Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 13 (9): 1467-1472, 2013
ISSN 1812-5654 / DOI: 10.3923/jas.2013.1467.1472

Improved Support Vector Machine Short-term Power Load Forecast Model Based on Particle Swarm Optimization Parameters

Xiaopeng Guo, Xiaodan Guo and Juan Su
School of Economics and Management, North China Electric Power University,
Beijing, 102206, China

Abstract: An improved short-term power load forecast model that uses Support Vector Machine (SVM) was developed. The new model, called the PSO-SVM forecast model, is based on Particle Swarm Optimization (PSO) parameters. In traditional SVM models, penalty factor C and kernel function parameter σ are generally dependent on particle experience. When power load forecast data change, however, obtaining satisfactory forecast precision using these empirical values is difficult to accomplish. Therefore, this study used PSO to optimize the parameter selection methods of SVM in accordance with training data and improved SVM forecast precision. PSO-SVM is generalizable and easily expandable. To verify the validity of the model, this study selected and analyzed integral point data on Fujian Province in October 2011. Data for October 1-25 were used for training and those for October 26-30 were employed for testing. The PSO-SVM model was then employed to forecast and analyze the October 31 data. Results show that the forecast efficiency of PSO-SVM was better than that of traditional SVM. In contrast to the forecast efficiency of GA-SVM, PSO-SVM was slightly better. In addition, PSO-SVM exhibited better operational performance than did GA-SVM.

Key words: PSO, SVM, short-term load forecasting

INTRODUCTION

Power load forecasting is an important operational component of electric power departments because it plays a vital role in guaranteeing the security and reliability of a power system. Decision-making in power planning, dispatching and power market transactions depends on load forecast. Load change is a complex process that is influenced by many conditions of uncertainty. From a macroscopic perspective, however, it also has recognizable regularity under certain time horizons. Many scholars have conducted research on long, medium and short-term power load forecasting. The induced ordered weighted geometry averaging operator, weighted Markov chain (Mao *et al*., 2010) model and BP neural network optimized by Particle Swarm Optimization (PSO) (Cui *et al*., 2009) were combined for medium-and long-term load forecasting. The Grey forecast model-based BP neural network and Markov chain were used to forecast China's electricity demand (Li and Wang, 2007). Self-organizing neural network (Zhao and Xu, 2010), least squares Support Vector Machine (SVM) (He *et al*., 2011) and combined SVM and rough set's model (Niu *et al*., 2010; Li *et al*., 2009; Yang *et al*., 2011) were used to forecast short-term power load.

Although, good progress has been made in using the above-mentioned algorithms for short-term load forecasting, neural networks and support vector machines suffer from shortcomings, such as easily falling into local extrema, overlearning and so on. Some scholars attempted to establish an SVM forecast model using a Genetic Algorithm (GA) (Wu *et al*., 2009) and ant colony algorithm (Long *et al*., 2011) to optimize parameters. However, GAs entail a series of more complex operations, including coding, selection, crossover and mutation, whereas PSO is relatively simple. In the current work, therefore, this study established a short-term load forecast model and employed PSO to optimize the core parameters of SVM. The proposed model was analyzed and validated using actual data on a region.

MATERIALS AND METHODS

Overview of SVM regression: SVM was put forward by Vapnik on the basis of small-sample statistical learning theory (Vapnik, 2000), which is used primarily to study small samples under statistical learning rules and is commonly adopted in pattern classification and nonlinear regression (Thissen *et al*., 2003; Kim, 2003).

The sample data set is given as D = {(xi, yi)| i = 1, 2, …, n}, where xi ∈ Rn represents the input variables and yi ∈ Rn denotes the output variables.

