan (market rate) [27].

2.2 Data sources

The data regarding the prices of natural gas delivered for electricity generation and system capacity of Texas are obtained from [28]. The data of consumer price index (CPI) and average wages of power plant operators are obtained from [29]. The data of lending rate in the USA are obtained from [30]. The data for day-ahead electricity prices and loads were obtained from [25].

Features for the predicted month should be predicted prior to electricity price forecasting. In these features, the lending rate and average wages change little and usually remain constant over a long period. However, the natural gas prices and CPI varies from month to month. The natural gas price may even fluctuate sharply at some times. Research regarding natural gas price forecasting has been reported and some well-performing methods are available [23, 31, 32]. CPI forecasting has also been well investigated in the economic field [33–35]. As this paper mainly focuses on the method of electricity price forecasting, forecasting of natural gas prices, CPI, average wages and lending rate for the predicted month are beyond the scope of this study. These parameters are assumed to be predicted accurately.

3 Nonlinear regression model with deviation compensation (NRM-DC)

3.1 Nonlinear regression model (NRM)

1) The compositions of costs

According to engineering economics, the cost is divided into the period cost and product cost [36]. The composition

of cost and the impact factor related to cost are shown in Fig. 1.

Period costs are defined as costs charged to expenses in the period in which they are incurred, and mainly consist of selling expenses, administrative expenses and financial expenses such as insurance and income tax expenses. Selling expenses and administrative expenses are mostly determined by the management level and price level. Financial expenses are closely related to the price level and lending rate.

Product costs consist of the costs involved in the purchase or production of goods, including the direct material, direct labor cost and manufacturing overhead. Direct material includes raw materials that can be processed into products. In power plants, direct material cost is directly related to fuel price. Direct labor cost is the wages paid to workers who produce the products and is directly related to social average wages of the industry. The remaining cost belongs to the manufacturing overhead, which is mainly influenced by the price level.

As analyzed above, cost (C_{cost}) can be described as a linear superposition of fuel price, price level, average wages and lending rate with different weights. It can be described as:

$$C_{cost} = a_1f + a_2w + a_3l + a_4c + a_5 \tag{1}$$

where f is fuel price; w is average wages; l is lending rate; c is CPI; a_1, a_2, a_3, a_4, a_5 are the fitting coefficients and a_5 is a fixed asset.

In ERCOT, electricity prices are driven to a large extent by changes in fuel prices, and natural gas prices in particular [37]. According to the natural gas prices, coal prices and heating rate, which can be obtained from the EIA, the comparison between costs of electricity generated with natural gas-fired and coal-fired power plants is shown in Fig. 2. Obviously, the cost of electricity generated with

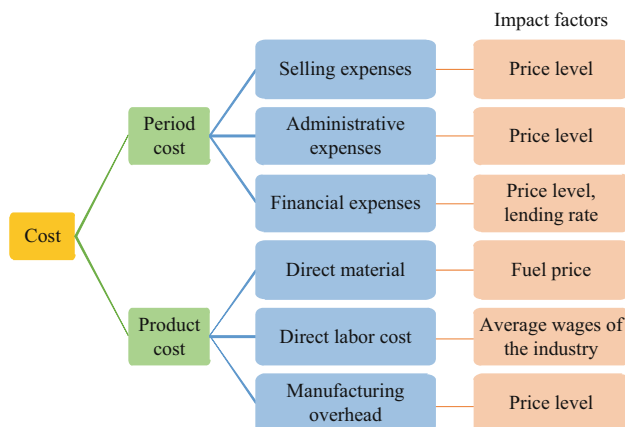


Fig. 1 Composition of cost and the impact factor related to cost

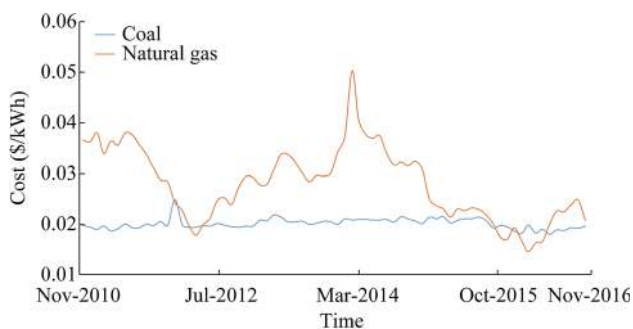


Fig. 2 Costs of electricity generated with natural gas and coal

natural gas is much higher than that generated with coal. As 62.33% of installed capacity fuel types in Texas is natural gas and only 23.51% of installed capacity fuel types is coal, coal-fired power plants cannot meet the system load. Typically, the fuel type of marginal units is natural gas. Therefore, the natural gas price is selected as the dominant impact factor in fuel price in this paper.

