Available online at www.sciencedirect.com
ScienceDirect
Physica A 382 (2007) 98–113
PHYSICA A
www.elsevier.com/locate/physa

Price forecast in the competitive electricity market by support vector machine

Ciwei Gao[a,b,*], Ettore Bompard[b], Roberto Napoli[b], Haozhong Cheng[a]

[a]Department of Electrical Engineering, Shanghai Jiaotong University, China
[b]Dipartimento di Ingegneria Elettrica Politecnico di Torino, Italy

Available online 4 April 2007

Abstract

The electricity market has been widely introduced in many countries all over the world and the study on electricity price forecast technology has drawn a lot of attention. In this paper, with different parameter C_i and ε_i assigned to each training data, the flexible C_i Support Vector Regression (SVR) model is developed in terms of the particularity of the price forecast in electricity market. For Day Ahead Market (DAM) price forecast, the load, time of use index and index of day type are taken as the major factors to characterize the market price, therefore, they are selected as the inputs for the flexible SVR forecast model. For the long-term price forecast, we take the reserve margin R^m, *HHI* and the fuel price index as the inputs, since they are the major factors that drive the market price variation in long run. For short-term price forecast, besides the detailed analysis with the young Italian electricity market, the new model is tested on the experimental stage of the Spanish market, the New York market and the New England market. The long-term forecast with the SVR model presented is justified by the forecast with the data from the Long Run Market Simulator (LREMS).

Keywords: Electricity market; Flexible C_i SVR; Price forecast; Particle swarm optimization

1. Introduction

Price forecast is a challenging task and is very important in competitive electricity market. Both the market players and the regulators concern very much about the price evolution. On the one hand, the market price prediction is a crucial information for the producers' production arrangement and bidding strategies, e.g. optimal schedule for storage of hydro energy production with reference to the hydrology, or unit commitment definition with regard to the minimum power constraints and the flexibility constraints of the thermal plants; as far as the bidding strategy, both the supply and the demand sides need the price information to adjust their price submission to gain more profit or hedge the bidding risk. Hence the accuracy of price forecast greatly impacts the players' benefit. On the other hand, the regulators need to analyze the market behavior and monitor the market evolution with the price forecast tool.

*Corresponding author. Department of Electrical Engineering, Shanghai Jiaotong University, China.
E-mail address: ciwei.gao@polito.it (C. Gao).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2007.03.050

The price forecast tools are classified into two categories. One is the detailed market simulation [1–4], which needs lot of market information. The other type of price forecast technology refers to those mathematical approaches without a thorough market modeling, but which try to discover the relation between some known inputs and the electricity price. The most popular approach is the time series algorithm, which has been extensively studied [5,6]. Artificial Intelligence (AI) approaches such as neural network, and support vector machine (SVM), which have been successfully applied in load forecast, are also suitable for price forecast [7].

In an another respect, price forecast can be classified into three types in terms of the time frame (short/medium/long term). Apparently, the short-term price forecast is the most elusive, which is due to the incomplete information or uncertain bidding strategy of the market participants. With the time frame enlarged, the price is more and more reasonable with reference to the corresponding context, the uncertainty of which will be a major difficulty for the prediction.

SVM is especially suitable for solving problems of small sample size [8]. In this paper, we propose a flexible C_i Support Vector Regression (SVR) model for price forecast based on classic ε-intensive SVR for forecasting the Day Ahead Market (DAM) price. According to the particularity of the price forecast, uniform C and ε have been replaced by various C_i and ε_i that are optimally assigned to each sample considered with the help of Particle Swarm Optimization (PSO). The flexible C_i SVR model has also been applied to long-term price forecast for Italian electricity market. The results show the effectiveness of the proposed SVM model to solve the long-term price forecast with some special market indexes (reserve margin R^m[1], HHI[2] and fuel price).

The remainder of the paper is organized as follows: Section 2 briefly introduces the SVR approach and presents the flexible C_i SVR model with regard to the particularity of the electricity price forecast. Section 3 analyses the methods for selecting the appropriate parameter C and ε, and an effective approach for our new model has been proposed; in Section 4 the data input for the DAM price forecast and the long run price forecast have been defined based on a careful analysis. Section 5 presents some numerical examples focusing on the Italian market, and the reference results with the Spanish market, the New England (NE) market and New York (NY) market are also provided. Finally, conclusions are drawn in Section 6.

2. SVM for electricity price forecast

SVM is firmly grounded in the framework of statistical learning theory or VC theory, which has been developed over the last three decades by Vapnik and Chervonenkis [9], and Vapnik [10,11]. Generally speaking, SVM is to minimize the structural risk instead of the usual empirical risk by minimizing an upper bound of the generalization error and it obtains an excellent generalization performance [9,12]. Moreover, SVM is especially suitable for solving problems of small sample size [8], and has already been used for classification [13], regression [14] and time series prediction [15]. Based on the classic ε-intensive SVR model [9], the flexible C_i SVR specially designed for price forecast in the electricity market is proposed.

SVR is to map the input data x into a higher dimensional feature space through a nonlinear mapping Φ and then a linear regression problem is obtained and solved in this feature space. With the given training data $\{(x_1, y_1), \dots (x_i, y_i), \dots, (x_n, y_n)\}$, the mapping function can be formulated as

$$f(x) = \sum_{i=1}^{n} \omega_i \Phi_i(x_i) + b, \tag{1}$$

where ω_i and b are the parameters that need to be defined. ε-SVR is to find a function $f(x)$ that has at most ε deviation from the actually obtained targets y_i for all the training data and at the same time is as flat as possible. Flatness in this case means to reduce the model complexity by minimizing $||\omega||^2$, so that we can write

[1]Reserve margin, the ratio between maximum capacity and the maximum load.

[2]A commonly accepted measure of market concentration expressed as: $HHI = s_1^2 + s_2^2 + s_3^2 + \dots + s_n{}^2$ (where s_n is the market share of the ith firm).

this problem as an optimization problem:

$$\begin{aligned} &\text{Min} \quad \tfrac{1}{2}||\omega||^2, \\ &\text{s.t.} \quad \begin{cases} y_i - \Phi(\omega, x_i) - b \leqslant \varepsilon, \\ \Phi(\omega, x_i) + b_i - y \leqslant \varepsilon, \end{cases} \end{aligned} \tag{2}$$

which means we do not care about errors as long as they are less than ε, but will not accept any deviation larger than this. To be more realistic, one can add slack variables ξ_i, ξ_i^* $i = 1,2,3,\ldots,n$, to cope with otherwise infeasible constraints of the optimization problem (2). Hence we arrive at the formulation stated in Ref. [11]:

$$\begin{aligned} &\text{Min} \quad \tfrac{1}{2}||\omega||^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*), && (3.1) \\ &\text{s.t.} \quad \begin{cases} y_i - \Phi(\omega, x_i) - b \leqslant \varepsilon + \xi_i, & (3.2) \\ \Phi(\omega, x_i) + b_i - y \leqslant \varepsilon + \xi_i^*, & (3.3) \\ \xi_i, \xi_i^* \geqslant 0, & (3.4) \end{cases} \end{aligned} \tag{3}$$

where C is a positive constant as regularization parameter. The optimization formulation can be transformed into a dual problem [12] and the solution is expressed as

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_i, x) + b, \tag{4}$$

where α_i,α_i* are the dual variables with reference to constraints (3.2) and (3.3) and $0 \leqslant \alpha_i$, $\alpha_i{*} \leqslant C$. The constant C is the trade-off between the flatness of f and the amount up to which deviation larger than ε is tolerated.

