

# Application of Seasonal SVR Model with Genetic Algorithm and Tabu Search on Prediction

Zhang Zheng\* and Yang Shanlin

*School of management, Hefei University of Technology, Hefei 230009, Anhui province, P. R. China*

*\*E-mail: zhzheng99@sina.com*

## Abstract

*The monthly electricity consumption data is fluctuant over time, but it also has regular pattern of variation. In order to predict the electricity consumption, a seasonal support vector regression model based on genetic algorithm and tabu search (GATS) is proposed. In this model, a hybrid strategy of genetic algorithm and tabu search is applied to parameter optimization. The hybrid strategy can overcome the drawbacks of genetic algorithm, such as the deficiency of climbing ability. The least absolute criterion is used in the model instead of least square criterion because it is more stable. In addition, a seasonal adjustment method is adopted. It can improve the accuracy of prediction. At last, the seasonal GATS-SVR model is used to predict the electricity consumption. It also compares with GATS-SVR, original SVR and ARIMA (1,1,1) models and the result shows the SVR model is better than these models in prediction accuracy.*

**Keywords:** Prediction; Support vector regression; GATS; Seasonal adjustment

## 1. Introduction

With the rapid economic development in China, electricity demand increases with incredible speed. In the production and living of society, electricity rationing still happens and it does significant influence to enterprise's production activities. The electricity is an indispensable part of people's daily life because people need light, computer, household appliances and so on. They can do nothing without electricity. So the electricity is indispensable. But if produce electricity too much, it will cause a great waste of resources and serious environmental pollution since power storage is extremely difficult. On the contrary, if produce electricity too less, the quality of people's life will not be able to obtain the safeguard. It will also lead to enterprises shutdown and affect the enterprises benefit. Thus, precise prediction of electricity consumption will make production, supply and sales of power system achieve dynamic balance. It will also guarantee economic development and increase the economic benefit and social benefit.

In past decades, many methods have been proposed in electricity prediction. The traditional well known methods are as follows: Box and Jenkins [1] proposed autoregressive integrated moving average (ARIMA) model in 1970. Vähäkylä *et al.* [2], Erkan [3] and Chen *et al.* [4] applied ARIMA model to the load forecasting. Christiaanse [5] used exponential smoothing models with Fourier series transformation in electricity load forecasting and Taylor [6] proposed a double seasonal exponential smoothing model, it improved the accuracy of electricity demand forecasting. Papalexopoulos and Hesterberg [7] as well as Charytoniuk [8] used the regression methods in load forecasting. With the development of computer technology, some researches such as Park *et al.* [9] and Shimakura [10] applied artificial neural network (ANN) in load forecasting since

---

\* Corresponding Author

ANN has outstanding nonlinear mapping abilities. In addition, grey models which proposed by Deng [11] are also applied to electricity prediction by lots of researchers [12-13]. These methods all have their disadvantages and pretty of researchers tried to modify the forecasting models to improve the prediction accuracy [14-16].

Support vector machine (SVM) was proposed by Vapnik [17] in 1995. It is based on structural risk minimization principle and can better solve over learning problem. Support vector machine is composed of support vector classifier (SVC) and support vector regression (SVR). Thereinto, support vector regression is an important research direction of Support vector machine. Since SVM theory was put forward, lots of scholars and researchers have used SVR model into their fields to solve practical problems. Certainly, there are also many people who focus on electricity prediction by SVR. For example, Kavaklioglu [18] proposed a SVR model to predict Turkey's electricity consumption. Pan *et al.* [19] used radial basis function as the kernel function of SVR to predict short load and the results showed that it was better than BP-neural network, *etc.*

Although lots of people find that the support vector regression models are better than the other models, such as linear regression model, ARIMA, ANN, *etc.*, due to the strong theoretical basis and generalization ability, the SVR models also have shortcomings and problems. On the one hand, the parameters of SVR model have a great influence on prediction accuracy. For example, the parameter  $C$  is the trade-off between model flatness and empirical risk and the parameter  $\varepsilon$  is the width of loss function. These parameters have a serious impact on prediction accuracy. In the early research, people determined parameter values according to the experience. Cherkassky and Mulier [20] determined the parameter values through the number of training samples and the noise level of training samples. It is not very accurate when people lack of experience. So in recent years, lots of researchers and scholars apply intelligent optimization method to determine the parameter values of SVR models. Hong[21] used immune algorithm to optimize the parameters of SVR model, then used the model to predict Taiwan electricity consumption; Niu and Wang *et al.* [22] used ant colony algorithm to optimize the training data and speeded up the SVM training time. And Zong *et al.* [23] proposed a modified particle swarm optimization (PSO) algorithm to search the parameters of SVR model. In this algorithm, certain particles at each iterative was added in order to broaden search area and avoid falling into local optimization. Zhao [24] used rough set theory to reduce attributes and established the least squares support vector machine to predict financial distress by parameters optimization of genetic algorithm (GA). On the other hand, the model performance is also affected by the kernel function. The kernel function plays a role of setting the input data space into a high dimensional feature space so that the problem will become a high dimensional space regression problem. So some researches are focus on the performance of different kernel functions. Shawe [25] pointed that a new kernel function could be constructed by the simple operation of the other kernel functions. Smits [26] mixed radial basis function kernel and polynomial kernel in SVR model. The mixtures of kernels could result in having both good interpolation and extrapolation abilities. Üstün *et al.* [27] applied a universal kernel function based on the Pearson VII function (PUK) into SVR model, and from the result, it was concluded that the PUK kernel is robust and has an equal or even stronger mapping power than the common kernel functions in the examinations.