2) The relationship between price and cost

Cost is the basis of price, but it is not the only factor that influences price. Price is determined by supply and demand in economic theory [38]. Spot price can even be forecasted directly by modeling the supply curve and demand curve in reference [39]. Figure 3 illustrates the relationship between electricity prices and demand in the PJM electricity market from January 1, 1999 to December 31, 2000 [40]. As demand increases, the price grows slowly at the beginning and rises significantly when the demand is high. An exponential function is used in [41], a Box–Cox transformation model is used in [42] and a hockey-stick shaped model is used in [40] to describe the relationship between price and supply-demand situation. The exponential function is employed in this paper.

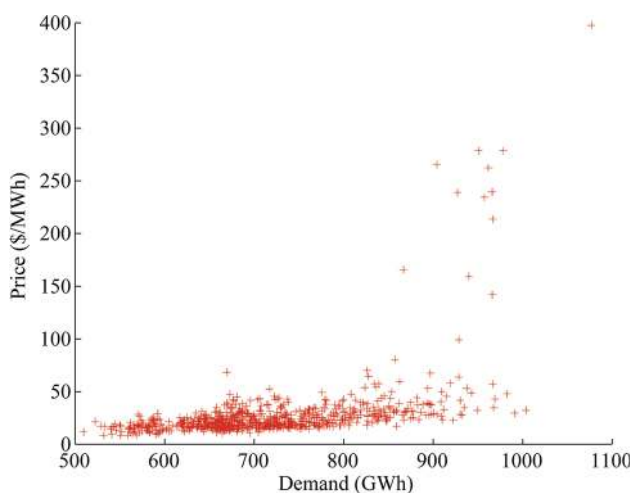


Fig. 3 Electricity prices and demand

3) Nonlinear regression model

The mean-reversion model can be used to describe electricity price and price fluctuation around the mean value [43]. Price (C_{price}) can be expressed by generation cost multiplying supply demand coefficient (SDC) (i.e. k_{SDC}), as in (2). It is worth noting that other forms of functions can also be used.

$$C_{price} = C_{cost}k_{SDC} \tag{2}$$

The relationship between supply and demand can be expressed by the average hourly loading rate in month horizon (AHLRMH) (i.e. ρ_{AHLRMH}), which is the average hourly load in a month horizon (P_{av}) divided by system capacity (P_s), as in (3).

$$\rho_{AHLRMH} = \frac{P_{av}}{P_s} \tag{3}$$

As each month has different numbers of days, the average hourly load in the month horizon rather than total hourly load in the month horizon is chosen. An example is given to show how to calculate average hourly load in month horizon. Assume that average hourly load at 10:00 a.m. in September (with 30 days) is needed. It is the sum of all the loads at 10:00 a.m. in September divided by 30.

Each hour of each month has its average loading rate (i.e. ρ_{AHLRMH}^{each}), but it cannot reflect the tension of supply and demand straightforwardly. Therefore, a benchmark average loading rate (BALR) (i.e. ρ_B) is defined. It is the average loading rates of all the hours of all the months. It is used to measure the tension of supply and demand. The supply and demand situation coefficient (SDSC) (i.e. k_{SDSC}) is proposed and is defined as (4).

$$k_{SDSC} = \frac{\rho_{AHLRMH}^{each}}{\rho_B} \tag{4}$$

SDSC is a number approximately 1.0 and can reflect the supply and demand situation straightforwardly. If SDSC is greater than 1.0, it means that there is a tense relationship between supply and demand in which prices may go up, and vice versa.

An exponential function is employed to describe the relationship between price and the supply-demand situation. The expression of SDC is as (5). It is worth noting that other forms of functions can also be used.

$$k_{SDC} = a^{k_{SDSC}-b} \tag{5}$$

By substituting (1) and (4) into (2), the electricity price can be expressed by a nonlinear regression model as:

$$C_{price} = (a_1f + a_2w + a_3l + a_4c + a_5)a_6^{k_{SDSC}-a_7} \tag{6}$$

where a_1, a_2, \dots, a_7 are the parameters to be fitted.



Due to the fluctuation of prices, to reduce the influence of price fluctuation, a logarithmic smoothing processing is applied, which is as:

$$\log(C_{\text{price}}) = \log(a_1f + a_2w + a_3l + a_4c + a_5) + (k_{\text{SDSC}} - a_7) \cdot \log a_6 \tag{7}$$

As the prices vary widely in different periods, the fitting may not perform well if all the historical data are used for fitting one model. Therefore, different periods are separated for regression in this paper. For example, to forecast prices at 10:00 a.m., only historical data at 10:00 a.m. are used for training.