The kernel function $K(x_i, x_j) = \Phi(x_i)\Phi(x_j)$ satisfies the Mercer's conditions and performs the non-linear mapping. Those sample points that appear with non-zero coefficients in (4) are the so called support vectors (SV). Kernel function can be a Gaussian function, a polynomial function, or a sigmoid function etc. In our SVR model, we apply the Gaussian kernel as (5), the convenience of which has been demonstrated in Refs. [16,17]:

$$K(x_i, x_j) = \exp\left(-\frac{||x_i - x_j||}{2p^2}\right). \tag{5}$$

The classic ε-intensive SVR will give a good performance in mapping with the appropriate parameter set (C, ε, p) based on the precondition that all the data in the training set are equally treated with respect to the uniform deviation ε which might be appropriate for the load forecast [18], since, in a rather long time, the mapping relation can be taken as constant, but for the price forecast, the mapping changes much faster, and different history samples have different mapping effectiveness with respect to the forecast accuracy. Therefore, we will make some changes to the classic ε-intensive SVR model to allow the model introduce the mechanism of different weight allocation for each sample. Namely, if the data are more relevant for the prediction, there would be bigger corresponding C_i, smaller ε_i and vice versa. A new model with different C_i, ε_i, assigned to each sample is, therefore, developed as following:

$$\begin{aligned} &\text{Min} \quad \tfrac{1}{2}||\omega||^2 + \sum_{i=1}^{n}C_i(\xi_i + \xi_i^*), && (6.1) \\ &\text{s.t.} \quad \begin{cases} y_i - \Phi(\omega, x_i) - b \leqslant \varepsilon_i + \xi_i, & (6.2) \\ \Phi(\omega, x_i) + b_i - y \leqslant \varepsilon_i + \xi_i^*, & (6.3) \\ \xi_i, \xi_i^* \geqslant 0, & (6.4) \end{cases} \end{aligned} \tag{6}$$

the solution is similar to (4), but α_i,α_i* are the dual variables with reference to the constraints (6.2), (6.3), respectively, and various C_i are considered for the dual as $0 \leqslant \alpha_i$, $\alpha_i{*} \leqslant C_i$. As shown in Fig. 1, different values of C and ε will result in completely different regression function, which will surely influence the forecast results hugely.

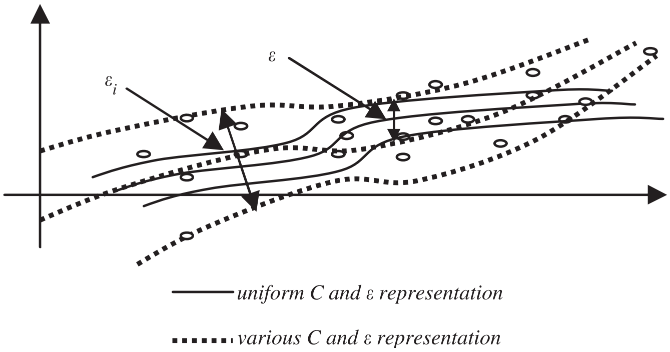


Fig. 1. Flexible C_i SVR with various C_i and ε_i and Vapnik SVR with uniform C and ε.

3. Parameters setting for the flexible c_i SVR

As we know, the choices of the parameters will largely influence the generalization performance of the SVM. But in our case, the generalization has different meaning with respect to the traditional one, the purpose of our model is to forecast the future, therefore, the generalization performance should be related to the forecast accuracy. Some traditional parameters selection methods are:

1. Parameter selection based on a priori knowledge or user expertise. Obviously, the approach is not easy to use by an inexperienced operator.
2. k-fold cross validation [19] and leave-one-out (LOO) [20], both approaches are very popular in machine learning algorithms. They are effective but very demanding computationally, especially in our model where there are so many parameters (various C_i and ε_i) to be defined.
3. Practical selection of the parameters based on statistics research, the value of ε should be proportional to the input noise level, whereas C is interpreted as a traditional regularization parameter. Practical formulas for parameters selection are deduced in Ref. [21] to avoid the heavy computation burden and get rather good results.

Since our model has to introduce various C_i and ε_i to each sample, the parameters that need to be defined will be much more than those in the traditional Vapnik ε-intensive model. Considering the computation constraints, the empirical formula deduced in Ref. [21] could be a good choice, where,

(1) for the width parameter p of the Gaussian kernel,

$$p \sim (0.1-0.5)\mathrm{range}(x), \tag{7}$$

for multivariate d-dimensional problems, the width parameter $p\sim$(0.1–0.5), where all d input variables are prescaled to [0,1].

(2) For the hyperparameter C,

$$C = \max(|\bar{y}+3\sigma_y|, |\bar{y}-3\sigma_y|), \tag{8}$$

where $\bar{y}$ and σ_y are the mean and the standard deviation of the y values of the training data.

(3) For the hyperparameter ε,

$$\varepsilon = 3\sigma\sqrt{\frac{\ln n}{n}}, \tag{9}$$

and σ is the standard deviation of noise, which can be estimated as

$$\sigma^2 = \frac{n^{1/5}k}{n^{1/5}k-1}\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2, \tag{10}$$

where n is the number of the training data, $\hat{y}_i$ is an estimated value by k-nearest-neighbors regression. With respect to the model (6), C_i and ε_i for each sample i is solved considering that since C determines the trade-off between the empirical loss and structural risk, the total effect should be kept.

$$\sum_{i=1}^{n} C_i = nC. \tag{11}$$

The proportion determination of C_i among the samples will be explained in Section 4 with the help of a PSO algorithm [22], which is approved to have better performance than the normal evolutionary-based optimization algorithm [23].

Once C_i is determined, Eq. (10) can be modified as

$$\sigma_i^2 = \frac{n^{1/5}k}{n^{1/5}k - 1} \frac{1/C_i}{\sum_{i=1}^{n} 1/C_i} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2, \tag{12}$$

and with (9)

$$\varepsilon_i = 3\sigma_i \sqrt{\frac{\ln n}{n}}. \tag{13}$$

Therefore, we can obtain ε_i for each sample. The basic idea is that, the samples with stricter (smaller) ε_i should have larger C_i corresponded and vice versa. Hence we assume an inverse proportion between C_i and ε_i.