The paper is organized as follows: in Section 2, we first introduce original SVR model. Then a hybrid algorithm combining genetic algorithm and tabu search is used to parameter optimization. A seasonal adjustment method is also applied to enhance the prediction accuracy. In Section 3, we construct the model proposed in this paper and apply it to electricity consumption prediction. Conclusion and further directions of study are underlined in Section 4.

## 2. The Process of Optimized SVR Model

### 2.1. Support Vector Regression (SVR)

Suppose a set of data  $S = (x_i, y_i)$ ,  $i = 1, 2, \dots, n$ ,  $x_i \in R^p$ ,  $y_i \in R$ ,  $x_i$  is input data of  $p$  dimensional vector.  $y_i$  is output data. By using a nonlinear map from input space to output space, the data is mapped to a higher dimensional feature space. And let the SVR function is

$$f(x) = w\phi(x) + b \quad (1)$$

Where  $\phi(\cdot)$  is the nonlinear map,  $w$  is weighed vector,  $b$  is threshold value. Through the structural risk minimization principle, the following problem needs to be solved in order to get  $f(x)$ :

$$R(f) = \frac{1}{2} \|w\|^2 + \frac{C}{N} \sum_{i=1}^n L(y_i, f(x_i)) \quad (2)$$

Where  $C$  is positive constant, it's the trade-off between model flatness and empirical risk.  $L(y_i, f(x_i))$  is loss function. Generally, it is  $\varepsilon$ -insensitive loss function which is shown as follows:

$$L(y_i, f(x_i)) = L_\varepsilon(y_i, f(x_i)) = \begin{cases} 0, & |y_i - f(x_i)| \leq \varepsilon \\ |y_i - f(x_i)| - \varepsilon, & |y_i - f(x_i)| > \varepsilon \end{cases} \quad (3)$$

For  $i = 1, 2, \dots, n$ .

Then, the regression solving problem turns to the optimization problem:

$$\min R(w, \xi_i, \xi_i^*) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (4)$$

$$s.t. \begin{cases} y_i - f(x_i) \leq \varepsilon + \xi_i \\ f(x_i) - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, n \end{cases}$$

In order to solve the optimization problem (4), Lagrange function is defined like this:

$$L(w, b, \xi_i, \xi_i^*, \alpha_i, \beta_i, \mu_i, \nu_i) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (5)$$

$$+ \sum_{i=1}^n \alpha_i (y_i - w\phi(x_i) - b - \varepsilon - \xi_i) + \sum_{i=1}^n \beta_i (w\phi(x_i) + b - y_i - \varepsilon - \xi_i^*) - \sum_{i=1}^n \mu_i \xi_i - \sum_{i=1}^n \nu_i \xi_i^*$$

Then the following equations are obtained in order to solve the optimization problem:

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^n \alpha_i \phi(x_i) + \sum_{i=1}^n \beta_i \phi(x_i) = w - \sum_{i=1}^n (\alpha_i - \beta_i) \phi(x_i) = 0 \quad (6)$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^n \beta_i - \sum_{i=1}^n \alpha_i = 0 \quad (7)$$

$$\frac{\partial L}{\partial \xi_i} = C - \alpha_i - \mu_i = 0 \quad (8)$$

$$\frac{\partial L}{\partial \xi_i^*} = C - \beta_i - \nu_i = 0 \quad (9)$$

Substituting (6)~(9) into (5) and assuming Kernel function  $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ , then the optimization problem is transformed into its dual problem, that is

$$\max J = \sum_{i=1}^n y_i (\alpha_i - \beta_i) - \varepsilon \sum_{i=1}^n (\alpha_i + \beta_i) - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \beta_i) (\alpha_j - \beta_j) \phi(x_i) \cdot \phi(x_j) \quad (10)$$

$$s.t. \begin{cases} \sum_{i=1}^n \beta_i - \sum_{i=1}^n \alpha_i = 0 \\ \alpha_i, \beta_i \in [0, C], i = 1, 2, \dots, n \end{cases}$$