3.2 Deviation compensation model based on SVM

It is widely acknowledged that there may be some system bias if only one model is employed to forecast the price. With the specific data for ERCOT from October 2010 to April 2016, examples of deviation analysis of April and September based on the proposed nonlinear regression model are shown in Fig. 4. The deviation is the actual price minus the predicted price.

It is apparent that the deviations exhibit certain distribution characteristics. Therefore, it is reasonable to adjust the results obtained from the nonlinear regression model. An improvement of the forecast accuracy can be expected. Therefore, a deviation compensation model is proposed.

As the deviation distribution is nonlinear and may be related to various factors, conventional regression methods may be ineffective.

As is known to all, SVM is an effective statistical machine learning method that is suitable for the high-order non-linear regression problem [44]. SVM is adopted to predict the deviations between predicted prices of the proposed nonlinear regression model and the actual prices. Natural gas price, CPI, average wages, lending rate, SDSC are preprocessed by principal component analysis (PCA) to extract the principal components [45]. The principal

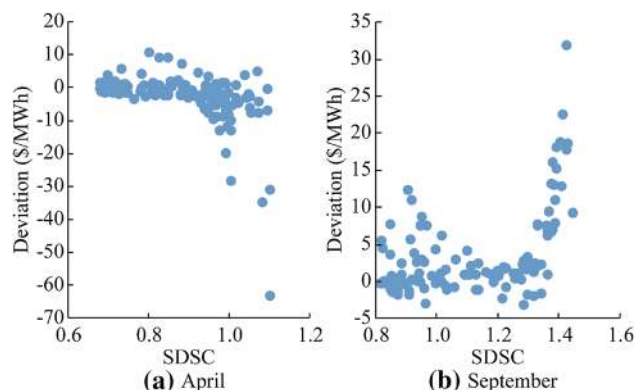


Fig. 4 Deviation analysis

components are selected as the input features of the SVM model.

The framework of the proposed deviation compensation model is shown in Fig. 5.

3.3 Performance evaluation

Root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) are widely used performance evaluation indices in forecasting. RMSE, MAE and MAPE are defined as:

$$RMSE = \sqrt{\frac{\sum_{n=1}^N |y_n - \hat{y}_n|^2}{n}} \tag{8}$$

$$MAE = \frac{1}{N} \sum_{n=1}^N |y_n - \hat{y}_n| \tag{9}$$

$$MAPE = \frac{1}{N} \sum_{n=1}^N \frac{|y_n - \hat{y}_n|}{y_n} \tag{10}$$

where N is the number of forecasted data; y_n is the actual value and \hat{y}_n is the forecasted value.

These three indices are adopted for performance evaluation in this paper.

3.4 ERCOT case study

Historical data from December 2010 to November 2016 are used in the case study. A rolling forecast for the month ahead average daily electricity price profiles from May 2016 to November 2016 is presented. Two benchmarks are designed to verify the performance of the proposed method. The actual price data from the same month and same periods in the previous year are used as the forecasted price in Benchmark 1. The actual price data of the same periods in the previous month are used as the forecasted

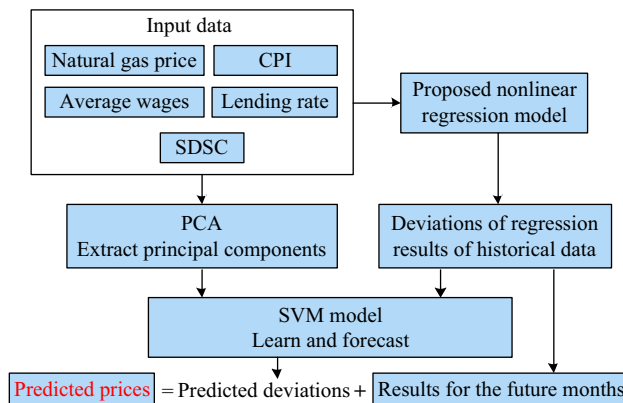


Fig. 5 Framework of the proposed deviation compensation model

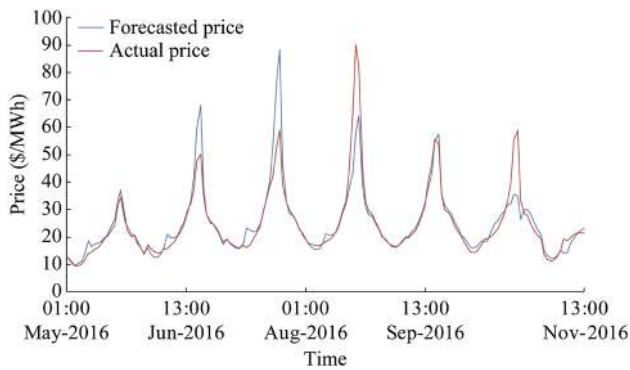


Fig. 6 Actual prices and forecasted prices based on NRM-DC

price in Benchmark 2. The results are shown in Figs. 6 and 7, respectively. Data in Fig. 6 are ordered by months and by periods.