As a way of simulation, the known data are divided into two sets Λ_1 and Λ_2 according to the time sequence. The first n_1 data in Λ_1 are taken as the training data, which will be used to define the forecast function by flexible C_i SVR, while the left n_2 data in set Λ_2 is the test data, which will be used to help PSO determine the parameters C_i, ε_i with reference to the data in set Λ_1. With the optimized C_i and ε_i, the SVR model is identified with Λ_1 and predict the unknown set Λ_3 (see Fig. 2).

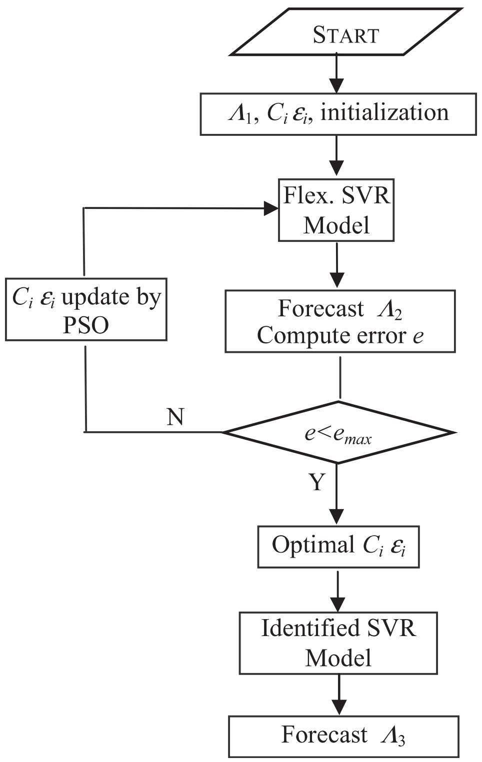


Fig. 2. Flow chart of $C_i\varepsilon_i$ optimization.

4. Input data analysis for SVR price forecast

It is very important to define and select meaningful inputs for the price forecast. For the short run DAM price forecast, only publicly known data are used, namely, the data of history/forecast load, history price, in which the short-term load forecast is of high accuracy with an error of 1–2% [1], hence it will be appropriate to take the actual load for price prediction for the test. For the long run price forecast, some related economic indexes that determine/reflect the market price in the long run can be the inputs for the price forecast.

4.1. Inputs design for short-term price forecast

Studying the market and the price evolution, we find that the price shows similar behavior within the class defined by load profile; it is assumed that with similar load profile, the market players take similar bidding strategies at least in the short run. It means that the prices may experience huge differences when they are not in the same class even if the corresponding loads are the same. Fig. 3 reports the load-price profile in the Italian market, look at points A and B, from profile point of view, they are not in the same class; A belongs to work day, and B belongs to the vacation day. A although the load $L_A = L_B$, the prices are very different. On the other hand, the impacts of the variation of the bidding strategies, and regulations cannot be precisely inferred. To alleviate the related negative effects, smaller size sampling is considered. In Fig. 3, the load profiles of 13/09/2004–17/09/2004 and 20/09/2004–24/09/2004 are very similar, but the corresponding prices are very different, obviously, the error cannot be diminished by enlarging the number of samples only by having smaller sampling size or to give the latest samples bigger weights. The essence of the problem is that we are not able to precisely take into account all the factors to fully represent the problem, bigger weight assigned to the latest samples that contains more reliable information would possibly improve the forecast accuracy.

Before the liberalization, electricity tariffs were defined by the Italian Regulatory Authority according to a classification of the different hours of the year into four "time of use", named F1 (peak load), F2 (mid-peak load), F3 (mid-base load), and F4 (base load). Market prices of the Italian power exchange in 2004 showed some relation with the "time of use". See Fig. 4, even with the same load level, the prices are different according to different time of use.

With all the aspects analysed above, the inputs (see Table 1) to characterize the market price are defined with the following procedures:

(a) define the number of history days that can be used for price prediction, e.g. 30 days;
(b) the inputs $X(t) = X(L_t,B_t,S_t,t)$ for characterizing the price $Y(t) = y_t$ $(t = 1,\ldots,n)$, n is the number of the training data.
  - L_t: load level, obviously it is one of the most important drive for the price variation;
  - B_t: coefficient with relation to time of use, the relation between the time of use and the market price is a particular phenomenon of Italian market at early stage;
  - S_t: load profile index;
  - t: time index, to indicate the effectiveness of the data with respect to the corresponding time.

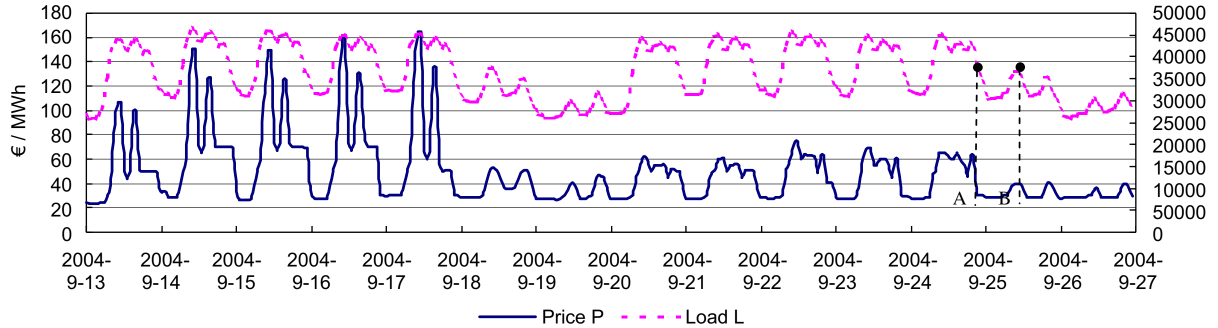


Fig. 3. Price load profile (13/09/2004–26/09/2004, Italian Electricity Market).

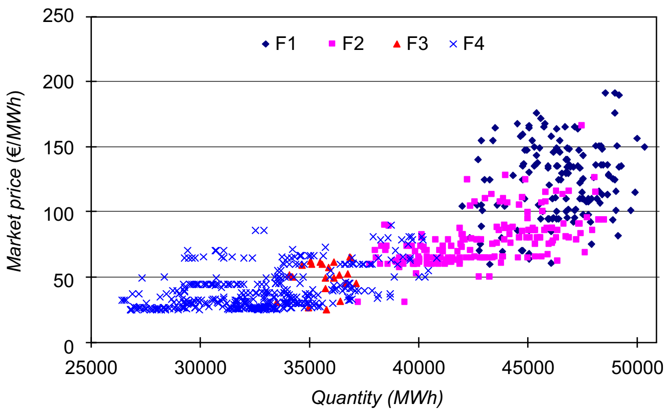


Fig. 4. Market prices in the different time bands (July 2004).