By solving the quadratic programming problem (10), we have

$$f(x) = \sum_{i=1}^n (\alpha_i - \beta_i) \phi(x_i) \cdot \phi(x) + b = \sum_{i=1}^n (\alpha_i - \beta_i) K(x_i, x) + b \quad (11)$$

$$w = \sum_{i=1}^n (\alpha_i - \beta_i) \phi(x_i) \quad (12)$$

Then the function  $f(x)$  can be obtained by using dot product operation from high dimensional feature space without getting the form of  $\phi(x)$ . And the kernel function which meets the Mercer's conditions is corresponding to a dot product operation of high dimensional space [17]. Generally, kernel function has a variety of forms, such as polynomial function, radial basis function, sigmoid function, *etc.* The radial basis function has fewer parameters and the model will have better approximation performance for higher frequency of nonlinear system when applied RBF kernel function. In addition, changing parameter of RBF kernel function in the valid range won't add space complexity [28]. Due to these reasons, the radial basis function is adopted and the form is

$$K(x, x_i) = \exp(-|x - x_i|^2 / \sigma^2) \quad (13)$$

Then, the SVR model is constructed to forecast electricity consumption.

## 2.2. Seasonal Adjustment

It is obvious that the electricity consumption has a rising trend with the time goes on because of the rapid development of the economy. The electricity consumption also shows a cyclic trend from month to month (Figure1), and numerous studies have shown the electricity consumption change rule [6,29-31]. It will be feasible to make use of the periodical change of electricity consumption since the electricity consumption has had such regularity. There are many methods on making use of cyclic trend of calculating seasonal index [32-34]. Some researchers applied simple weighted average method to calculate seasonal index. This approach is widely used because it is easy and convenient. And in this paper, squared average method will be applied to solve the seasonal index. This approach is more sensitive to the data variation. It is shown as follows:

$$S_i = \frac{y_i}{\hat{y}_i} = \frac{y_i}{\sum_{i=1}^n (\alpha_i - \beta_i) K(x_i, x) + b} \quad (14)$$

Where  $y_i$  is true value,  $\hat{y}_i$  is predictive value,  $i = j, j+d, \dots, j+(m-1)d$ ,  $0 < j \leq d$ ,  $d$  is the step length of time series,  $m$  is time length. Then the seasonal index is

$$SI_j^2 = \frac{1}{m} \sum_{i=1}^m S_{j+(i-1)d}^2 \quad (15)$$

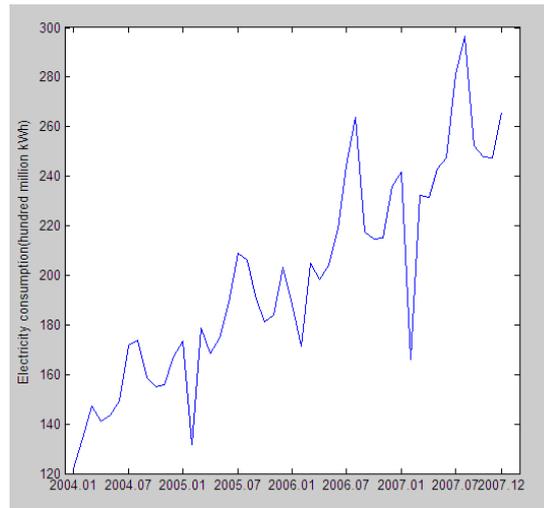
$$0 < j \leq d, \quad j \in N$$

$$SI_j = \sqrt{\frac{1}{m} \sum_{i=1}^m S_{j+(i-1)d}^2} \quad (16)$$

And the predictive value will be adjusted by Eq.(17), that is

$$\tilde{\hat{y}}_{T+l} = \left( \sum_{i=1}^n (\alpha_i - \beta_i) K(x_i, x_{T+l}) + b \right) * SI_l \quad (17)$$

Where  $l = j+(i-1)d$ ,  $i = 1, 2, \dots, 0 < j \leq d$ ,  $T$  is the latest end point of the year. The electricity consumption data are of volatility and randomness, but they also shows a cyclic trend. Thus, making use of the characteristic by seasonal method could enhance the prediction accuracy.



**Figure 1. Electricity Consumption from 2004 to 2007 of Jiangsu Province**

### 2.3. Parameters Optimization of Genetic Algorithm (GA) and Tabu Search (TS)

Through SVR algorithm process, the prediction results of SVR model depend largely on the variable  $\varepsilon$  in insensitive loss function, the positive constant  $C$  and the variable  $\sigma^2$  in radial basis function. So the selection of parameters of SVR model is certainly important.

**2.3.1. Least Absolute Criterion:** Through the least square criterion is widely used in regression analysis due to its convenience in differentiation, and hence in optimization, it is also preponderant in estimation, for example in the method of least squares, and for the same reason the square-error loss function is widely used in statistical decision theory [35]. However, the least square criterion has several drawbacks, for example, it will lead the model error to be bigger when the singular points exist.