The SVM algorithm seeks one misalignment mapping from the input space to output space φ. Through this

Corresponding Author: Xiaopeng Guo, School of Economics and Management, North China Electric Power University, Beijing, 102206, China

mapping, data × is mapped to a feature space Γ and linear regression is carried out in the feature space with the following function:

$$\begin{aligned} & f(x) = [\omega \times \varphi(x)] + b \\ & \phi: R^m \to \Gamma \end{aligned} \tag{1}$$

In Eq. 1, b is a threshold value. According to statistical learning theory, SVM determines the regression function through objective function minimization:

$$\begin{aligned} & \min \left\{ \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}\left(\xi_i^* + \xi_i\right) \right\} \\ & \text{s.t.} \quad y_i - \omega\cdot\phi(x) - b \le \varepsilon + \xi_i^*; \\ & \qquad \omega\cdot\phi(x) + b - y_i \le \varepsilon + \xi_i; \\ & \qquad \xi_i, \xi_i^* \ge 0 \end{aligned} \tag{2}$$

where, C is a weight parameter for balancing the complex items of the model and training error, also called the penalty factor; ε is the insensitive loss function; and ξi* and ξi are the relaxation factors. ξi* is expressed as follows:

$$\xi_i^* = \begin{cases} 0, |f(x) - y_i| \le \varepsilon \\ |f(x) - y_i| - \varepsilon, \quad |f(x) - y_i| > \varepsilon \end{cases} \tag{3}$$

By solving the dual problem in Eq. 2, lag range factors ai, ai* can be obtained, so that the regression equation coefficient is:

$$\omega = \sum_{i=1}^{n}(a_i - a_i^*)x_i \tag{4}$$

The SVM regression equation is as follows:

$$f(x) = \sum_{i=1}^{n}(a_i - a_i^*)\, K(X_i, X) + b \tag{5}$$

where, K (Xi, X) is the SVM kernel function. Kernel function types include linear kernels, polynomial kernels and radial basis functions.

Penalty factor C, insensitive loss function ε and kernel function parameter σ determine SVM performance. The σ responds to the training data set characteristics, determines the complexity of the solution and affects the generalizability of the learning machine. Parameter C determines the penalty to large fitting deviation: An excessively large value may cause overlearning but one too small easily results in less learning. The optimization of these parameters is therefore important in improving SVM performance.

Overview of PSO: PSO is an evolutionary computation based on swarm intelligence; it was proposed by Kennedy and Eberhart (1995). Its basic concept stems from the study of bird predation.

In PSO, particles identify potential optimal solutions in a solution space. Three targets-the position, speed and fitness value-express the characteristics of the particles. The fitness value is obtained by the fitness function and is used to express whether the particle is fit or unfit. Individual positions are updated through the track individual extreme value (denoted as Pbest) and the group extreme value (denotes as Gbest). The individual extreme value is the optimal solution of the fitness value in the particle's experiences and the group extreme value is the optimal solution of the fitness value in the entire particle population.

Assuming an N dimension search space, this study define a population set X, including n particles $X = (X_1, X_2, \ldots, X_n)$. In X, the ith particle is the position in the N dimension search space (i.e., a potential solution), denoted as an N dimension vector $Xi = (X_{i1}, X_{i2}, \ldots, X_{in})$. The speed of the ith particle is denoted as $Vi = (V_{i1}, V_{i2}, \ldots, V_{in})T$. Individual extreme value Pbest is denoted as $Pi = (P_{i1}, P_{i2}, \ldots, P_{in})T$ and group extreme value Gbest is designated as $Pg = (V_{g1}, V_{g2}, \ldots, V_{gn})T$.

In the PSO algorithm iterative process, the particle updates its own speed and position using Eq. 6 and 7:

$$V_{id}^{k+1} = \omega V_{id}^{k} + c_1 r_1 (P_{id}^{k} - X_{id}^{k}) + c_2 r_2 (P_{gd}^{k} - X_{gd}^{k}) \tag{6}$$

$$X_{id}^{k+1} = X_{id}^{k} + V_{id}^{k+1} \tag{7}$$

In Eq. 6, ω is the inertia weight; d = 1, 2, …, N; i = 1, 2, …, n; k denotes the current iteration times; Vid is the particle speed; c_1 and c_2 represent nonnegative constants called acceleration factors; and r_1 and r_2 are random numbers distributed between (0, 1).