RMSE and MAPE of NRM, NRM-DC, Benchmark 1 and Benchmark 2 are shown in Table 2.

It is apparent from Table 2 that the proposed NRM and NRM-DC models outperform Benchmark 1 and Benchmark 2 in this case. The forecasted prices coincide with the trend of the actual electricity prices. NRM-DC performs better than NRM, especially in off-peak periods (periods except period 7, 8, 15, 16, 17 and 18). It is apparent from Table 3 that the proposed NRM and NRM-DCR methods perform better than the SVM model in off-peak periods in this case.

However, errors of the peak hours (period 7, 8, 15, 16, 17 and 18) are much higher than those of the off-peak hours. A specialized model is proposed in Section 4 to forecast prices in peak hours.

4 Price forecasting for peak hours

The supply and demand situation is usually tense in peak hours, which often results in price spikes. The prices in off-peak hours may be similar to the prices of the same period in the previous time window [46], but the average electricity prices in peak hours may become remarkably higher than usual. Large forecasting errors are produced. The price spike is considered to be an abnormal price in many studies and several models have been proposed for price spike forecasting [47, 48]. Price forecasting methods for peak hours should be differentiated from those for off-peak hours.

As is shown in Fig. 6, the maximum hourly average price is approximately 9 times the minimum hourly average price. The hourly average prices of periods 7, 8, 15, 16, 17 and 18 from December 2010 to April 2016 are shown in Figs. 8 and 9. Price spikes can deteriorate the fitness of nonlinear regression model.

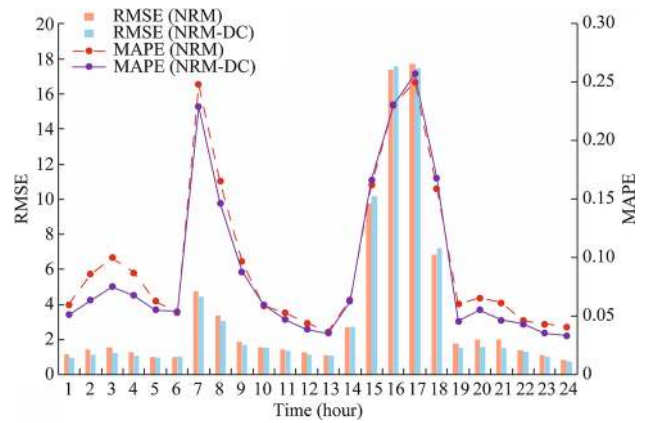


Fig. 7 Forecasted results of each hour

Table 2 RMSE and MAPE of NRM, NRM-DC, benchmark 1 and benchmark 2

Method	RMSE	MAE	MAPE (%)
NRM	5.90	2.91	9.74
NRM-DC	5.89	2.80	8.99
Benchmark 1	18.04	5.07	15.32
Benchmark 2	7.48	3.90	13.95

Table 3 RMSE and MAPE of NRM, NRM-DC and SVM model in off-peak periods

Method	RMSE	MAE	MAPE (%)
NRM	1.55	1.25	6.24
NRM-DC	1.39	1.09	5.33
SVM model	1.80	1.33	6.57

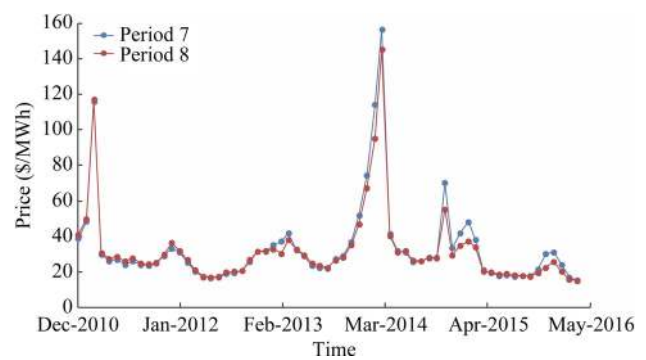


Fig. 8 Month ahead average hourly prices of periods 7 and 8

As one can observe, prices in peak hours are not always very high. According to the prices and SDSC, months can be divided into peak months and off-peak months. Peak months of periods 7 and 8 are December, January and February in winter. Peak months of periods 15 and 18 are