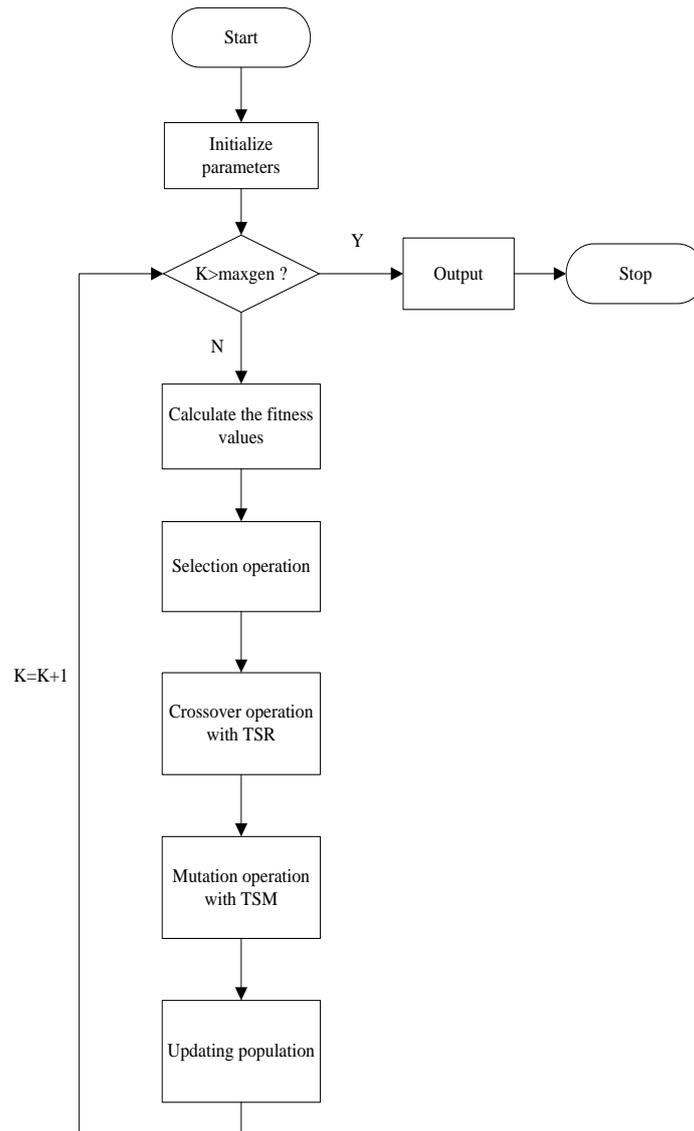
On the other hand, the least absolute criterion is also very important in regression analysis. But the research and application on least absolute criterion is fewer than least square absolute criterion. In least absolute criterion, it uses the first power of error instead of the quadratic error. Thus the singular points have less influence on the model and error will not be expanded. So the least absolute method has better stability than least square method, and also has better performance on statistics. The least absolute criterion uses absolute error to describe the deviation, that is

$$\begin{aligned} & \min \sum_k |y_k - \hat{y}_k| \\ & = \min \sum_k \left| y_k - \sum_{i=1}^n (\alpha_i - \beta_i) K(x_i, x_k) - b \right| \end{aligned} \quad (18)$$

Because Eq. (18) is non-smooth and non-analytical, it is difficult to solve the parameters. Thus intelligent optimization algorithms are considered.

**2.3.2. Hybrid Strategy of Genetic Algorithm (GA) and Tabu Search (TS):** Genetic algorithm is especially useful for complex optimization problems where the number of parameters is large and the analytical solutions are difficult to obtain [36]. But the genetic algorithm also has some drawbacks with large scale optimization problem. One reason is that this algorithm needs to maintain a large population in every generation. It will cost a lot of time and space with the access and storage of population. Another drawback is

premature phenomena. The reason of premature phenomena is that the crossover operator makes the population have partial similarity between chromosomes. It may lead the search procedure stop early. Another reason is that the probability of mutation operator is generally low. And this leads the diversity of population to be not enough and it is also the cause of deficiency of climbing ability. Compared with the traditional optimization algorithms, the tabu search algorithm has its advantages. Firstly, tabu search algorithm can accept bad solutions and it has better ability in climbing. Secondly, it has a good balance between local search and global search. Hence, the tabu search algorithm is used into the genetic algorithm to change the crossover operator and mutation operator. And the flow chart of GATS is shown in Figure 2.



**Figure 2. GATS Algorithm Flow Chart**

The hybrid algorithm procedure is as follows:

Step 1: Set initial population size  $N=100$ . Then generating randomly  $N$  chromosomes. The coding method is binary coding method.

Step 2: Judge whether the algorithm reaches termination conditions. If doesn't, turn to Step 3. Otherwise, turn to Step 8. Here, the termination conditions are maximum iteration

number and minimum error bound. And maximum iteration number is 1000, minimum error bound is  $10^{-8}$ .

Step 3: The fitness function is

$$F = \sum_k \left| y_k - \sum_{i=1}^n (\alpha_i - \beta_i) K(x_i, x_k) - b \right| \quad (19)$$

Decoding every chromosome in the population and calculate the fitness values.