Table 1
Input, output definition for the flexible C_i SVR price forecast model

Input $X(t)$				Distance	Output Y_t
L_t	B_t	S_t	t	D	y_t
L_1	B_1	S_1	1	$D_{1,n+1}$	y_1
L_2	B_2	S_2	2	$D_{2,n+1}$	y_2
…	…	…	…	…	…
L_{n-1}	B_{n-1}	S_{n-1}	$n-1$	$D_{n-1,\ n+1}$	y_{n-1}
L_n	B_n	S_n	n	$D_{n,\ n+1}$	y_n
L_{n+1}	B_{n+1}	S_{n+1}	$n+1$	0	$\hat{y}_{n+1}$

(c) Normalize each item of the input, and compute the Euclidean distance D between the input (X_{n+1}) of the day to be forecasted and the history data, and list them in descending order. With the distance decreasing, the ranking index $I = 1, 2,\ldots,n$ are assigned to the training data with the distance from the most to the least.
More similar the load (L_t), the time of use coefficient (B_t), the load profile (S_t) and closer the time (t) with respect to those of X_{n+1}, the distance ($D_{t,n+1}$) is smaller and the index I is bigger, accordingly, the corresponding sample will be given bigger C_i. It can be interpreted by the function :

$$C_i = g(I_i)nC, \tag{14}$$

$$\sum_{i=1}^{n} g(I_i) = 1, \tag{15}$$

where $g(I_i)$ is a monotone increasing function and (15) is for complying with (11).

(d) PSO for C_i allocation
To obtain the optimal $g(I_i)$, PSO is resorted to. With respect to $g_i = g(I_i)$ $(i = 1,2,\ldots,n)$, the particle is a vector that contains n elements and corresponds to a type of mapping $\boldsymbol{g}^j$, which is updated in each iteration j. For each particle:
Current position: $\boldsymbol{g}^j = (g_1^j, g_2^j, \ldots, g_n^j)$.
Best individual position: $\boldsymbol{P}^j = (P_1^j, P_2^j, \ldots, P_n^j)$ considering the history of only this particle.
Best global position: $\boldsymbol{P}_g^j = (P_{g1}^j, P_{g2}^j, \ldots, P_{gn}^j)$ considering the history of all the particles in the swarm.

Flying velocity: $\boldsymbol{V}^j = (v_1^j, v_2^j, \ldots, v_n^j)$.

$$\boldsymbol{V}^{j+1} = \omega \times \boldsymbol{V}^j + \mu_1 \times \text{rand}() \times (\boldsymbol{P}^j - \boldsymbol{g}^j) + \mu_2 \times \text{rand}() \times (\boldsymbol{P}_g^j - \boldsymbol{g}^j), \tag{16}$$

$$\boldsymbol{g}^{j+1} = \boldsymbol{g}^j + \boldsymbol{V}^{j+1}, \tag{17}$$

$$\boldsymbol{g}^{j+1} = \boldsymbol{g}^{j+1} / \sum_{i=1}^{n} \boldsymbol{g}_i^{j+1}, \tag{18}$$

where, j is the iteration number, (16) is to update the velocity, (17) is to update the particle position, and (18) is to comply with (15). With the iteration going on, $\boldsymbol{g}^j$ is optimized with the forecast accuracy improvement of the test data set. The parameters ω, μ_1, and μ_2 play important roles for the optimization performance, hence they should be selected carefully [24,25]. We also study the particles initialization, considering (11), $g(I_i)$ is the function to allocate the share of nC to each training data with regard to the different distance index I_i. Since we think that the data with larger I_i are more relevant, the particles are initialized by those assumed monotone increasing functions in Fig. 5, where each function represents one type of particle.

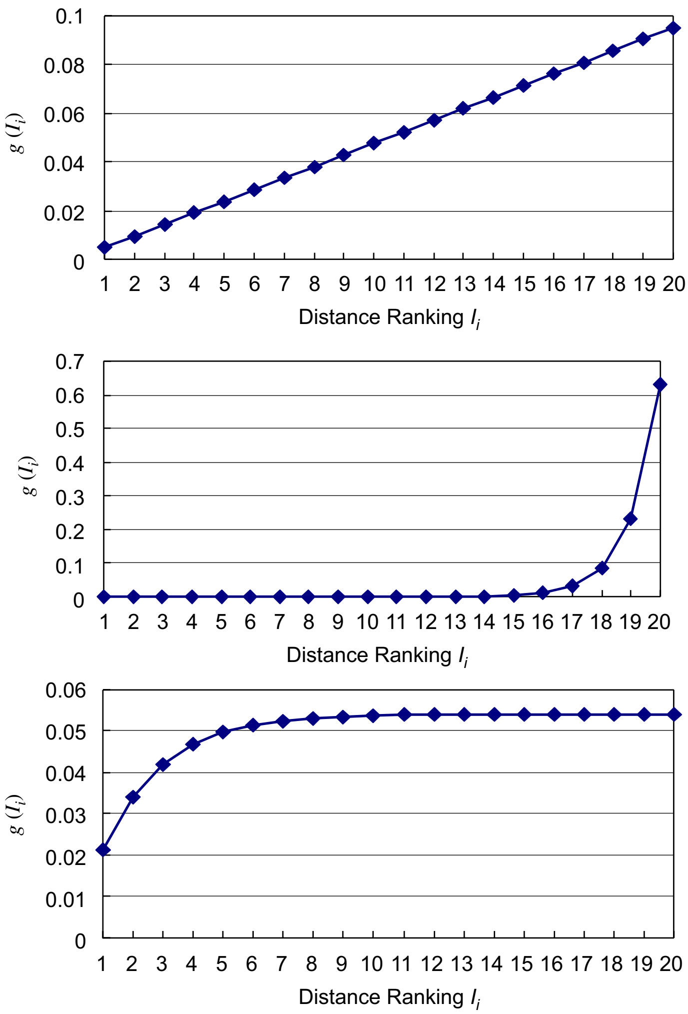


Fig. 5. C_i allocation function $g(I_i)$ initialization for PSO optimization, $\boldsymbol{g}(\boldsymbol{I}_i)$ is a monotone increasing function.

(e) Knowing the training data X_t, Y $(t = 1,2,\ldots,n)$ and ε, the flexible SVR model can be solved. With the discovered function and $_{n+1}$, the forecast $\hat{y}_{n+1}$ can be obtained. Note that data in are normalized to avoid the side effect of the number value and is not included in when we refer to the mapping of the SVR, it is only used to indicate the priority of newer data for the distance definition.

4.2. Inputs design for long-term price forecast

For the long run price forecast, we are not going to forecast the price of a specific hour, but only to predict the average price of a certain period (i.e. yearly average price). The factors that cause the price variation are analyzed in following three respects.

The strategic behavior of the players can be possibly represented by some market indexes like *HHI* (market concentration index), which varies in long run and can be an important reason to explain the behavior of the market players and hence the variation of the price as well.