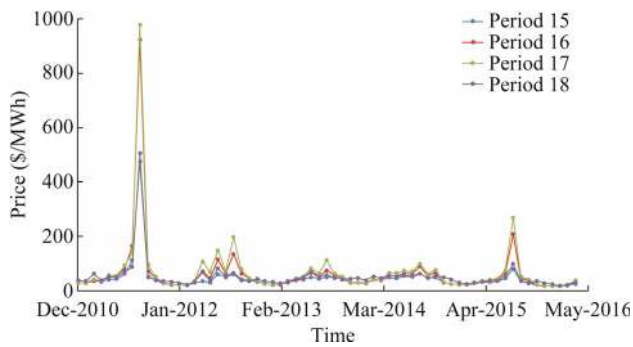


Fig. 9 Month ahead average hourly prices of periods 15, 16, 17 and 18

July and August in summer. Peak months of periods 16 and 17 are June, July and August in summer. Prices in peak hours of off-peak months are rational, whereas prices in peak hours of peak months are extremely high. Therefore, different forecasting methods should be adopted to forecast the prices in different scenarios.

4.1 Price forecasting for peak hours in off-peak months

Prices of peak hours in peak months are significantly high, which deteriorates the fitness of nonlinear price regression model. However, prices of peak hours in off-peak months still exhibit strong patterns, which means the NRM-DC model may remain applicable.

Historical data of off-peak months are used for training. The results are shown in Tables 4, 5 and 6, respectively.

The proposed method for forecasting prices of peak hours in off-peak months can significantly improve the prediction accuracy in this case.

4.2 Price forecasting for peak hours in peak months

Prices of peak hours in peak months can be extremely high and uncertain. The relationship between price and SDSC can be strongly nonlinear. SVM shows advantages for strong nonlinear problems. Therefore, it is employed to forecast the prices of peak hours in peak months.

The influences of CPI, wages and lending rate on prices become minor as SDSC increases. When supply and demand are tight, prices are mainly related to supply, demand and variable costs. Therefore, only SDSC, natural gas prices, month and period are selected as the input features. Results are shown in Table 7.

It is apparent that the RMSE and MAPE of the SVM model are much lower than those of the other two models.

Table 4 Forecasting results of periods 7, 8 in off-peak months

Period	Method	RMSE	MAE	MAPE (%)
Period 7	NRM-DC with peak months	4.43	3.78	22.92
	NRM-DC without peak months	3.00	2.29	13.66
Period 8	NRM-DC with peak months	3.06	2.53	14.67
	NRM-DC without peak months	2.63	2.03	11.58

Table 5 Forecasting results of periods 15, 18 in off-peak months

Period	Method	RMSE	MAE	MAPE (%)
Period 15	NRM-DC with peak months	4.56	3.98	11.11
	NRM-DC without peak months	2.48	2.30	6.77
Period 18	NRM-DC with peak months	7.32	5.51	16.95
	NRM-DC without peak months	3.59	2.90	8.77

Table 6 Forecasting results of periods 16, 17 in off-peak months

Period	Method	RMSE	MAE	MAPE (%)
Period 16	NRM-DC with peak months	10.17	6.44	13.90
	NRM-DC without peak months	7.12	4.94	10.39
Period 17	NRM-DC with peak months	12.43	8.65	18.03
	NRM-DC without peak months	9.96	6.90	12.59

Table 7 Forecasting results for peak hours in peak months

Method	RMSE	MAE	MAPE (%)
Forecasting results based on NRM-DC with peak months	19.96	18.09	30.73
Forecasting results based on NRM-DC without peak months	19.75	17.32	28.94
Forecasting results based on SVM method	12.55	11.23	20.73

5 Framework and final forecasting results of hybrid nonlinear regression and SVM model

Section 3 proposes a nonlinear regression model with deviation compensation. Different methods are employed to forecast prices of peak hours in peak months in Section 4. A hybrid nonlinear regression and SVM model is proposed to synthesize the advantages of these methods. In summary, the framework of the hybrid model is shown in Fig. 10.

Historical data are used for training forecasting models. Data of off-peak hours are used as the training sets to forecast the prices of off-peak hours based on the NRM-DC model. Data of peak hours in off-peak months are used to forecast the prices of peak hours in off-peak months based on the NRM-DC model. All of the data are used to forecast the prices of peak hours in peak months based on the SVM model.

