International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:4, No:11, 2010

Annual Power Load Forecasting Using Support Vector Regression Machines: A Study on Guangdong Province of China 1985-2008

Zhiyong Li, Zhigang Chen, Chao Fu, Shipeng Zhang

***Abstract*—**Load forecasting has always been the essential part of an efficient power system operation and planning. A novel approach based on support vector machines is proposed in this paper for annual power load forecasting. Different kernel functions are selected to construct a combinatorial algorithm. The performance of the new model is evaluated with a real-world dataset, and compared with two neural networks and some traditional forecasting techniques. The results show that the proposed method exhibits superior performance.

***Keywords*—**combinatorial algorithm, data mining, load forecasting, support vector machines

I. Introduction

LOAD forecasting is always defined as basically the art or science of predicting the future electricity demand on a given system for a specified period of time ahead. This issue plays a dominant part in the economic optimization and secure operation of electric power systems.

Depending on different time horizons, load forecasting can be generally divided into short-term and mid-long-term categories. Short-term load forecasting, ranging from an hour to a week, is essential in unit commitment, economic dispatch and real-time control, and thus, has been under great focus worldwide [1]-[3].

Mid-long-term load forecasting, on the other hand, covers horizons of a few months to several years. It represents the first step in developing future generation, transmission, and distribution facilities. Compared with extensive research throughout the globe on short-term load forecast, fewer technical studies discuss how to acquire advance knowledge of electrical load in a year or even longer period. However, the accuracy of long-run electricity demand forecast has significant effect on planning future generation and distribution networks. Therefore, for developing countries where the electricity consumption increases with high growth rate, mid-long-term load forecasting still plays a decisive role to provide a reliable energy supply for economic development. Several researches of mid-term or long-term load prediction on typical developing countries, like Vietnam [4], Iran [5], Thailand [6], Egypt [7], Brazil [8], Saudi Arabia [9] and China [10]-[11], were reported recently.

The relationship between electric load growth in developing countries and its exogenous factors is always complex and nonlinear, making it quite difficult to be modeled through traditional deterministic techniques such as linear regression or elastic coefficient method. With the development of artificial intelligence techniques, different intelligent algorithms have been put forward and applied in mid-long-term load forecasting. Artificial Neural Networks (ANN) becomes the most widely used method. Some variants of neural networks, such as Back Propagation (BP) neural network [9], wavelet neutral network [10], Radial Basis Function (RBF) neutral network [12] and genetic neural network [13] have been adopted for the mid-long-term load forecasting. Besides, several other algorithms including Genetic Algorithm (GA) [14], Decision Tree (DT) [15], Hidden Markov Model (HMM) [16] and fuzzy rules [17] have also been employed and proved feasible.

Support Vector Machine (SVM), proposed by Vapnik and his coworkers [18], is a novel powerful machine learning method based on statistical learning theory. This Structural Risk Minimization (SRM) framework has been receiving increasing attention in many real-world applications due to several attractive features, including: 1) guaranteed global extremum by solving quadratic programming problem as opposed to gradient-based training for neural networks that might converge to local optimization; 2) the ability to handle high dimensional input spaces by adopting kernel functions; 3) a reduced sensitivity to noisy inputs and sparse training samples.

On the basis of author's previous work [19], this paper builds an improved SVM prediction model, and estimates its performance using the statistics of a fast developing region of China. This work differs from the existing reports by carefully selected inputs based on association analysis and the combinatorial algorithm using different kernel functions.

The rest of this paper is organized as follows. Section 2 introduces the basic concepts of SVM for regression. Section 3 illustrates which indices are selected as the input variables. In section 4 and 5, the empirical results of the proposed method and comparison studies using some existing approaches like BP neural network and the regularized RBF neural network are shown respectively. Finally, section 6 draws the conclusion.