Step 4: Do selection operation and copy operation according to each individual fitness value. Proportional strategy is used for selection strategy. That is, for individual  $i$ , population size  $N$ , set the fitness value is  $F_i$ . Thus, the selective probability of individual  $i$  is

$$P_i = (1/F_i) / \sum_{j=1}^N (1/F_j) \quad (20)$$

After selective probability is got, roulette wheel method is used to accomplish selective operation. That is, let  $A_0 = 0$ ,  $A_i = \sum_{j=1}^i P_j$ , for a random number  $\xi$ , if  $A_{i-1} \leq \xi \leq A_i$ , then individual  $i$  is to be selected.

Step 5: The crossover operator is TSR. And crossover probability is

$$p_c = \begin{cases} p_{c \max} - \frac{p_{c \max} - p_{c \min}}{\max gen} * curgen, & F' < F_{avg} \\ p_{c \max} & , F' \geq F_{avg} \end{cases} \quad (21)$$

Where  $p_c$  is crossover probability,  $p_{c \max}$  is maximum crossover probability,  $p_{c \min}$  is minimum crossover probability,  $F'$  is the smaller fitness value of two individuals,  $\max gen$  is maximum iteration number,  $curgen$  is current iteration number. In order to make the algorithm have a good ability in global search and avoid to fall into local optimal solution, here set  $p_{c \max} = 0.9$ ,  $p_{c \min} = 0.1$ .

The crossover operation procedure is shown below:

Step 5.1: For each chromosome, generating a random number  $RN$  between 0 and 1. If  $RN \leq p_c$ , the chromosome is chosen as parent chromosome.

Step 5.2: Multi-point crossover operation is used to generate offspring from parent chromosomes.

Step 5.3: Set the average fitness value of parent chromosomes as the level of aspiration value. And the fitness value is regarded as tabu object. If the fitness value of offspring is better than the level of aspiration value, it is accepted. Otherwise, turn to Step 5.4.

Step 5.4: If the fitness value of offspring is not in the tabu list, it is accepted. Otherwise the better one of two parents is chosen to the next generation.

Step 6: The mutation operator is TSM. And mutation probability is

$$p_m = \begin{cases} p_{m \max} - \frac{p_{m \max} - p_{m \min}}{\max gen} * curgen, & F < F_{avg} \\ p_{m \min} & , F \geq F_{avg} \end{cases} \quad (22)$$

Where  $p_m$  is mutation probability,  $p_{m \max}$  is maximum mutation probability,  $p_{m \min}$  is minimum mutation probability,  $F$  is fitness value of individual. In order to make the chromosomes be diverse in the beginning iteration and improve the local search ability later period, here set  $p_{m \max} = 0.4$ ,  $p_{m \min} = 0.01$ .

The mutation operation procedure is shown below:

Step 6.1: For each chromosome, generating a random number  $RN$  between 0 and 1. If  $RN \leq p_m$ , the chromosome is mutated by TSM operator.

Step 6.2: Set the current chromosome as the initial solution and initialize the tabu search algorithm.

Step 6.3: Judge whether the iteration of tabu search algorithm reaches the stopping criterion. If it does, then stop tabu search and turn to Step 7. Otherwise turn to Step 6.4.

Step 6.4: Judge whether the best candidate solution reaches the level of aspiration. If it does, updating the level of aspiration and current solution. Otherwise, turn to Step 6.5.

Step 6.5: The best candidate solution which is not in the tabu list is chosen as the current solution.

Step 6.6: Update the tabu list. Then turn to Step 6.3.

Step 7: Set  $curgen = curgen + 1$ , turn to Step 2.

Step 8: Output the best individual of the population. That is the optimal parameters  $C$ ,  $\sigma^2$  and  $\varepsilon$ .

From the above steps, it can be inferred that the GATS algorithm reserves their advantages and the parameters of SVR model could be found by this algorithm.

### 3. Model Construction and Prediction

#### 3.1. Data Selection and Preprocessing

In this paper, the data are monthly whole society electricity consumption of Jiangsu province from 2004 to 2011. There are 96 data points. The previous 84 data points from 2004 to 2010 are chosen to be training data. They are used to construct the best seasonal GATS-SVR model. Then the established model is applied to predict the electricity consumption of Jiangsu province from Jan. 2011 to Dec. 2011 which is the last 12 data points.

In order to compare the prediction performance, the ARIMA model, the original SVR model and the GATS-SVR model without seasonal adjustment are adopted to comparative analysis.