Supply–demand relation is always an essential and direct factor to define the price in every market and it can be represented by the reserve margin R^{m}.

In most cases, the market prices are defined by the thermal plants, hence the fuel price ρ plays a very important role in determining the price and is taken as another key drive for the price variation. The flexible C_i *SVR* model is also applied to the long run price forecast. X_t $(HHI_t, R_t^{\mathrm{m}}, \rho_t)$ is the input and the yearly average price $Y_t(y_t)$ is the output.

5. Numerical examples for price forecast with flexible C_i SVR

In this section, the flexible C_i SVR and classic ε-intensive SVR models are both applied to the short-term (DAM price forecast) and long-term price forecast.

For the short run price forecast, it is interesting to study the various market behavior in the transition stage of the Italian market from the market's unveiling (1st April 2004) to 31st March 2005. The flexible C_i SVR model and the Vapnik ε-intensive SVR model are compared with their application to the data of the New England market (February–December, 2004) and the New York market (February–December, 2004). Moreover, to justify the forecast accuracy of the new model, we forecast the Spanish market with contrast to the results of the two ARIMA models reported in Ref. [5].

For the long run price forecast, due to the limited database, we have to make use of the 40 years simulation results of the Long Run Electricity Market Simulator (LREMS) [26], where the future scenarios are assumed, to test the flexible C_i SVR algorithm for long-term price forecast. The target is to verify that our new model can give a good mapping between the input (R^{m}, ρ and *HHI*) and output (yearly average price y) and hence provide a tool other than the simulator to forecast the long run electricity price.

The popular index to measure the forecast is the mean absolute percentage error (MAPE), and the square root of the forecast mean square error (FMSE), which are defined as:

$$\mathrm{MAPE} = \frac{100}{N}\sum_{i=1}^{N}\frac{|\lambda_A - \lambda_F|}{\lambda_A}, \tag{19}$$

$$\sqrt{\mathrm{FMSE}} = \sqrt{\sum_{i=1}^{N}(\lambda_A - \lambda_F)^2}, \tag{20}$$

where N is the number of the forecast hours, λ_A is the actual price, and λ_F is the forecast price. To avoid the difficulty caused by $\lambda_A = 0$, the expression of the index MAPE is modified as

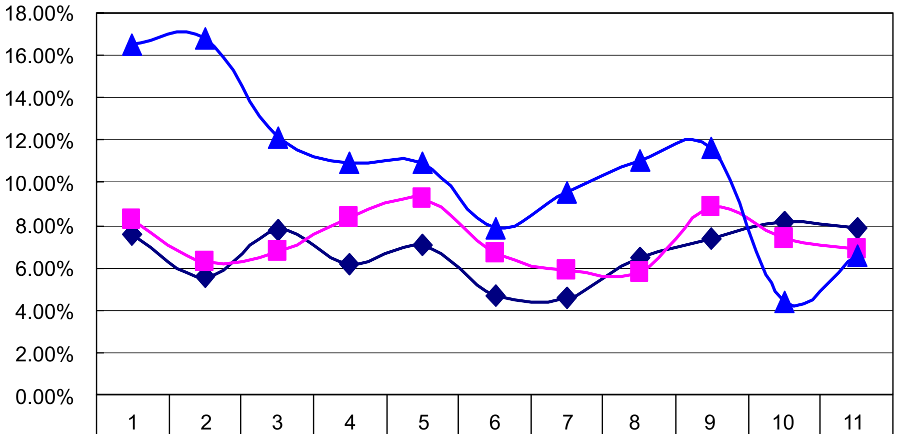
$$\mathrm{MAPE} = \frac{100}{N}\sum_{i=1}^{N}\frac{|\lambda_A - \lambda_F|}{\bar{\lambda}_A}, \tag{21}$$

where $\bar{\lambda}_A$ is the average actual price.

5.1. Numerical examples for short run price forecast

5.1.1. Price forecast for studying the Italian market evolution

Fig. 6 shows the price forecast results for Italy, New England, and New York markets with our new model. The average forecast errors are different with difference months. Usually, the more volatile price is more difficult to be forecasted, hence bigger MAPE. In the Italian market, it takes on an apparently decreasing tendency. From the results of the new model, we see that, at beginning (May, June), the average error is rather huge (about 16%) and gradually it decreases to 10% or so (July 2004–Jan 2005), and finally it reached to about 5%. The less and less prediction error can be possibly explained by the fact that at the beginning, the market players in the market are trying to become familiar with market, the bidding strategies are different and the corresponding market results seem to be very volatile. Gradually, they become more used to the market and the market behavior seems to be more tangible and less elusive, which will surely result in the price forecast accuracy improvement as we find in Fig. 6. In the later evolution of the market, the possible change of the market regulation and other external context change will bring some instability to the price and hence propose more difficulties to price forecast. Figs. 7–9, respectively, show the typical price profiles and forecast in the early, middle and late stage of the studied time frame. The price profile in case A (Fig. 7) is rather volatile. There is even one whole day where the price stays at zero price, the prices seems randomly scattered and hard to be forecasted. The flexible C_i SVR model obtained MAPE equal to 27.10%, a little higher than that by the Vapnik model. In case B (Fig. 8), the price has some transition among the period but much less volatility than the case A, the flexible C_i model captures the transition much more efficiently than the Vapnik model by a 10% accuracy improvement. Case C in Fig. 9 seems to be in a stable period, the price profile is rather mild and the forecast accuracy of our model is 4.82% as shown in Table 2. Generally speaking, the market is rather young, and the market participants are not experienced for their strategic bidding, hence compared with the load, the price profiles are much more volatile. With the volatility of the load, market structure, and season change, the Italian market still needs much time to let the players become familiar with. In contrast to the Italian market, the price forecasts for the other two markets have much better results and their MAPE are much less volatile, which can be explained by the fact that the other two markets have longer history, hence the price is more regular and easier to be forecasted.



	1	2	3	4	5	6	7	8	9	10	11
New England Market	7.51%	5.52%	7.74%	6.21%	7.02%	4.64%	4.56%	6.43%	7.40%	8.13%	7.88%
New York Market	8.26%	6.22%	6.77%	8.33%	9.29%	6.64%	5.91%	5.72%	8.87%	7.34%	6.83%
Italian Market	16.55	16.83	12.18	10.94	10.98	7.87%	9.57%	11.02	11.65	4.36%	6.60%

Fig. 6. MAPE during May 2004–March 2005.