Zhiyong Li, Zhigang Chen, Chao Fu, Shipeng Zhang are with Guangdong Electric Power Design Institute, Guangzhou, China (corresponding author to provide phone: +86-20-32117303; e-mail: lizhiyong@gedi.com.cn)

II. Support Vector Regression Machines

SVM was developed in the early 1990s for pattern recognition [20]. Its application was then extended to nonlinear regression by bringing in Vapnik's ε-insensitive loss function. Regression is a statistical technique for estimating the relationships among variables. Given a set of training data: $\{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \ldots (\boldsymbol{x}_n, y_n)\}$, where $\boldsymbol{x}_i \in \boldsymbol{R}$ is an input vector and y_i is the corresponding output, SVM approximates the function using the following form:

$$f(\boldsymbol{x}) = (\boldsymbol{w} \cdot \boldsymbol{x} + b) \tag{1}$$

Based on Vapnik's ε-insensitive loss function depicted in Fig. 1, the Support Vector Regression (SVR) finds the approximating function by solving a constrained optimization problem:

$$\begin{aligned} \min \quad & \frac{1}{2}\|\boldsymbol{w}\|^2 + \mathrm{C}\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\ \text{S.T.} \quad & (\boldsymbol{w}\cdot\boldsymbol{x}_i + b) - y_i \le \varepsilon + \xi_i \\ & y_i - (\boldsymbol{w}\cdot\boldsymbol{x}_i + b) \le \varepsilon + \xi_i^*, i = 1,2,\ldots,n \\ & \xi_i, \xi_i^* \ge 0 \end{aligned} \tag{2}$$

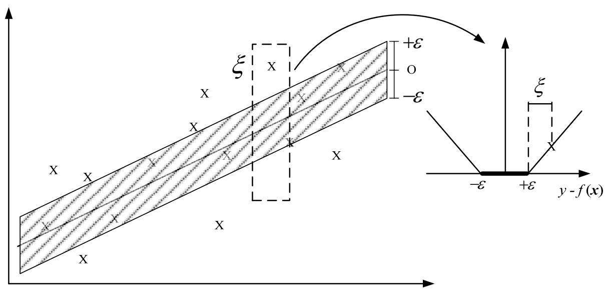

Fig. 1 The ε-insensitive loss function.

where C is the cost of error, ξ_i and ξ_i^* are slack variables, defined as the upper and lower training errors subject to Vapnik's ε-insensitive loss function.

The constraints of (2) imply that SVR tries to put most of the data $\boldsymbol{x}_i$ in the tube $|y - (\boldsymbol{w}\cdot\boldsymbol{x} + b)| \le \varepsilon$. If $\boldsymbol{x}_i$ locates within the tube, the loss is zero, whereas if $\boldsymbol{x}_i$ is outside the tube, there is an error ξ_i or ξ_i^* which is minimized in the cost function. This constrained optimization problem is usually solved in dual form by employing Lagrange multipliers:

$$\begin{aligned} \min \quad & \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(\boldsymbol{x}_i \cdot \boldsymbol{x}_j) + \varepsilon\sum_{i=1}^{n}(\alpha_i^* + \alpha_i) \\ & - \sum_{i=1}^{n} y_i(\alpha_i^* - \alpha_i) \\ \text{S.T.} \quad & \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0, \quad 0 \le \alpha_i, \alpha_i^* \le C, i = 1,2,\ldots,n \end{aligned} \tag{3}$$

For non-linear problem solving, input data are first mapped into a higher dimensional feature space $\phi(\boldsymbol{x})$ and then the linear regression can be carried out. In SVM theory, fortunately, kernel function can be applied to facilitate the computation. The value of the kernel is equal to the inner product of two vectors $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ in the feature space $\phi(\boldsymbol{x}_i)$ and $\phi(\boldsymbol{x}_j)$:

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{\Phi}(\boldsymbol{x}_i)\cdot\boldsymbol{\Phi}(\boldsymbol{x}_j) \tag{4}$$

Functions that satisfy Mercer's condition can be used as the kernels. Common examples of the kernel function include:

Linear kernel $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i \cdot \boldsymbol{x}_j$ (5)

Radial basis function kernel $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp(-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{2\sigma^2})$. (6)

Sigmoid kernel $K(\boldsymbol{x}, \boldsymbol{x}') = \tanh(\beta_0 \boldsymbol{x}\cdot\boldsymbol{x}' + \beta_1)$ (7)

After finding Lagrange multiplier vectors $\hat{\boldsymbol{\alpha}}^*$ and $\hat{\boldsymbol{\alpha}}$, an optimal desired weights vector of the Kernels expansion and an optimal bias $\hat{b}$ can be determined. Finally, the best non-linear regression hyper function is given by

$$f(\boldsymbol{x}) = \sum_{i=1}^{n}(\hat{\alpha}_i^* - \hat{\alpha}_i)K(\boldsymbol{x}_i, \boldsymbol{x}) + \hat{b} \tag{8}$$

III. Data Selection and Preprocessing

A. Data Selection

Annual data of Guangdong province, covering the period 1985-2008 were investigated in this study. The socio -economic data is collected from official statistical almanac and the load data is obtained directly from Guangdong Power Grid Company. The data set is divided into two parts. The first part, up to 2003, is used to construct the forecasting model. While the next part, from 2004 to 2008, is used to evaluate the forecasting process.