The classical PSO algorithm features fast convergence and strong currency but also suffers from shortcomings such as premature convergence, low precision search and low efficiency of late period search. Therefore, the PSO precedent derived from GA introduces a random factor into1 the iterative process (mutation factor (Higasshi and Iba, 2003) via a probability. This probability is used to re-initialize the particle and expand the search space. Through this method, the algorithm is prevented from getting caught in local extrema.

A high inertia weight value is advantageous to global search, while a low value is beneficial to local search. To balance the global and local search ability of PSO, this study applied a series of weight selection methods, including linearly decreasing inertia weight (Shi and Eberhart, 1999; Jin *et al*., 2006).

SVM's Penalty factor C and kernel function parameter σ optimized by PSO: The construction of the SVM forecast model based on PSO optimization parameters (PSO-SVM) entails seven steps (Fig. 1).

Step 1: Initialize information, including population size, particle position, initial particle speed, range of speed, acceleration factor, penalty factor C, kernel function parameter σ and mutation probability
Step 2: The fitness value is calculated with input data. In this step, the SVM kernel function is used to calculate the fitness value in the training process
Step 3: The optimal solution of the iterator is solved
Step 4: Evaluate whether the iteration suspension conditions, including maximum iterator times or ata precision, are satisfied
Step 5: When the suspension conditions are not satisfied Eq. 6 and 7 are used to update the particle information. Step 2 is then repeated
Step 6: When the suspension conditions are satisfied, output optimal solutions C and σ
Step 7: C and σ are used to construct the SVM forecast model and then regression forecasting is executed

RESULTS

Forecast data selection and pretreatment of application case: Integral point data were collected from selected areas in Fujian Province in October 2011. The data were designated as training, testing and forecast data sets.

Data for October 1-25 were used for training and those for October 26-30 were employed for testing. On the basis of the training and testing, this study constructed the SVM forecast model. The model was then used to forecast the October 31 data. Finally, this study analyzed the predicted and actual data.

To obtain better convergence results, this study normalized the training data, while data testing and forecasting were carried out at a distribution between (0, 1). The entire forecasting process was coded by MATLAB and the LibSVM (Chang and Lin, 2011) toolbox.

PSO and GA are heuristic algorithms. The derived parameter optimization results differ each time. The predicted values and precision also fluctuate at a small scale. Through repeated testing and analyses, however, the operational efficiency and forecast precision of these algorithms are generally stable.

Forecast results: Table 1 shows the actual load data, PSO-SVM-predicted load data, GA-SVM-predicted load data, traditional SVM-predicted data and each predicted data set error and Root Mean Squared Relative Error (RMSRE). RMSRE is expressed as Eq. 8, where e_i is relative error:

$$\text{RMSRE} = \sqrt{\frac{\sum_{i=1}^{N} e_i^2}{N}} \tag{8}$$

Errors and RMSRE are important evaluation indicators of forecast results. The smaller the values obtained, the more accurate the model forecast.

The contrasting results for the actual, PSO-SVM, GA-SVM and SVM data are shown in Fig. 2a. The contrasts in error of the three forecast models are shown in Fig. 2b.

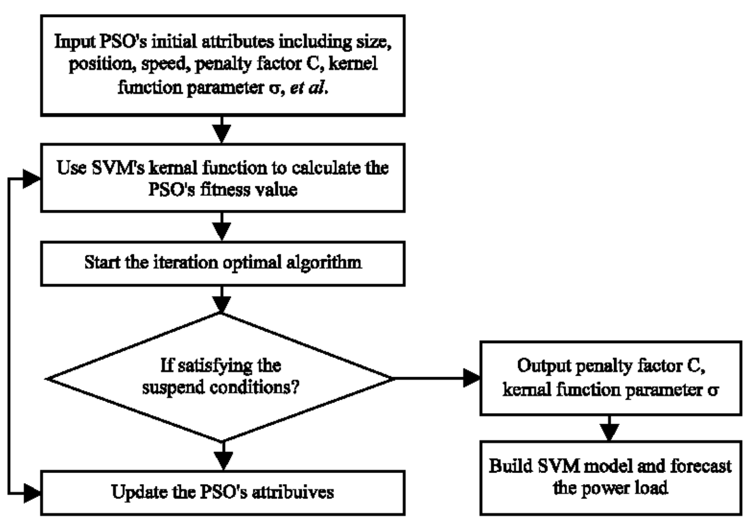

Fig. 1: Iteration algorithm of PSO Optimize SVM's penalty factor C, kernel function parameter σ