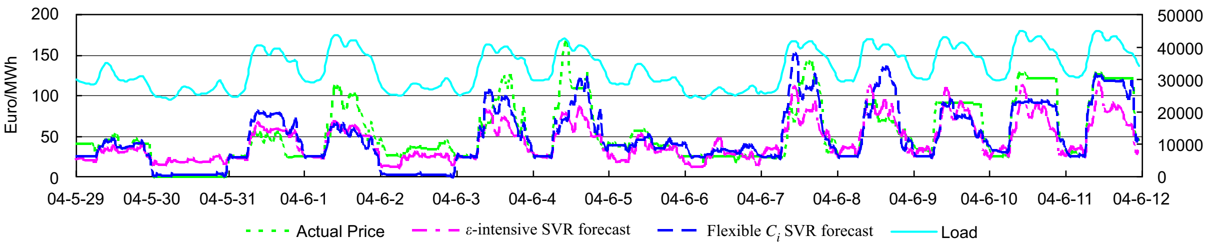


Fig. 7. Case A 29th May–12th June 2004 price forecast in early stage of the studied time frame.

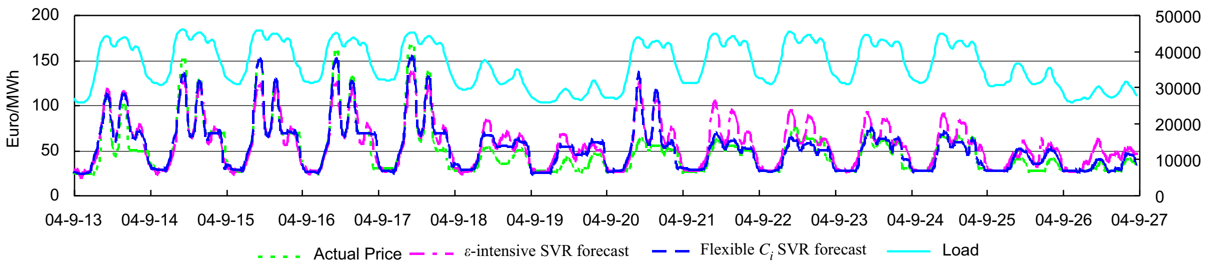


Fig. 8. Case B 13th Sept.–27th Sept. 2004 price forecast in middle stage of the studied time frame.

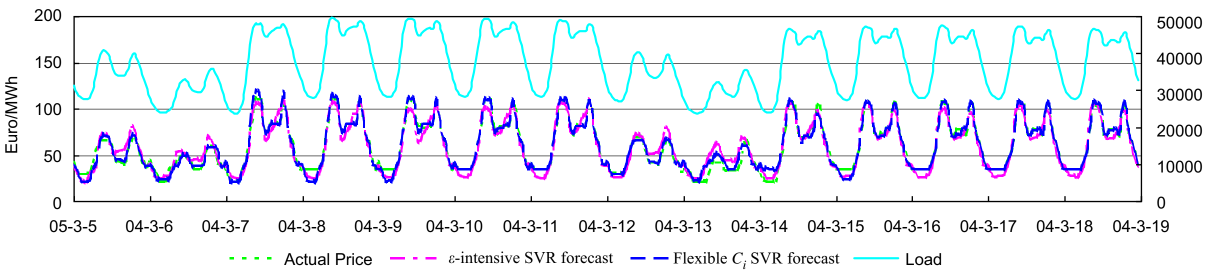


Fig. 9. Case C 13th Sept.–27th Sept. 2004 price forecast in late stage of the studied time frame.

Table 2
MAPE in three typical cases (%)

	ε-intensive SVR	Flexible C_i SVR
Case A	31.53	27.10
Case B	23.69	14.43
Case C	10.56	4.82

5.1.2. Results comparison between the ε-intensive SVR and the flexible C_i SVR

Table 3 shows the results between classic ε-intensive SVR (with the practical parameter selection method [21]) and the flexible C_i model. The flexible C_i SVR model always presents better results than the Vapnik ε-intensive SVR. Especially, when the market is not so mature and the volatility of the price is high, the new model can capture much better the actual price, e.g. in the Italian market, there is almost 6% more accurately obtained by the flexible C_i SVR model. In the longer time operated market, there is much less price volatility, the Vapnik model can produce good results, although still 1.5–2.0% less accurate than that of the flexible C_i model. The reason for these behaviors is that, the Vapnik model treats the history training data equally, while actually, the different history data play different roles with reference to the price to be forecasted, especially

Table 3
MAPE comparison between Vapnik intensive ε-SVR and flexible C_i SVR

Market	ε-intensive SVR (%)	Flexible C_i SVR (%)
Italian market	17.45	10.71
New England	8.48	6.98
New York	9.16	7.39

Table 4
Price forecast result comparison between ARIMA models and flexible C_i SVR in Spanish market 2000

	MAPE			$\sqrt{\text{FMSE}}$		
	SVR (%)	ARIMA1 (%)	ARIMA2 (%)	SVR	ARIMA1	ARIMA2
Jan	8.33	12.06	9.97	62.52	71.98	64.72
Feb	5.89	8.05	8.13	32.96	36.77	45.1
Mar	9.29	11.28	10.50	63.47	71.75	71.57
Apr	16.22	19.37	14.68	55.25	61.51	45.24
May	6.22	4.99	7.75	28.03	19.91	33.25
Jun	10.81	9.97	10.80	91.93	81.14	80.99
Jul	9.12	9.39	8.83	55.77	42.59	41.8
Aug	7.08	8.17	9.39	42.79	48.13	49.35
Sept	7.93	12.01	10.72	50.07	70.82	65.5
Oct	9.85	13.63	13.69	51.60	80.33	77.57
Nov	6.64	7.32	9.88	39.80	47.51	73.73
Aver.	8.85	10.57	10.39	52.20	57.49	58.98

when the price is volatile. Particularly designed for the price forecast task, the flexible C_i model tries to catch the different role of the training data by different C_i and ε_i assigned, and hence it obtained better results.

5.1.3. *Results comparison between the ARIMA models and the flexible C_i SVR*

Two ARIMA models are proposed in Ref. [5]; one is ARIMA model with explanatory variables and the other one without, and they both are tested with the data of the Spanish market. Table 4 shows the comparison of the result achieved by our model and the results provided by Ref. [5].

In Table 4, SVR refers to the flexible C_i model and ARMIAs 1 and 2 are, respectively, the ARIMA model without and with explanatory variables. With the index MAPE and $\sqrt{\text{FMSE}}$, our model presents, obviously, better results than the two ARIMA models.

5.1.4. *Typical $\boldsymbol{g}(\boldsymbol{I})$ assignment*

In the Italian market, with the training data of 1–20 April 2004 and the test data of 21–30 April 2004, Fig. 10 shows the typical $\boldsymbol{g}(\boldsymbol{I})$ assignment, which is efficiently found by the PSO approach.

Obviously, with reference to the distance D, the data closer to the point to be forecasted are assigned larger $g(I_i)$, hence the corresponding C_i is bigger as well.