B. Variables Considered

The relationship between electric load and its exogenous factors is complex and nonlinear, making it quite difficult to find the key factors in so many options. Based on the author's association analysis of dominant factors presented in [19], nine socio-economic indices were declared as dominant factors in power load growth, and thus chosen as the inputs to the SVM:

1) Gross Domestic Product (GDP) [$\boldsymbol{x}$1(100 million yuan)]
2) Secondary industry GDP [$\boldsymbol{x}$2(100 million yuan)]
3) Per Capita GDP [$\boldsymbol{x}$3(yuan)]
4) Per capita annual consumption [$\boldsymbol{x}$4(yuan)]
5) Number of employed persons [$\boldsymbol{x}$5(10000 persons)]
6) Total investment in fixed assets [$\boldsymbol{x}$6(100 million yuan)]
7) Total value of imports and exports [$\boldsymbol{x}$7(100 million US dollars)]
8) Output value of industry [$\boldsymbol{x}$8(100 million yuan)]
9) Per capita disposable income of urban households [$\boldsymbol{x}$9(yuan)]

Table I shows the value of these inputs during the investigated period, as well as y (unit: 100GWh), namely the SVM's output: electricity demand. Fig.2 displays the curves of growth rate in which we can find potential relations between electricity demand and its influencing factors.

International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:4, No:11, 2010

TABLE I
SOCIO-ECONOMIC INDICES AND ANNUAL POWER CONSUMPTION OF GUANGDONG DURING 1985-2008

Year	$x1$	$x2$	$x3$	$x4$	$x5$	$x6$	$x7$	$x8$	$x9$	y
1985	1396.95	399.8	2481.87	1626.91	2731.11	184.59	—	661.68	3471.32	175.68
1986	1574.36	432.19	2744.95	1781.47	2811.92	216.5	—	764.28	3828.86	198.4
1987	1882.94	551.04	3211.6	1895.48	2910.99	251.01	210.37	1019.57	4070.08	241.42
…	…	…	…	…	…	…	…	…	…	…
2002	13340.85	6292.68	15168.47	6119.36	4134.37	3970.69	2210.92	23455.29	11616.81	1687.83
2003	15315.29	7570.09	17201.05	7159.65	4395.93	5030.57	2835.22	29907.77	12824.96	2031.35
2004	17581.96	8993.27	19454.39	8298.04	4681.89	6025.53	3571.29	37004.16	13761.19	2387.14
2005	20008.27	10342.26	21866.73	9152.73	5022.97	7164.11	4280.02	44097.64	14628.14	2673.56
2006	22929.47	12090.1	24796.87	9985.63	5250.09	8132.37	5272.07	53384.8	15578.97	3004.03
2007	26300.1	14145.42	28045.26	11353.66	5402.65	9596.95	6340.35	64681.13	16607.18	3394
2008	28956.42	15758	30485.2	12261.96	5553.67	11165.06	6834.92	73264.19	17553.79	3506.78