The forecast results based on the hybrid model proposed in this paper are shown in Table 8.

It is apparent that the proposed hybrid model performs better than the NRM-DC and SVM models in this case based on the ERCOT dataset.

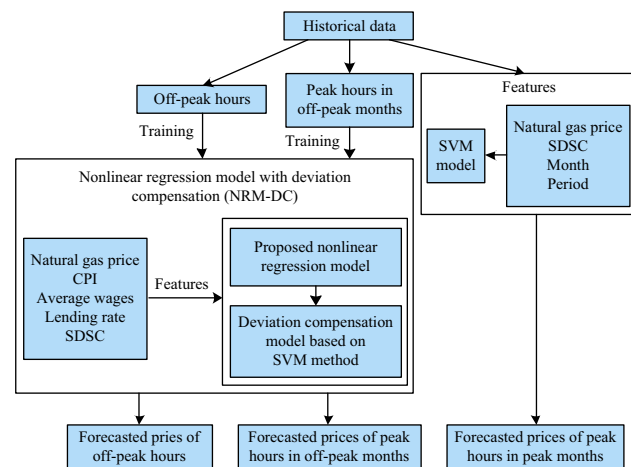


Fig. 10 Framework of hybrid nonlinear regression and SVM model

Table 8 Forecast results based on hybrid model and a brief comparison with other models

Method	RMSE	MAE	MAPE (%)
The proposed hybrid model	3.95	2.10	7.30
NRM-DC	5.89	2.80	8.99
SVM model only	4.33	2.49	8.63

6 Conclusion

Month ahead average daily electricity price profile forecasting is an essential task for retailers and investment decision making in electricity markets. A hybrid nonlinear regression and SVM model is proposed in this paper for month ahead average hourly price forecasting. In this model, prices of different periods in different months are forecasted by different methods. Three methods are adopted for prices of off-peak hours, peak hours in off-peak months and peak hours in peak months. A nonlinear price regression model with deviation compensation is proposed to forecast the prices of off-peak hours and prices of peak hours in off-peak months. SVM is adopted to forecast the prices of peak hours in peak months. The case study suggests that the hybrid method proposed in this paper performs well in month ahead average daily electricity price profile forecasting based on the ERCOT dataset.

Future work will investigate the applications of month ahead average daily electricity price profile forecasting, especially in retail market and investment decision.

Acknowledgements This work was supported by National Natural Science Foundation of China (No. 51537005) and State Grid Corporation of China “Research on the model and application of power supply and demand technology under the market trading environment”.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- [1] Wang J, Zhong H, Tang W et al (2017) Optimal bidding strategy for microgrids in joint energy and ancillary service markets considering flexible ramping products. *Appl Energy* 205:294–303
- [2] Khodaei A, Bahramirad S, Shahidehpour M (2014) Microgrid planning under uncertainty. *IEEE Trans Power Syst* 30(5):2417–2425
- [3] Wang J, Zhong H, Xia Q et al (2017) Optimal planning strategy for distributed energy resources considering structural transmission cost allocation. *IEEE Trans Smart Grid*. <https://doi.org/10.1109/TSG.2017.2685239>
- [4] Wu QH, Zheng J, Jing Z (2015) Coordinated scheduling of energy resources for distributed DHCs in an integrated energy grid. *CSEE J Power Energy Syst* 1(3):59–67
- [5] Yajing G, Fushen X, Wenhai Y et al (2017) Optimal operation modes of photovoltaic-battery energy storage system based power plants considering typical scenarios. *Prot Control Modern Power Syst* 2(2):397–406