5.2. *Long-term price forecast for Italian market*

Based on the basic data of Italian market in 2004, load, fuel price increase by 2% each year till 2043, LREMS obtained the evolution of the market with HHI_i, R_i^{m} and average yearly price y_i. Related data are shown in Appendix A. We suppose that the first 20 years are known, and we need to forecast for the last 20 years. The data of the first 15 years are used as the training data and the left 5 years are the data for C_i and ε_i optimization. Note that all the data has been preprocessed to be not greater than 1 by dividing the maximum value in its column (Appendix A Table A.1).

Fig. 11 shows the plot of the R^m, HHI, and the yearly average price, the forecast results by the flexible C_i SVR and the Vapnik ε-intensive SVR. The flexible SVR presents accurate forecast with the $MAPE$ equal to 3.28%. The precise forecast demonstrates that the flexible SVR successfully discovered the relation hidden behind the indexes and the price produced by the LREMS. Moreover, the result by the flexible SVR is much better than that by the Vapnik ε-intensive SVR (18.61%). Different C_i setting gives to the model the flexibility to distinguish the effectiveness of each data according to the test set forecast. Uniform C in Vapnik ε-intensive SVR cannot make the distinction among the data, and takes all the history data with the same effectiveness.

Fig. 12 shows the optimal value of C_i, ε_i for year 2004–18 to obtain the best forecast for the year 2019–23. From this plot we find that some of the years are much more important to forecast the future, like in 2006, 2007 and 2017, C_i are, respectively, 4.2018, 7.7394 and 4.2384, rather higher than the other years (0.0002), which corresponds to the large α_i-α_i^* value and small ε_i, hence the data of those years are not only the support vectors ($\alpha_i-\alpha_i^*\neq 0$) but also contribute the most to the forecast.

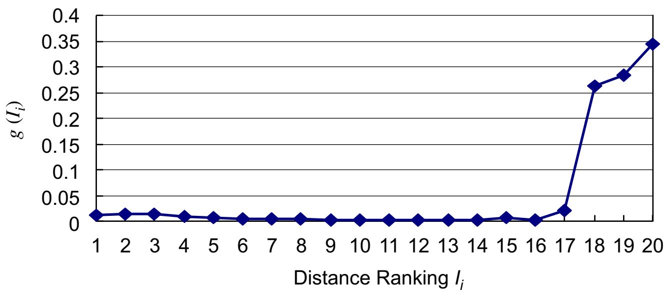


Fig. 10. Optimal $g(I_i)$ found by PSO.

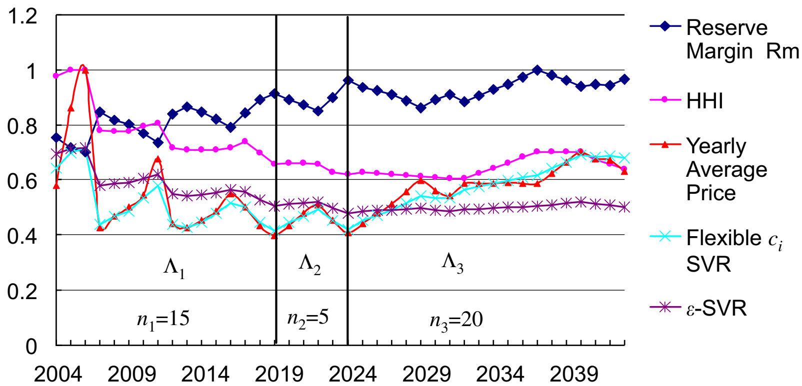
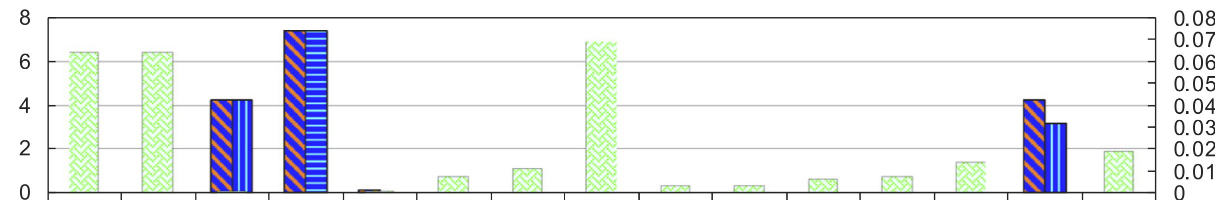


Fig. 11. Long-term price forecast with the data by LREMS based on Italian electricity market, 2004.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_i	0.0002	0.0002	4.2018	7.3794	0.1067	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	4.2384	0.0002
α_i-α_i^*	0.0002	0	4.2018	7.3794	0	0.0002	0	0.0002	0	0	0.0002	0.0002	0.0002	3.1767	0
ε_i	0.0644	0.0644	0.0008	0	0.0004	0.0072	0.0108	0.069	0.0028	0.0028	0.0059	0.0075	0.0138	0.0001	0.0192

Fig. 12. Identified parameters in the flexible C_i SVR model.

6. Conclusions

In this paper, a new model based on the classic ε-intensive SVR is proposed to the particular price forecast task. The classic Vapnik model aims to obtain the regression function with the best generalization by appropriate uniform parameters setting, the generalization being based on the uniform importance among the data. But in electricity price forecast, the history data have different importance, the data from the same type of day, or the newer data are more relevant, and this distinction should be highlighted. The flexible C_i model assigns various C_i, ε_i to each training data and hence is able to capture the distinction effectively. Moreover, combined with the practical empirical method for parameter selection, PSO with particularly designed initialization are presented to find the optimal values of the parameters C_i, ε_i efficiently.

Using the flexible C_i model, we forecast and analyse the market price of the Italian market, the NY market, and the NE market. Compared with the Vapnik ε-intensive SVR, the new model always presents much better results with the same platform of practical parameters selection. Moreover, with the numerical test, the flexible C_i SVR model is proved to have a better performance than the two ARIMA models. For long-term price forecast, the flexible C_i SVR discovered the mapping between the input (reserve margin, fuel price, *HHI*) and output (yearly average market price) with the data provided by LREMS. From this point of view, the crucial point to forecast the long-term price is how to define the future context correctly. In another angle, the regulator can make right policy that will influence those factors (*HHI*, reserve margin) to make a 'perfect' market in terms of the foreseen market behavior.

Acknowledgment

This research has been supported by the European Commission under Grant ASI/B7-301/98/679-026-ECLEE Project.

Appendix A

Long-term price forecast of Italian electricity market by LREMS, notice here the fuel price ρ is recorded by unit representation in 2004 and 2% increase in each year till 2034 (Table A.1).