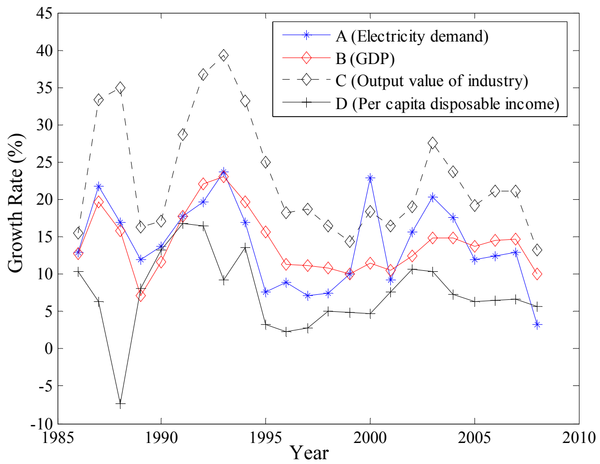

Fig.2 The growth rate of electricity demand and some influencing factors

IV. EMPIRICAL RESULTS

Estimation accuracy of SVM regression depends on the choice of key parameters. Because of the complexity of historical load data and the uncertainty of the influencing factors, no single kernel function can be proved exclusive feasible and perform well enough in different context. Therefore, an idea of combinatorial model is adopted in this work to mitigate forecasting risk. Three different kernels were embedded, and average value was calculated as the forecasted output. Table II shows the results.

TABLE II
FORECASTING RESULTS USING SUPPORT VECTOR REGRESSION MACHINES

Year	Actual Power Load	Linear Kernel		RBF Kernel	
		Forecasted Value	Error	Forecasted Value	Error
2004	2387.14	2445.8	2.46%	2088.5	-12.5%
2005	2673.56	2847.6	6.51%	2606.2	-2.52%
2006	3004.03	3112.7	3.61%	2987.9	-0.54%
2007	3394.00	3420.9	0.79%	3392.1	-0.05%
2008	3506.78	3694.5	5.35%	3576.6	1.99%
Year	Actual Power Load	Spline Kernel		Combinatorial Algorithm	
		Forecasted value	Error	Forecasted value	Error
2004	2387.14	2272.6	-4.8%	2269.0	-4.95%
2005	2673.56	2680.7	0.27%	2711.5	1.42%
2006	3004.03	2586.6	-13.9%	2895.6	-3.61%
2007	3394.00	3238.9	-4.57%	3350.6	-1.28%
2008	3506.78	3693.7	5.33%	3654.8	4.22%

V. COMPARISON OF FORECASTING PERFORMANCE WITH OTHER REPORTED METHODS

The performance of several other methods with the same historical data is also tested in order to reveal the advantage of the proposed method. The comparative methods include two widely used neural networks (BP and RBF algorithms) and three traditional techniques: elastic coefficient method I and II (averaging elastic coefficients of the preceding five years and 10 years, respectively), and conic fitting method.

The forecasting results for each model are presented in Table III, and the maximum errors in each column are highlighted with shadow. The improved accuracy of the proposed method indicates that optimized SVM algorithm is effective in mid-long-term power load forecasting.

TABLE III
COMPARISON OF THE PROPOSED METHOD AND OTHER MODELS

Year	Actual Power Load	BP Neural Network		RBF Neural Network	
		Forecasted Value	Error	Forecasted Value	Error
2004	2387.14	2248.2	-5.82%	2184.1	-8.50%
2005	2673.56	2583.5	-3.37%	2761.3	3.28%
2006	3004.03	2826.9	-5.90%	3062.4	1.94%
2007	3394.00	3148.5	-7.23%	3285.2	-3.20%
2008	3506.78	3436.5	-2.00%	3586.7	2.28%
Year	Actual Power Load	Elastic Coefficient Method I		Elastic Coefficient Method II	
		Forecasted Value	Error	Forecasted Value	Error
2004	2387.14	2422.3	1.47%	2331.1	-2.35%
2005	2673.56	2828.5	5.80%	2726.2	1.97%
2006	3004.03	3109.2	3.50%	3090.2	2.87%
2007	3394.00	3493.6	2.93%	3478.0	2.48%
2008	3506.78	3747.6	6.87%	3770.4	7.52%
Year	Actual Power Load	Conic fitting method		Proposed Method	
		Forecasted Value	Error	Forecasted Value	Error
2004	2387.14	2047.1	-14.24%	2269.0	-4.95%
2005	2673.56	2394.3	-10.45%	2711.5	1.42%
2006	3004.03	2747	-8.56%	2895.6	-3.61%
2007	3394.00	3118.9	-8.11%	3350.6	-1.28%
2008	3506.78	3525.1	0.52%	3654.8	4.22%

VI. Conclusions

In this paper, we have proposed a new approach to solve the load forecasting problem. This approach adopted three kernel functions to construct a support vector regression algorithm. It is verified that a combinatorial model is more reasonable comparing with arbitrarily selecting a single parameter to mitigate forecasting risk under uncertainty.