Table 1: Integral point load data and analysis of forecast results, 31 October 2011

		PSO-SVM		GA-SVM		SVM	
Time	Original (kW h^{-1})	Predict (kW h^{-1})	Error (%)	Predict (kW h^{-1})	Error (%)	Predict (kW h^{-1})	Error (%)
0	957	945.53	-1.19850	951.35	-0.59039	941.43	-1.62700
1	940	929.21	-1.14790	926.05	-1.48400	928.19	-1.25640
2	909	911.42	0.26623	903.40	-0.61606	915.56	0.72167
3	876	889.66	1.55940	896.76	2.36990	899.64	2.69860
4	881	890.10	1.03290	878.85	-0.24404	900.86	2.25430
5	883	892.80	1.10990	873.33	-1.09510	912.17	3.30350
6	948	951.76	0.39662	961.79	1.45460	945.30	-0.28481
7	994	995.92	0.19316	999.83	0.58652	1001.5	0.75453
8	1000	996.50	-0.35000	989.06	-1.09400	1002.4	0.24000
9	1011	987.73	-2.30170	999.20	-1.16720	988.95	-2.18100
10	1035	1002.9	-3.10140	1002.9	-3.10140	1012.2	-2.20290
11	1116	1066.9	-4.39960	1069.2	-4.19350	1056.7	-5.31360
12	1008	1009.8	0.17857	1006.7	-0.12897	1022.5	1.43850
13	959	945.92	-1.36390	938.14	-2.17520	940.34	-1.94580
14	970	961.18	-0.90928	984.44	1.48870	954.47	-1.60100
15	1005	98600	-1.89050	1002.2	-0.27861	973.23	-3.16120
16	1019	1013.5	-0.54073	1006.4	-1.23650	999.30	-1.93330
17	1155	1103.8	-4.43290	1086.8	-5.90480	1089.8	-5.64500
18	1215	1146.3	-5.65430	1139.4	-6.22220	1131.4	-6.88070
19	1105	1072.3	-2.95930	10590	-4.16290	1078.8	-2.37100
20	1077	1036.5	-3.76140	1029.9	-4.37330	1037.6	-3.65830
21	1061	1026.5	-3.25160	1021.7	-3.70410	1022.4	-3.63810
22	1035	1012.6	-2.16430	1003.4	-3.05120	1026.6	-0.81159
23	1007	985.90	-2.09530	999.04	-0.79047	979.68	-2.71300
RMSRE		0.0244		0.0277		0.0295	

