Application of LS-SVM in the Short-term Power Load Forecasting Based on QPSO

Liao Xiaohui
School of Electrical Engineering, Zhengzhou University
Zhengzhou, China
E-mail:liaoxiaohui0001@zzu.edu.cn

Ding Qian
School of Electrical Engineering, Zhengzhou University
Zhengzhou, China
E-mail:D.Q.OK@163.com

***Abstract*—Electricity power load forecasting is the basis of power system planning and construction. In order to improve the precision, this paper proposes a model, in which the parameters in least square support vector machine (LS-SVM) are optimized by Quantum-behaved Particle Swarm Optimization (QPSO) and considering the weather and date factors. In the quantum space, particles can be search in the whole feasible solution space. We can obtain the global optimal solution. Therefore, QPSO algorithm is a global guaranteed algorithm, which is better than the original PSO algorithm in search capability. The simulation results show that the adaptive particle swarm optimization-based SVM load forecasting model is more accurate than the neural networks model and traditional LS-SVM model.**

Keywords-short-term load forecasting;least square support vector machine(LS-SVM);quantum-behaved particle swarm optimization(QPSO);weather factor; date factor

I. INTRODUCTION

With the increasing marketization of electric power production and consumption, the accuracy, instantaneity, reliability and intelligent of load forecasting was put forward higher requirements. Therefore, load forecasting has become a very important research part of the modern power system operation and management[1].

Support vector machine (SVM)[2] is a kind of machine learning algorithms based on statistical learning theory. It achieves actual risk minimization by seeking structure risk minimization, thus under the condition of the small sample size, it also can get a very good study effect, with a stronger generalization performance. On the basis of SVM, Suykens et al proposed least squares support vector machines (LS-SVM)[3]. Instead of inequality constraints in the standard SVM with equality constraints, LS-SVM reduce some uncertain factors and change the convex quadratic programming problem into the linear equations to solve, thus it reduces the computational complexity and improves the solving speed. When using the LS-SVM model for load forecasting, the choice of parameters is the key to affect the prediction precision.

In particle swarm optimization (PSO)[4], the searching space of particles is a limited and gradually reducing area, which fails to cover the whole feasible solution space and cannot guarantee the global convergence. For this shortcoming, using some theories of quantum mechanics as a background and according to the basic convergence properties of the particle swarm, Sun proposed Quantum-behaved Particle Swarm Optimization (QPSO)[5][6]. This algorithm has been greatly improved in the search speed and accuracy.

In this paper, it achieves the optimal selection of model parameters by using QPSO to optimize LS-SVM parameters and applies it to the actual load forecast of Hefei. The instance simulation shows that this method has a higher precision and faster training speed.

II. THE PRINCIPLE OF LS-SVM

A sample sets $\{(x_i, y_i)\}(i=1,2,\ldots,n)$, x_i is the input vector, y_i is the corresponding target output data. The estimate function is defined as:

$$f(x)=\omega\varphi(x)+b \tag{1}$$

Where $\varphi(x)$ is a nonlinear mapping from the input space to the high dimensional feature space; ω is the weight vector; b is the bias.

According to the principle of structure minimization, the LS-SVM optimization target can be represented as:

$$\begin{cases} \min\limits_{\omega,b,\xi_i} J=\frac{1}{2}\omega^T\omega+\frac{1}{2}c\sum\limits_{i=1}^{n}\xi_i \\ s.t. y_i=\omega^T\varphi(x_i)+b+\xi_i, i=1,2,\ldots,n \end{cases} \tag{2}$$

Where c is a regularization function, which can take a compromise between the training error and complexity of the model, so that the required functions have good generalization; b is the error variable.

Through its dual form can find its optimal solution. And the dual form can establish the Lagrange function by the objective function and constraint condition:

$$L=\frac{1}{2}\omega^T\omega+\frac{1}{2}c\sum_{i=1}^{n}\xi_i^2-\sum_{i=1}^{n}\alpha_i\left\{\omega^T\varphi(x_i)+b+\xi_i-y_i\right\} \tag{3}$$

Where α_i is the inductive Lagrange multiplier.