#### 3.2. Modeling and Prediction

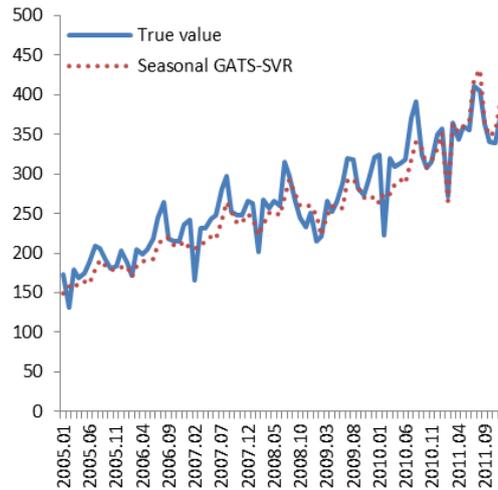
The rolling-base forecasting procedure is used to construct seasonal GATS-SVR model in this paper. It can be described like this: For  $x_i \in R$ ,  $i = 1, 2, \dots, n$ , set the seasonal training length to be  $l$ , then

$$\begin{aligned}\hat{x}_{13} &= f(x_1, x_2, \dots, x_l) \\ \hat{x}_{14} &= f(x_2, x_3, \dots, x_{l+1}) \\ &\vdots \\ \hat{x}_n &= f(x_{n-l}, x_{n-(l-1)}, \dots, x_{n-1})\end{aligned}$$

And the error between  $\hat{x}_i$  and the true value  $x_i$  is calculated based on Least absolute criterion.

Since the electricity consumption data has their special regular patterns in each month, the seasonal length is set to be 12. Then the seasonal index of each month can be calculated.

The regression and prediction curve of seasonal GATS-SVR model proposed in this paper is shown in Figure3. The unit of electricity consumption is hundred million kWh.



**Figure 3. The Regression and Prediction Curve of Seasonal GATS-SVR**

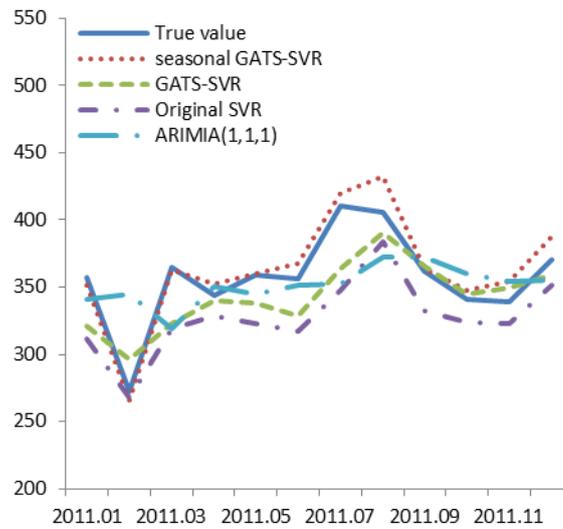
**Table 1. The Predictive Value of Different Models**

Time	True value	seasonal GATS-SVR	GATS-SVR	Original SVR	ARIMIA(1,1,1)
2011.01	357.54	351.5873116	320.9728333	311.5069186	340.74573
2011.02	272.46	265.840813	296.3520904	267.8071804	345.17376
2011.03	364.47	362.4346321	322.4722574	319.3547584	318.65833
2011.04	343.51	352.4800163	340.2920092	328.3776743	350.25519
2011.05	359.26	360.0730861	338.3917748	322.5943185	345.05319
2011.06	355.89	367.2788593	328.1192002	317.0632188	351.84961
2011.07	410.82	419.7093528	364.1228376	347.3564397	352.40514
2011.08	405.29	432.3987356	390.063396	384.1585385	372.23865
2011.09	362.22	364.2338671	365.6146422	332.764815	372.41785
2011.10	341.18	347.3572899	344.9717101	323.8180578	360.09873
2011.11	338.91	354.0893753	349.1270058	322.5618416	354.74567
2011.12	370.07	388.0616959	359.0572162	351.5404146	355.44098

Table 1 shows the predictive value of different models, such as seasonal GATS-SVR model, GATS-SVR model, original SVR model and ARIMA (1,1,1) model. The seasonal index of different months is shown in Table 2. Figure 4 gives graphic depictions of these models in predictive months. Prediction performance of different models is shown in Table 3.

**Table 2. Seasonal Index of Different Months**

Time	Seasonal index
January	1.095380279
February	0.897043826
March	1.123925001
April	1.035816319
May	1.064071626
June	1.119345833
July	1.15265869
August	1.10853451
September	0.996223414
October	1.006915291
November	1.014213651
December	1.080779548



**Figure 4. Prediction Curve of Different Models**

**Table 3. Prediction Performance of Different Models**

	seasonal GATS-SVR	GATS-SVR	Original SVR	ARIMIA(1, 1, 1)
MAPE	0.0260418	0.0568591	0.08016267	0.0753066
RMSE	11.917958	25.088011	33.5543982	33.370499
$C$	18.1013	6.9644	--	--
$\sigma^2$	69.9051	17.1484	--	--
$\varepsilon$	0.02621	0.01	--	--

$$\text{Where } RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|$$

**Table 4. Relative Error of Different Models**

Time	seasonal GATS-SVR	GATS-SVR	Original SVR	ARIMIA(1, 1, 1)
2011.01	-0.01665	-0.10227	-0.1287495	-0.0469717
2011.02	-0.02429	0.08769	-0.0170771	0.26687866
2011.03	-0.00558	-0.11523	-0.1237831	-0.1256939
2011.04	0.026113	-0.00937	-0.0440521	0.01963608
2011.05	0.002263	-0.05809	-0.1020589	-0.0395446
2011.06	0.032001	-0.07803	-0.1090977	-0.0113529
2011.07	0.021638	-0.11367	-0.1544802	-0.1421909
2011.08	0.066887	-0.03757	-0.0521391	-0.0815499
2011.09	0.00556	0.009372	-0.0813185	0.02815375
2011.10	0.018106	0.011114	-0.0508879	0.05545088
2011.11	0.044789	0.030147	-0.0482375	0.0467253
2011.12	0.048617	-0.02976	-0.0500705	-0.0395304