- [6] Weron R (2014) Electricity price forecasting: a review of the state-of-the-art with a look into the future. *Int J Forecast* 30(4):1030–1081
- [7] Hong T, Pinson P, Fan S et al (2016) Probabilistic energy forecasting: global energy forecasting competition 2014 and beyond. *Int J Forecast* 32(3):896–913
- [8] Aggarwal SK, Saini LM, Kumar A (2009) Electricity price forecasting in deregulated markets: a review and evaluation. *Int J Electr Power Energy Syst* 31(1):13–22
- [9] Chatzidimitriou KC, Chrysopoulos AC, Symeonidis AL et al (2012) Enhancing agent intelligence through evolving reservoir networks for predictions in power stock markets. In: *International workshop on agents and data mining interaction*, Springer, Berlin, Germany, pp 228–247
- [10] Portela J, Munoz A, Alonso E (2017) Forecasting functional time series with a new Hilbertian ARMAX model: application to electricity price forecasting. *IEEE Trans Power Syst*. <https://doi.org/10.1109/tpwrs.2017.2700287>
- [11] Hassan S, Khosravi A, Jaafar J et al (2014) Electricity load and price forecasting with influential factors in a deregulated power industry. In: *Proceedings of 9th international conference on system of systems engineering*, Adelaide, Australia, 9–13 June 2014, pp 79–84
- [12] Shiri A, Afshar M, Rahimi-Kian A et al (2015) Electricity price forecasting using support vector machines by considering oil and natural gas price impacts. In: *Proceedings of 2015 IEEE international conference on smart energy grid engineering*, Oshawa, Canada, 17–19 August 2015, pp 1–5
- [13] Alamaniotis M, Bargiotas D, Bourbakis NG et al (2015) Genetic optimal regression of relevance vector machines for electricity pricing signal forecasting in smart grids. *IEEE Trans Smart Grid* 6(6):2997–3005
- [14] Wang L, Zhang Z, Chen J (2016) Short-term electricity price forecasting with stacked denoising autoencoders. *IEEE Trans Power Syst*. <https://doi.org/10.1109/tpwrs.2016.2628873>
- [15] Vaccaro A, El-Fouly THM, Cañizares CA et al (2015) Local learning-ARIMA adaptive hybrid architecture for hourly electricity price forecasting. In: *Proceedings of PowerTech*, Eindhoven, Netherlands, 29 June–2 July 2015, pp 1–6
- [16] Darudi A, Bashari M, Javidi MH (2015) Electricity price forecasting using a new data fusion algorithm. *IET Gener Transm Distrib* 9(12):1382–1390
- [17] Wan C, Xu Z, Wang Y et al (2014) A hybrid approach for probabilistic forecasting of electricity price. *IEEE Trans Smart Grid* 5(1):463–470
- [18] Yan X, Chowdhury NA (2013) Mid-term electricity market clearing price forecasting: a hybrid LSSVM and ARMAX approach. *Int J Electr Power Energy Syst* 53(1):20–26
- [19] Torbaghan SS, Motamedi A, Zareipour H et al (2012) Medium-term electricity price forecasting. In: *Proceedings of 2012 North American power symposium*, Champaign, USA, 9–11 September 2012, pp 1–8
- [20] Mohamed A, El-Hawary ME (2016) Mid-term electricity price forecasting using SVM. In: *Proceedings of 2016 IEEE Canadian conference on electrical and computer engineering*, Vancouver, Canada, 15–18 May 2016, pp 1–6
- [21] Torghaban SS, Zareipour H, Le AT (2010) Medium-term electricity market price forecasting: a data-driven approach. In: *Proceedings of 2010 North American power symposium*, Arlington, USA, 26–28 September 2010, pp 1–7
- [22] Maciejowska K, Weron R (2016) Short- and mid-term forecasting of baseload electricity prices in the U.K.: the impact of intra-day price relationships and market fundamentals. *IEEE Trans Power Syst* 31(2):994–1005
- [23] Buchanan WK, Hodges P, Theis J (2001) Which way the natural gas price: an attempt to predict the direction of natural gas spot price movements using trader positions. *Energy Econ* 23(3):279–293
- [24] Daneshi H, Srivastava AK (2011). ERCOT electricity market: transition from zonal to nodal market operation. In: *Proceedings of 2011 IEEE power and energy society general meeting*, Detroit, USA, 24–29 July 2011, pp 1–7
- [25] ERCOT (2017) Texas. <http://www.ercot.com/>. Accessed 12 June 2017
- [26] Public Utility Commission of Texas (2015) Report to the 84th Texas legislature: scope of competition in electric markets in Texas. Public Utility Commission of Texas. Available via DIALOG. http://www.puc.texas.gov/industry/electric/reports/scope/2015/2015scope_elec.pdf. Accessed 12 June 2017
- [27] Utili-Facts: types of electric plans, electric reliability council of Texas. <http://www.puc.texas.gov/consumer/facts/factsheets/elecfacts/Electricplans.pdf>. Accessed 12 June 2017
- [28] U.S. Energy Information Administration (2017) Electricity, Washington, DC. <https://www.eia.gov/electricity/>. Accessed 12 June 2017
- [29] Bureau of labor statistic. <https://www.bls.gov/regions/southwest/texas.htm>. Accessed 12 June 2017
- [30] Trading Economics. <https://tradingeconomics.com>. Accessed 12 June 2017
- [31] Taal M, Bulatov I, Klemeš J et al (2003) Cost estimation and energy price forecasts for economic evaluation of retrofit projects. *Appl Therm Eng* 23(14):1819–1835
- [32] Amjady N, Hemmati M (2006) Energy price forecasting-problems and proposals for such predictions. *IEEE Power Energy Mag* 4(2):20–29
- [33] Iqbal M, Naveed A (2016) Forecasting inflation: autoregressive integrated moving average model. *Eur Sci J* 12(1):83–92
- [34] Xun XX, Zhang DS, Wang Y et al (2014) CPI forecasting based on combination model of TS fuzzy neural network. *J Shaanxi Univ Sci Technol (Nat Sci Ed)* 32(3):173–176
- [35] Koop GM (2013) Forecasting with medium and large Bayesian VARs. *J Appl Econ* 28(2):177–203
- [36] Park Chan S (2014) *Contemporary engineering economics*. Pearson, London
- [37] Potomac Economics, Ltd (1999) 2015 state of the market report for the ERCOT wholesale electricity markets. Potomac Economics, Ltd. Available via DIALOG. <https://www.potomaceconomics.com/wp-content/uploads/2017/01/2015-ERCOT-State-of-the-Market-Report.pdf>. Accessed 19 Jun 2017
- [38] Mankiw NG (2008) *Principles of economics*, 5th edn. Cengage Learning, Boston
- [39] Ziel F, Steinert R (2016) Electricity price forecasting using sale and purchase curves: the X-model. *Energy Econ* 59(3):435–454
- [40] Kanamura T, Ôhashi K (2007) A structural model for electricity prices with spikes: measurement of spike risk and optimal policies for hydropower plant operation. *Energy Econ* 29(5):1010–1032
- [41] Skantze P, Ilic M, Gubina A (2004) Modelling locational price spreads in competitive electricity markets; applications for transmission rights valuation and replication. *IMA J Manag Math* 15(4):291–319
- [42] Barlow MT (2002) A diffusion model for electricity prices. *Math Finance* 12(4):287–298
- [43] Deng S (2000) Stochastic models of energy commodity prices and their applications: mean-reversion with jumps and spikes. University of California Energy Institute, Berkeley
- [44] Yang M, Lin Y, Zhu S et al (2015) Multi-dimensional scenario forecast for generation of multiple wind farms. *J Mod Power Syst Clean Energy* 3(3):361–370
- [45] Wang J, Zhong H, Lai X et al (2017) Exploring key weather factors from analytical modeling toward improved solar power