According to the KKT conditions of optimal theory, we can get:

$$\begin{cases} \dfrac{\partial L}{\partial \omega} = 0 \to w = \sum_{i=1}^{n} \alpha_i \varphi(x_i) \\ \dfrac{\partial L}{\partial b} = 0 \to \sum_{i=1}^{n} \alpha_i = 0 \\ \dfrac{\partial L}{\partial \xi_i} = 0 \to \alpha_i = \gamma \xi_i \\ \dfrac{\partial L}{\partial \alpha_i} = 0 \to w^T \varphi(x_i) + b + \xi_i - y_i = 0 \end{cases} \quad (4)$$

We can obtain parameters α、b by eliminating variables w, ξ_i .And the forecasting mode is the equation (5):

$$f(x) = \sum_{i=1}^{n} \alpha_i \varphi(x_i)^T \varphi(x_j) + b \quad (5)$$

According to the Mercer condition, there is a mapping function and the kernel function .Using the kernel function method, $K(x_i, x_j)$ is defined as:

$$K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j) \quad (6)$$

So the forecasting model of LS-SVM is:

$$f(x) = \sum_{i=1}^{n} \alpha_i K(x_i, x_j) + b \quad (7)$$

III. THE PRINCIPLE OF QPSO

In QPSO, we think that particles have quantum behavior. When each particle moves in the search space, there is a DELTA potential well which centers on the optimal solution ($pbest$) that been found by the particle itself. Due to the nature of the particles' aggregative state is completely different in quantum space, there is no determined trajectory when particles move. This make the particles in the whole feasible solution space can find the global optimal solution($gbest$),and the evolution equation of QPSO have no velocity vector, less parameters, its form is simpler and easier to control. So QPSO is better than PSO in search ability [7]. Its principle is as follows:

In a N-dimensional search space, the particle swarm composed of M particles which represented potential problem solutions is defined as $X(t) = \{X_1(t), X_2(t), \mathrm{L}, X_M(t)\}$. In t-th moment, the position of the i-th particle is:

$$X_i(t) = [X_{i,1}(t), X_{i,2}(t), \mathrm{L}, X_{i,N}(t)] \quad i = 1, 2, \mathrm{L}, M \quad (8)$$

The particle has no velocity vector. The individual best position can be expressed as $P_i(t) = [P_{i,1}(t), P_{i,2}(t), \mathrm{L}, P_{i,N}(t)]$ .The global best position of the group is $G(t) = [G_1(t), G_2(t), \mathrm{L}, G_N(t)]$,and $G(t) = P_g(t)$, g is the index of the particles in the global best position, $g \in \{1, 2, \mathrm{L}, M\}$.

For the minimization problem, the objective function value is smaller, the corresponding adaptive value is better. The individual best position of the particle is determined as:

$$P_i(t) = \begin{cases} X_i(t) & when\ f[X_i(t)] < f[P_i(t-1)] \\ P_i(t-1) & when\ f[X_i(t)] \geqslant f[P_i(t-1)] \end{cases} \quad (9)$$

The global optimum position of the group is determined as:

$$g = \arg \min_{1 \leqslant i \leqslant M} \{ f[P_i(t)] \} \quad (10)$$

$$G(t) = P_g(t) \quad (11)$$

In the actual algorithm running, because it is needed to compute the global optimal position before every updating particle position.Therefore only need to compare the current individual fitness with the global optimum position of each particle. If the former is better, then $G(t)$ update:

$$p_{i,j}(t) = \varphi_j(t) \times P_{i,j}(t) + [1 - \varphi_j(t)] \times G_j(t)$$
$$\varphi_j(t) : U(0,1) \quad (12)$$

The evolution equation of particle is:

$$X_{i,j}(t+1) = p_{i,j}(t) \pm \alpha \times |C_j(t) - X_{i,j}(t)| \times \ln[1/u_{i,j}(t)]$$
$$u_{i,j}(t) : U(0,1) \quad (13)$$

Where α is the contraction coefficient of expansion, and it can control the convergence rate; $C_j(\mathrm{t})$ is the average best position of the particle swarm .It is defined as:

$$C_j(t) = \frac{1}{M} \sum_{i=1}^{M} P_{i,j}(t) \quad (14)$$

IV. APPLICATION OF LS-SVM IN THE SHORT-TERM POWER LOAD FORECASTING BASED ON QPSO

A. *The choice of kernel function*

Commonly used kernel functions are: linear kernel function , polynomial kernel function, the radial basis function (RBF),sigmoid kernel function. When the data in the different system process regression estimate, it exists the corresponding optimal kernel function. The most widely used is the radial basis function. Whether small sample or large sample, high dimension or low dimensional case, the radial basis kernel function are applicable. Its form of expression is simple. Even for multi-variable input, it does not increase too much complexity. Its arbitrary derivatives are present, analyticity is good, and it is advantageous for the theoretical analysis[8].