From Table 3, the original SVR model has the worst prediction performance of all comparative models according to MAPE and RMSE. The GATS-SVR model is better than both original SVR model and ARIMA (1,1,1) model. It indicates that the hydride algorithm GATS has a good performance in parameter optimization. And the seasonal

GATS-SVR model is better than the other models in electricity consumption prediction. So it means that the seasonal adjustment method can help to improve the prediction accuracy of SVR model.

Table 4 gives the relative error of seasonal GATS-SVR model and comparative models. Even through the relative error of seasonal GATS-SVR is not the best in all months when compared with other models, it still has the best performance when considered the twelve months as a whole. In that way, the seasonal GATS-SVR model has the best prediction accuracy of all the comparative models.

#### 4. Conclusion

Electricity consumption prediction is very important for decision process because the electricity is hard to store and it will cause problems whenever the electricity is produced too much or too less. In this paper, a seasonal GATS SVR model is proposed to predict the electricity consumption of Jiangsu province. Since the parameters of SVR determine the framework of model and have a great influence on the prediction accuracy, a hybrid strategy of genetic algorithm and tabu search is applied to the parameter optimization of SVR. The hybrid algorithm combines advantages of both algorithms and overcome the disadvantages of the genetic algorithm. The least absolute criterion is also used in the parameter optimization of SVR model. It is more stable than least square criterion. According to the monthly electricity consumption curve, the electricity consumption shows a special regular pattern and the same month in every year also has internal relations. Thus a seasonal adjustment method is used in the SVR model to improve the prediction accuracy. At the end, the result shows that the seasonal GATS-SVR model proposed in this paper has better performance in electricity consumption prediction than GATS-SVR model, original SVR model and ARIMA(1,1,1) model.

There is still a lot of work need to be done in the future, such as to use different loss functions and kernel functions, the better way to search the parameters and so on. The work will be finished in a follow-up research.

#### Acknowledgements

This research is conducted with the support of national “863” key project of smart grid (2011AA05A116).

#### References

- [1] Box G. E. P. and Jenkins G. M., “Time series analysis, forecasting and control”, San Francisco: Holden-Day, (1970).
- [2] Vähäkylä P., Hakonen E. and Léman P., “Short-term forecasting of grid load using Box-Jenkins techniques”, International Journal of Electrical Power & Energy Systems, vol. 2, no. 1, (1980), pp. 29-34.
- [3] Erdogdu E., “Electricity demand analysis using co-integration and ARIMA modeling: A case study of Turkey”, Energy Policy, vol. 35, no. 2, (2007), pp. 1129-1146.
- [4] Chen J. F., Wang W. M. and Huang C. M., “Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting”, Electric Power Systems Research, vol. 34, no. 3, (1995), pp. 187-196.
- [5] Christiaanse W. R., “Short-Term Load Forecasting Using General Exponential Smoothing”, IEEE Transactions on Power Apparatus and Systems, , PAS-90, no. 2, (1971), pp. 900-911.
- [6] Taylor J. W., “Short-term electricity demand forecasting using double seasonal exponential smoothing”, Journal of the Operational Research Society, vol. 54, (2003), pp. 799–805.
- [7] Papalexopoulos A. D. and Hesterberg T. C., “A regression based approach to short-term system load forecasting”, IEEE Transactions on Power Systems, vol. 5, no. 4, (1990), pp. 1535-1547.
- [8] Charytoniuk W., Chen M. S. and Van O. P., “Nonparametric regression based short-term load forecasting”, IEEE Transactions on Power Systems, vol. 13, no. 3, (1998), pp. 725-730.