forecasting. *IEEE Trans Smart Grid*. <https://doi.org/10.1109/tsg.2017.2766022>

- [46] Ziel F (2016) Forecasting electricity spot prices using lasso: on capturing the autoregressive intraday structure. *IEEE Trans Power Syst* 31(6):4977–4987
- [47] Zhao JH, Dong ZY, Li X et al (2007) A framework for electricity price spike analysis with advanced data mining methods. *IEEE Trans Power Syst* 22(1):376–385
- [48] Koban V, Zlatar I, Pantoš M et al (2015) A remark on forecasting spikes in electricity prices. In: *Proceedings of 2015 12th international conference on the European energy market (EEM)*, Lisbon, Portugal, 19–22 May 2015, pp 1–5

Ziming MA received his Bachelor of Engineering degree in electrical engineering from Tsinghua University, Beijing, China, in 2015, where he is currently pursuing the Ph.D. degree. His research interests include electricity price forecasting, electricity market and energy internet.

Haiwang ZHONG received his B.E. and Ph.D. degrees from Tsinghua University, China in 2008 and 2013 respectively. He is currently working as a Associate Professor at Department of Electrical Engineering, Tsinghua University. His research interests include: Electricity market, smart grid, power system operations, demand response, demand dispatch, and renewable energy integration. He received the 2014 ProSPER.Net-Scopus Young Scientist Award.

Le XIE received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2004, the M.Sc. degree in engineering sciences from Harvard University, Cambridge, MA, in June 2005, and the Ph.D. degree from the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA in 2009. Currently, he is an Associate Professor at the Department of Electrical and Computer Engineering, Texas A&M University, College Station. His industry experience included an internship at ISO-New England and an internship at Edison Mission Energy Marketing and Trading. His research interest includes modeling and control of large-scale complex systems, smart grids application with renewable energy resources, and electricity markets.

Qing XIA received the Ph.D. degree from the Electrical Engineering Department of Tsinghua University, Beijing, China, in 1989. He is now a Professor at the same university. His research interests include electricity market, generation scheduling optimization, power system planning, and load forecasting.

Chongqing KANG received the Ph.D. degree from the Department of Electrical Engineering in Tsinghua University, Beijing, China, in 1997. He is now a Professor at the same university. His research interests include low-carbon electricity, power system planning, power market, power system reliability, and power system load forecasting.