In this paper, we choose the radial basis function as the kernel function in the regression model, the expression is as follows:

$$K\left(x, x_i\right) = \exp\left(-\left\|x - x_i\right\|^2 / \sigma^2\right) \tag{15}$$

Where x is a m -dimensional input vector; x_i is the center of the first i radial basis function, and its dimension is same as x . σ is the standardized parameters, and it determines the width that the function around the center. $\left\|x - x_i\right\|$ is the norm of vector $x - x_i$,and it is the distance between x and x_i .

B. Parameter selection

Through a large number of experimental studies we have found that penalty factor C and the parameters of kernel function play a very important role on the performance of the SVM[9][10]. Penalty factor C regulate the fiducial range of machine learning and the proportion of empirical risk in determining the characteristics space. The parameter of kernel function σ reflects the degree of correlation of the support vector and influences the complexity that learning samples distribute in high dimensional feature space. It plays a vital role to the generalization ability of SVM. In this article we use QPSO to choose σ and C .The initialization of parameters in QPSO is as follows:

The number of particles $m = 20$; Maximum iterations $T_{\max} = 200$; The range of inertia weight coefficient ω is $\left[0.4, 0.9\right]$;Acceleration constant $c_1 = 1.5$, $c_2 = 1.7$ ..

C. QPSO-LS-SVM model

(1) Initialization of particle swarm: Ordering the number of particles m , maximum iterations $T_{\max}$, inertia weight coefficient ω ,acceleration constant c_1 and c_2 .

(2)Define the fitness function:

$$F = \frac{1}{N}\sum_{i=1}^{n}\left|\frac{Y_i^{'} - Y_i}{Y_i}\right| \tag{16}$$

Where N is the number of samples; $Y_i^{'}$ is the forecasting sample estimated value; Y_i is the forecasting sample actual value. If the fitness function value of particles is larger, the degree of adaptation is smaller. If the fitness function value of particles is smaller, the degree of adaptation is larger.

(3)Using the equation (14) to calculate the average optimal position of particle swarm;

(4)Calculate the fitness value in its current position for each particle. According to the equation (9) to update particle individual optimal position.

(5)For the i-th particle, compare the fitness value of $P_i\left(t\right)$ to the global optimal position $G\left(t-1\right)$ .If the former is superior to the latter, then $G\left(t\right) = P_i\left(t\right)$. Otherwise, $G\left(t\right) = G\left(t-1\right)$.

(6)Using the equation (13) to calculate the particle's new position.

(7)If it does not meet the end condition of algorithm, ordering $t = t + 1$, return to step 3. Otherwise, the optimization ends.

(8)Put the optimal position $\left(C, \sigma\right)$ to LS – SVM model.

V. INSTANCE ANALYSIS

In this paper, we use the sample of 96 points daily load data, weather factors (including the daily maximum temperature, minimum temperature, weather type) and date type from October 10, 2012 to October 17, 2012 of Hefei to train LS-SVM model and forecast the 96 points load of October 18, 2012.

Firstly preprocess the load data(including the repair of missing data load, vertical and horizontal data processing). In order to avoid calculating saturation phenomenon, the load data need normalization processing to make the input load data in $\left[0, 1\right]$.The equation is as follows:

$$L = \frac{L_t - L_{\min}}{L_{\max} - L_{\min}} \tag{17}$$

Where $L_{\max}$, $L_{\min}$ are the maximum and the minimum of training sample set load.

We do quantitative data processing to the weather factors (temperature, weather type). According to the characteristics of the effect that the date type to load forecasting, date type can be divided into 4 classes: Monday, workday(from Tuesday to Friday),Saturday, Sunday. Then quantify the four date types. Thereby forming a sample set of training and prediction, and then based on QPSO optimized LS-SVM model to forecast load.

In order to compare the forecast result, at the same time, we use artificial neural network, the traditional LS - SVM ,PSO optimize LS-SVM to forecast in this paper. Figure 1 is the area chart that the four methods to predict the results.

In this article, Mean Absolute Percentage Error($MAPE$)is used as the predict performance evaluation index. It is defined as:

$$MAPE\left(\%\right) = \frac{1}{N}\sum_{i=1}^{n}\left|\frac{Y_i^{'} - Y_i}{Y_i}\right| \times 100\% \tag{18}$$

Where N is the number of samples; $Y_i^{'}$ is the forecasting sample estimated value; Y_i is the forecasting sample actual value.