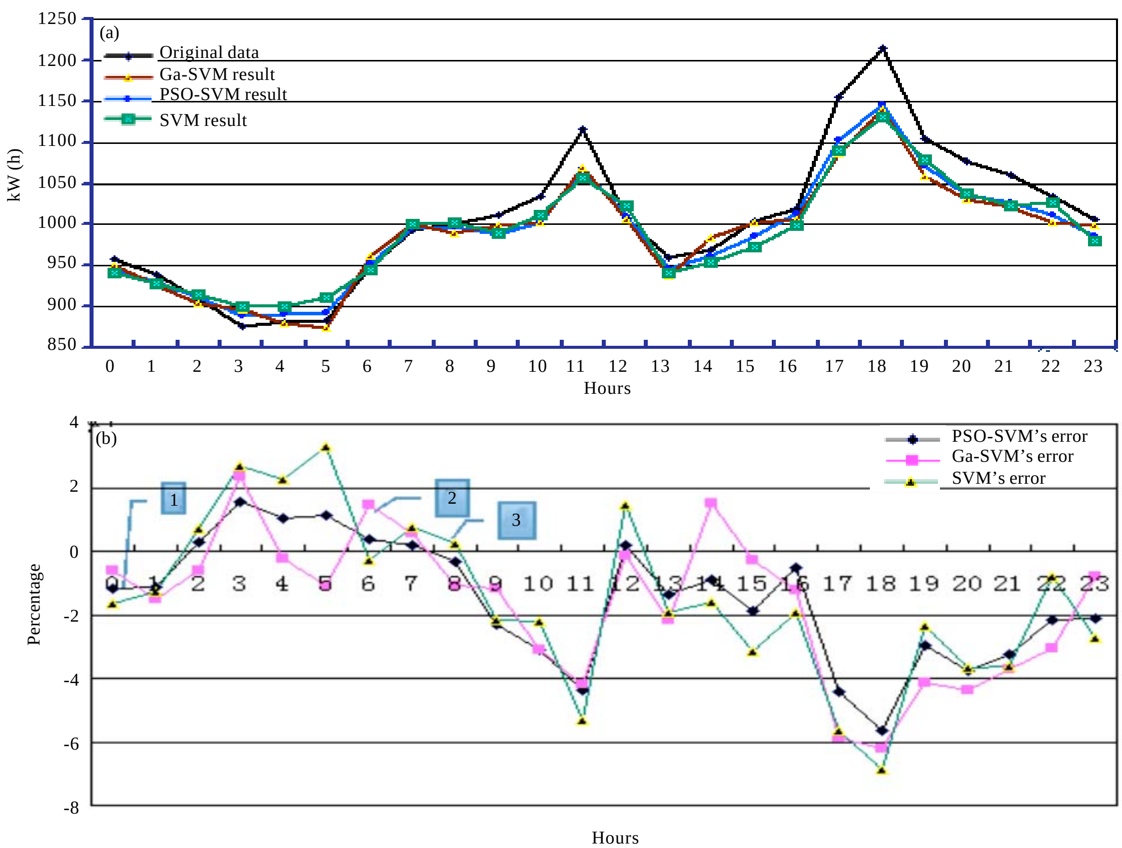

Fig. 2(a-b): Contrast of original data, the forecast results (by PSO-SVM, Ga-SVM and SVM) and the contrast of three forecast model's errors

DISCUSSION

According to (Li *et al*., 2009), when the absolute value of an error is smaller than 3%, the forecast results can be considered ideal.

As shown in Table 1 and Fig. 2b, of the 24 errors identified by PSO-SVM, 18 were smaller than 3%, with the smallest error amounting to only 0.178%. The largest error was that observed in 31 October 2011 at 18:00, at a value of -5.65%. Furthermore, two errors approached 3% and three amounted to about 4%. These results are better than those derived via GA-SVM (16 errors were smaller than 3%, with the smallest at -0.244% and the largest at -6.22%) and traditional SVM (17 errors were smaller than 3%, with the smallest and largest being -0.24 and -6.88%, respectively).

For RMSRE values, those of PSO-SVM, GA-SVM and SVM were 0.0244, 0.0277 and 0.0295, respectively. All the forecast results are ideal. These results follow the ranking PSO-SVM>GA-SVM>SVM. During the testing, the efficiency of PSO was higher than that of GA. The operation time of PSO-SVM was 92.786364 sec, while that of GA-SVM was 271.097241 sec. The parameter selection for traditional SVM was affected by the particle experiences. An inappropriate parameter causes a large prediction error. In this study, the SVM forecast data yielded better results given the frequent adjustment in parameters C and σ.

For different training and forecast data, parameters should not be fixed. The dynamic adjustment of SVM parameters according to data characteristics can obtain results that correspond with actual forecasts. This advantage also emphasizes the necessity of using intelligent algorithms such as PSO or GA in selecting SVM parameters from the sample data.