Table A.1
40 years Italian market simulation by LREMS

Year	R^{m}	*HHI*	ρ	y (€/MWh)
2004	1.1691	3770	1.0000	53.84
2005	1.1091	3864	1.0200	79.99
2006	1.0873	3864	1.0404	92.62
2007	1.3134	3020	1.0612	39.47
2008	1.2641	3001	1.0824	43.37
2009	1.2411	2994	1.1041	46.29
2004	1.1691	3770	1.0000	53.84
2005	1.1091	3864	1.0200	79.99
2006	1.0873	3864	1.0404	92.62
2007	1.3134	3020	1.0612	39.47
2008	1.2641	3001	1.0824	43.37
2009	1.2411	2994	1.1041	46.29
2010	1.1881	3071	1.1262	50.69
2011	1.1406	3116	1.1487	62.49
2012	1.3020	2766	1.1717	40.85
2013	1.3382	2742	1.1951	39.70
2014	1.3135	2738	1.2190	42.10
2015	1.2704	2746	1.2434	45.16
2016	1.2225	2773	1.2682	50.76
2017	1.3066	2859	1.2936	46.51

Table A.1 (*continued*)

Year	R^m	HHI	ρ	y (€/MWh)
2018	1.3813	2696	1.3195	40.41
2019	1.4163	2541	1.3459	36.77
2020	1.3804	2561	1.3728	40.10
2021	1.3503	2550	1.4002	44.48
2022	1.3164	2541	1.4282	47.17
2023	1.3944	2429	1.4568	41.91
2024	1.4895	2401	1.4859	37.83
2025	1.4507	2420	1.5157	41.04
2026	1.4299	2408	1.5460	44.52
2027	1.4084	2394	1.5769	47.52
2028	1.3745	2384	1.6084	51.60
2029	1.3368	2365	1.6406	55.53
2030	1.3811	2347	1.6734	52.00
2031	1.4101	2345	1.7069	50.14
2032	1.3696	2333	1.7410	54.33
2033	1.4049	2413	1.7758	54.31
2034	1.4355	2483	1.8114	54.53
2035	1.4671	2556	1.8476	54.87
2036	1.5087	2635	1.8845	54.48
2037	1.5482	2709	1.9222	54.43
2038	1.5178	2709	1.9607	57.88
2039	1.4881	2709	1.9999	61.51
2040	1.4561	2715	2.0399	65.19
2041	1.4661	2600	2.0807	62.51
2042	1.4600	2534	2.1223	62.27
2043	1.4959	2467	2.1647	58.65

References

[1] D.W. Bunn, Forecasting demands and prices in competitive power markets, Proc. IEEE 88 (2) (2000) 163–169.

[2] S. de la Torre, A.J. Conejo, J. Contreras, Simulating oligopolistic pool-based electricity markets: a multiperiod approach, IEEE Trans. Power Syst. 18 (4) (2003) 1547–1555.

[3] A. Botterud, M. Korpas, K.-O. Vogstadetc, A dynamic simulation model for long-term analysis of the power market, 14th Power Systems Computation Conference, 2002, June 24–28, Sevilla, Spain.

[4] J. Valenzuela, M. Mazumdar, On the computation of the probability distribution of the spot market price in a deregulated electricity market, PICA 2001 innovative computing for power—electric energy meets the market. The 22nd International Conference on Power Industry Computer Applications, pp 268–271.

[5] J. Contreras, R. Espínola, F.J. Nogales, A.J. Conejo, ARIMA models to predict next-day electricity prices, IEEE Trans. Power Syst. 18 (3) (2003) 1014–1020.

[6] A.J. Conejo, M.A. Plazas, R. Espinola, A.B. Molina, Day-ahead electricity price forecasting using the wavelet transform and ARIMA models, IEEE Trans. Power Syst. 20 (2) (2005) 1035–1042.

[7] B.R. Szkuta, L.A. Sanabria, T.S. Dillon, Electricity price short-term forecasting using artificial neural networks, IEEE Trans. Power Syst. 14 (3) (1999) 851–857.

[8] N. Chen, W. Lu, J. Yang, G. Li, Support Vector Machine in Chemistry, World Scientific Publishing, Singapore, 2005.

[9] V. Vapnik, A. Chervonenkis, Theory of Pattern Recognition, Nauka, Moscow, 1974.

[10] V. Vapnik, Estimation of Dependences Based on Empirical Data, Springer, Berlin, 1982.

[11] V. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 1995.

[12] V. Vapnik, Statistical Learning Theory, Wiley, New York, 1998.

[13] S.-X. Lu, X.-Z. Wang, A comparison among four SVM classification methods: LSVM, NLSVM, SSVM and NSVM, in: Proceedings of 2004 International Conference on Machine Learning and Cybernetics, vol. 7, 2004, pp. 4277–4282.

[14] A.J. Smola, B. Scholkopf, A tutorial on support vector regression, Stat. Comput. 14 (2004) 199–222.

[15] U. Thissen, R. van Brakel, A.P. de Weijer, W.J. Melssen, L.M.C. Buydens, Using support vector machines for time series prediction, Chemometr. Intell. Lab. Syst. 69 (1–2) (2003) 35–49.

[16] S.S. Keerthi, C.J. Lin, Asymptotic behaviors of support vector machines with Gaussian kernel, Neural Comput. 15 (7) (2003) 1667–1689.

[17] H.T. Lin, C.J. Lin, A Study on Sigmoid Kernels for SVM and the Training of Non-PSD Kernels by SMO-type Methods, Department of Computer Science and Information Engineering, National Taiwan University, Taipei, 2003.
[18] P.-F. Pai, W.-C. Hong, Support vector machines with simulated annealing algorithms in electricity load forecasting, Energy Convers. Manage. 46 (17) (2005) 2669–2688.
[19] K. Duan, S.S. Keerthi, A.N. Poo, Evaluation of simple performance measures for tuning SVM hyperparameters, Neurocomputing 51 (2003) 41–59.
[20] M.M.S. Lee, S. Sathiya Keerthi, C. JinOng, D. DeCoste, An efficient method for computing leave-one-out error in support vector machines with Gaussian kernels, IEEE Trans. Neural Networks 15 (3) (2004) 7500–7757.
[21] V. Cherkassky, Y. Ma, Practical selection of SVM parameters and noise estimation for SVM regression, Neural Networks 17 (2004) 113–126.
[22] J. Kennedy, R. Eberhart, Particle swarm optimization, in: Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, Piscataway, 1995, 1942–1948.
[23] E. Elbeltagi, T. Hegazy, D. Grierson, Comparison among five evolutionary-based optimization algorithms, Adv. Eng. Inform. 19 (1) (2005) 43–53.
[24] Y.H. Shi, R.C. Eberhart, Parameter selection in particle swarm optimization, The Seventh Annual Conference on Evolutionary Programming, San Diego, USA.
[25] Y.H. Shi, R.C. Eberhart, A modified particles swarm optimizer, in: Proceedings of the IEEE International Conference on Evolutionary Computation, Piscataway, 1998, pp. 69–73.
[26] G. Migliavacca, M. Benini, M. Gallanti, E. Bompard, E. Carpaneto, G. Ciwei, LREMS: a long term electricity market simulator based on game theory, Proceedings of the 2006 Power Systems Conference and Expo, October 29–November 1, 2006, Atlanta, Georgia, USA.