Through analysis, we obtained the $MAPE\left(\%\right)$ of artificial neural network, the traditional LS-SVM, PSO-LS-SVM, as shown in table 1.Compared with $MAPE\left(\%\right)$ of the PSO-LS-SVM model, QPSO-LS–SVM model was reduced by 0.56%,and the relative errors

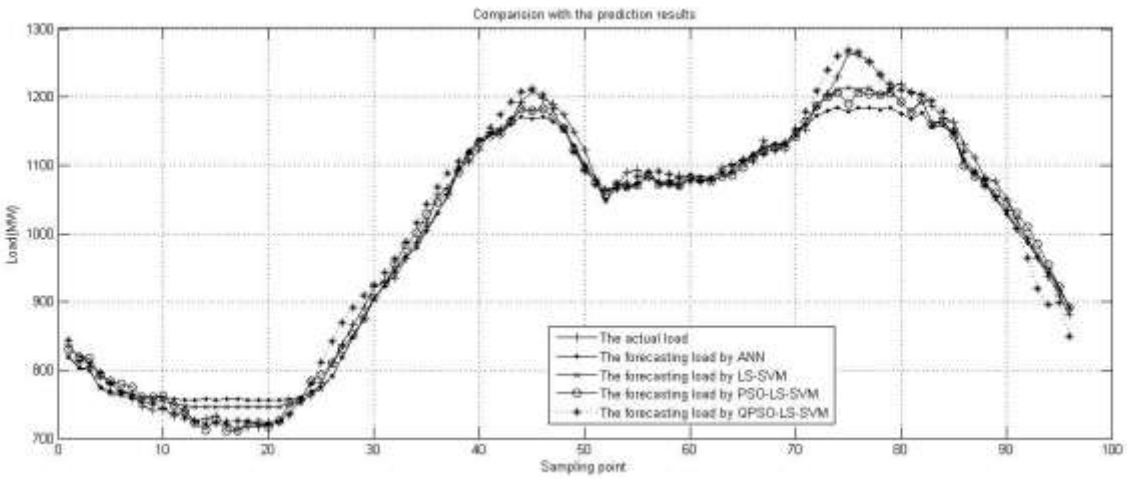

Figure 1. Comparision of the prediction results of ANN、LS-SVM、PSO-LS-SVM、QPSO-LS-SVM

of 96 point prediction are less than 3% and in the acceptable range .The maximum relative error is 2.38% and the minimum relative error is 0.08%. This shows that QPSO-LS-SVM model has higher predicted accuracy, the sensory ability of peaks and troughs improve remarkably.

TABLE I. THE FORECASTING EVALUATION ANALYSIS TABLE

	ANN	LS-SVM	PSO-LS-SVM	QPSO-LS-SVM
MAPE	2.67	1.89	1.47	0.91

VI. CONCLUSIONS

It is proposed a LS-SVM model that based on QPSO optimize the parameters to load forecast. It overcomes the limitations and blindness of the traditional SVM parameter selection.The instance simulate results show that the model convergence is good, have higher precision and faster training speed.

REFERENCES

[1] ZHOU Chao, XING Wen-yang, LI Yu-long. Summarization on load forecasting method of electrical power system[J]. Journal of Power Supply,2012,06:32-39.

[2] LI Peng-fei,Jiamalihan•Kumashi,CHANG Xi-qiang,et al.Research on the application of load forecasting based on support vector machine[J].Heilongjiang Electric Power,2012,34(4):259-262.

[3] CHEN Gan,YAN Fei,GONG Xiao,et al. State estimate based on parameter-optimized least square support vector machines[J]. Power System Protection and Control,2011,39(19):83-88.

[4] TANG Jun.Principle and application of PSO algorithm[J].COMPUTER TECHNOLOGY AND DEVELOPMENT,2010,20(2):213-216.

[5] SUN Jun.Research on quantum-behaved particle swarm optimization algorithm[D].Wuxi:Jiangnan University,2009.

[6] LIU Zi-xi, ZHANG Hang. Optimal design of RBF NN based on QPSO algorithm [J]. Journal of Central South University (Science and Technology),2013,44:27-30.

[7] SUN Jun,FANG Wei,WU Xiao-bo,et al. Quantum-behaved particle swarm optimization: principles and applications[M]. Beijing: Tsinghua University Press,2011.

[8] ZHANG Zheng-guo,WU Yan-zeng. The short-term power load forcasting based on genetic algorithm optimized LS-SVM[J]. Journal of Lanzhou Jiaotong University,2012,06:44-48.

[9] JIANG Yan,LAN Sen,SUN Yan-xue. Short-term load forecasting based on least square support vector machine[J]. Heilongjiang Electric Power,2012,35(5):349-352.

[10] WANG Ben,LENG Bei-xue,ZHANG Xi-hai,et al. Application profiles of support vector machine in short-term load forecasting[J]. Proceedings of the CSU-EPSA,2011,23(4):115-121.