CONCLUSION

- Parameter selection is important for SVM forecasting. Traditional SVM suffers from problems such as overlearning or underlearning. These problems diminish algorithm performance and affect forecast precision
- Using the PSO optimization choice, SVM penalty factor and kernel function parameter yielded good results. PSO-SVM also exhibited more efficient operation than did GA-SVM
- The proposed model enables good forecast results through actual data confirmation, making it a valid model. However, because power load is affected by many external factors, a multi-factor forecast model should be explored

ACKNOWLEDGMENTS

Project supported by National Natural Science Fund of China (No. 71071054) and the Fundamental Research Funds for the Central Universities of China (No. 11QR34).

REFERENCES

Chang, C.C. and C.J. Lin, 2011. Libsvm: A library for support vector machines. ACM Trans. Intell. Syst. Technol., Vol. 2, No. 3.

Cui, J.F., J.X. Qi and S.D. Yang, 2009. Combined forecasting model based on BP improved byPSO and its application. J. Central South Univ. (Sci. Technol.), 40: 190-194.

He, Y.X., H.Y. He, Y.J. Wang and T. Luo, 2011. Forecasting model of residential load based on general regression neural network and PSO-Bayes least squares support vector machine. J. Central South Univ. Technol., 18: 1184-1192.

Higasshi, N. And H. Iba, 2003. Particle swarm optimization with Gaussian mutation. Proceedings of the Swarm Intelligence Symposium, April 24-26, 2003, Technical Communications Ltd., Tokyo, pp: 72-79.

Jin, J., X.Y. Wang, X.G. Luo and B. Wang, 2006. Regression algorithm of PSO-ε-SVM. J. East China Univ. Sci. Technol. (Natl. Sci. Edn.), 32: 872-875.

Kennedy, J. and R.C. Eberhart, 1995. Particle swarm optimization. Proc. IEEE Int. Conf. Neural Networks, 4: 1942-1948.

Kim, K.J., 2003. Financial time series forecasting using support vector machines. Neurocomputing, 55: 307-319.

Li, C.B. and K.C. Wang, 2007. A new grey forecasting model based on BP neural network and Markov chain. J. Central South Univ. Technol., 14: 713-718.

Li, Y.B., N. Zhang and C.B. Li, 2009. Support vector machine forecasting method improved by chaotic particle swarm optimization and its application. J. Central South Univ. Technol., 16: 478-481.

Long, W., X.M. Liang, Z.Q. Long and Z.H. Li, 2011. Parameters selection for LSSVM based on modified ant colony optimization in short-term load forecasting. J. Central South Univ. (Sci. Technol.), 42: 3408-3414.

Mao, L.F., J.G. Yao, Y.S. Jin, H.L. Chen and W.J. Li *et al*., 2010. Theoretical study of combination model for medium and long term load forecasting. Proc. CSEE., 30: 53-59.

Niu, D.X., Y.L. Wang and X.Y Ma, 2010. Optimization of support vector machine power load forecasting model based on data mining and Lyapunov exponents. J. Pattern Recognit. Res., 17: 406-412.

Shi, Y. and R.C. Eberhart, 1999. Empirical study of particle swarm optimization. Proceedings of the IEEE International Conference on Evolutionary Computer, July 6-9, 1999, Piscataway, N.J., USA., pp: 1945-1950.

Thissen, U., R.V. Brakel, D.A.P. Weijer, W.J. Melssen and L.M.C. Buydens, 2003. Using support vector machines for time series prediction. Chemomet. Intell. Lab. Syst., 69: 35-49.

Vapnik, V.N., 2000. The Nature of Statistical Learning Theory. Tsinghua University Press, Beijing, China.

Wu, J.L., S.X. Yang and C.S. Liu, 2009. Parameter selection for support vector machines based on genetic algorithms to short-term p ower load forecasting. J. Central South Univ. (Sci. Technol.), 40: 180-184.

Yang, S.X., Y. Cao, D. Liu and C.F. Huang, 2011. RS-SVM forecasting model and power supply-demand forecast. J. Central South Univ. Technol., 18: 2074-2079.

Zhao, J. and K.M. Xu, 2010. Application of neural network and fuzzy theory in short-term load forecasting. Proc. CSU-EPSA, 22: 129-133.