- [9] Park D. C., El-Sharkawi M. A. and Marks R. J. II., "Electric load forecasting using an artificial neural network", *IEEE Transactions on Power Systems*, vol. 6, no. 2, (1991), pp. 442-449.
- [10] Shimakura Y., Fujisawa Y. and Maeda Y., "Short-term load forecasting using an artificial neural network", *Proceedings of the Second International Forum on Applications of Neural Networks to Power Systems*, (1993), pp. 233-238.
- [11] Deng J. L., "Control problems of grey systems", *Systems & Control Letters*, vol. 1, no. 5, (1982), pp. 288-294.
- [12] Wang C. S., Yang J. and Zhang C. J., "Application of Grey System Theory in City Electricity Demand Forecasting", *Power System Technology*, vol. 23, no. 2, (1999), pp. 15-18.
- [13] Zhang J. F., Wu Y. A. and Wu J. J., "Application of gray system theory in load forecasting", *Electric Power Automation Equipment*, vol. 24, no. 5, (2004), pp. 24-27.
- [14] Amjady N., "Short-term hourly load forecasting using time-series modeling with peak load estimation capability", *IEEE Transactions on Power Systems*, vol. 16, no. 3, (2001), pp. 498-505.
- [15] Srinivasan D., "Evolving artificial neural networks for short term load forecasting", *Neurocomputing*, vol. 23, no. 1-3, (1998), pp. 265-276.
- [16] Hsu C. C. and Chen C. Y., "Applications of improved grey prediction model for power demand forecasting", *Energy Conversion and Management*, vol. 44, no. 14, (2003), pp. 2241-2249.
- [17] V. N. Vapnik, "The Nature of Statistic Learning Theory", New York: Springer, (1995).
- [18] Kavaklioglu K., "Modeling and prediction of Turkey's electricity consumption using Support Vector Regression", *Applied Energy*, vol. 88, no. 1, (2011), pp. 368-375.
- [19] Pan F., Cheng H. Z. and Yang J. F., "Power System Short-Term Load Forecasting Based on support vector machines", *Power System Technology*, vol. 28, no. 21, (2004), pp. 39-42.
- [20] Cherkassky V. and Mulier F., "Learning from Data: Concepts, Theory and Methods", NY: John Wiley & Sons, (1997).
- [21] Hong W. C., "Electric load forecasting by support vector model", *Applied Mathematical Modeling*, vol. 33, no. 5, (2009), pp. 2444-2454.
- [22] Niu D. X., Wang Y. L. and Wu D. S. D., "Power load forecasting using support vector machine and ant colony optimization", *Expert Systems with Applications*, vol. 37, no. 3, (2010), pp. 2531-2539.
- [23] Zong Q., Liu W. J. and Dou L. Q., "Parameters selection for SVR based on PSO", *The Sixth World Congress on Intelligent Control and Automation*, vol. 1, (2006), pp. 2811-2814.
- [24] Zhao G. H., "Study on Financial Distress Prediction Model of Least Squares Support Vector Machine of Dual Constraint Type Based on Attribute Reduction of Neighborhood Rough Set", *Operations Research and Management Science*, vol. 20, no. 3, (2011), pp. 132-139.
- [25] Shawe T. J. and Cristianini N., "Kernel methods for pattern analysis", London, Cambridge University Press, (2004).
- [26] Smits G. F. and Jordaan E. M., "Improved SVM regression using mixtures of kernels", *Proceedings of the International Joint Conference on Neural Networks*, vol. 3, (2002), pp. 2785-2790.
- [27] Üstün B., Melssen W. J. and Buydens L. M. C., "Facilitating the application of Support Vector Regression by using a universal Pearson VII function based kernel", *Chemometrics and Intelligent Laboratory Systems*, vol. 81, no. 1, (2006), pp. 29-40.
- [28] Rong H. N., Zhang G. X. and Jin W. D., "Selection of Kernel Functions and Parameters for Support Vector Machines in System Identification", *Journal of System Simulation*, vol. 18, no. 11, (2006), pp. 3204-3208.
- [29] Ahmed S., "Seasonal models of peak electric load demand", *Technological Forecasting and Social Change*, vol. 72, no. 5, (2005), pp. 609-622.
- [30] Wang Y. Y., Wang J. Z., Zhao G. and Dong Y., "Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: A case study of China", *Energy Policy*, vol. 48, (2012), pp. 284-294.
- [31] Taylor J. W., "Triple seasonal methods for short-term electricity demand forecasting", *European Journal of Operational Research*, vol. 204, no. 1, (2010), pp. 139-152.
- [32] Deo R., Hurvich C. and Lu Y., "Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment", *Journal of Econometrics*, vol. 131, no. 1-2, (2006), pp. 29-58.

- [33] Wang J. Z., Zhu W. J., Zhang W. Y. and Sun D. H., “A trend fixed on firstly and seasonal adjustment model combined with the  $\varepsilon$ -SVR for short-term forecasting of electricity demand”, *Energy Policy*, vol. 37, no. 11, (2009), pp. 4901-4909.
- [34] Martens M., Chang Y.C. and Taylor S. J., “A comparison of seasonal adjustment methods when forecasting intraday volatility”, *Journal of Financial Research*, vol. 25, no. 2, (2002), pp. 283-299.
- [35] Pham-G. T. and Hung T. L., “The mean and median absolute deviations”, *Mathematical and Computer Modeling*, vol. 34, no. 7-8, (2001), pp. 921-936.
- [36] Lam H. K., Ling S. H., Leung F. H. F. and Tam P. K. S., “Tuning of the structure and parameters of neural network using an improved genetic algorithm”, *IEEE Transactions on Neural Networks*, vol. 14, no.1, (2003), pp.79-88.