A real-world dataset of Guangdong, one of the most economically prosperous provinces has been used to evaluate the performance of the proposed model. In addition, several existing methods have been tested with the same historical statistics. The improved accuracy implies that this method is a potentially useful alternative for real application.

References

[1] E. E. Elattar, J. Goulermas, and Q. H. Wu, "Electric load forecasting based on locally weighted support vector regression," *IEEE Transactions on Systems, Man, and Cybernetics — Part C: Applications And Reviews*, vol. 40, pp. 438–447, July 2010.

[2] Y. Zhang, Q. Zhou, C. X. Sun, S. L. Lei, Y. M. Liu, and Y. Song, "RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment," *IEEE Transactions on Power Systems*, vol. 23, pp. 853–858, August 2008.

[3] V. H. Hinojosa, and A. Hoese, "Short-term load forecasting using fuzzy inductive reasoning and evolutionary algorithms," *IEEE Transactions on Power Systems*, vol. 25, pp. 565–574, February 2010.

[4] Ng. Q. Thi, Tr. A. Dung, Ph. Th. Th. Binh, and Tr. Q. D. Khoa, "Application wavelet for middle term demand forecasting," in *Proc. 2005 IEEE/PES Transmission and Distribution Conference & Exhibition: Asia and Pacific*, 2005, pp. 1-6.

[5] A. Ghanbari A. Naghavi, S. F. Ghaderi, and M. Sabaghian, "Artificial neural networks and regression approaches comparison for forecasting Iran's annual electricity load," in *Proc. International Conference on Power Engineering, Energy and Electrical Drives*, 2009, pp. 675-679.

[6] P. Bunnoon, K. Chalermyanont, and C. Limsakul, "Mid term load forecasting of the country using statistical methodology case study in Thailand," in *Proc. 2009 International Conference on Signal Processing Systems*, 2009, pp. 924-928.

[7] H. A. E. Salama, A. F. A. El-Gawad, S.M. Sakr, E. A. Mohamed, and H. M. Mahmoud, "Applications on medium-term forecasting for loads and energy scales by using artificial neural network," in *Proc. 20th International Conference on Electricity Distribution*, June 2009.

[8] L. Pinto, and J. Szczupak, "Load forecast under uncertainty accounting for the economic crisis impact," in *Proc. 2009 IEEE Bucharest Power Tech Conference*, June 2009.

[9] H. M. Mahmoud, S. M. Elkhodary, S. El-Debeiky, M. Khafagy, and A. A. Twijri, "Special features of energy forecast methodology in fast growing countries," in *Proc. 12th International Middle-East Power System Conference*, 2008, pp. 555-558.

[10] Z. Ji, P, Zhang, and Z. Zhao, "Application of wavelet neutral network and rough set theory to forecast mid-long-term electric power load," in *Proc. 2009 First International Workshop on Education Technology and Computer Science*, 2009, pp. 1104-1108.

[11] C. X. Li, and L. M. Meng, "Comprehensive model for mid-long term load forecasting basing three-target quantities and RBFNN," in *Proc. Fourth International Conference on Natural Computation*, 2008, pp. 486-491.

[12] M. A. Mamun, and K. Nagasaka, "Artificial neural networks applied to long-term electricity demand forecasting," in *Proc. the Fourth International Conference on Hybrid Intelligent Systems*, 2004.

[13] Y. Shi, H. Yang, Y. Ding, and N. Pang, "Research on long term load forecasting based on improved genetic neural network," in *Proc. 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application*, 2008, pp. 80-84.

[14] A. Azadeh, S. F. Ghaderi, S. Tarverdian, and M. Saberi, "Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption," *Applied Mathematics and Computation*, vol. 186, pp. 1731–1741, 2007.

[15] Q. Ding, "Long-term load forecast using decision tree method," in *Proc. 2006 IEEE Power System Conference and Exposition*, 2006, pp. 1541-1543.

[16] D. Niu, B. Kou, and Y. Zhang, "Mid-long term load forecasting using hidden markov model," in *Proc. Third International Symposium on Intelligent Information Technology Application*, 2009, pp. 481-483.

[17] Q. Zhang, and T. Liu, "Research on the mid-long term electric load forecasting based on fuzzy rules," in *Proc. The 2nd IEEE International Conference on Information Management and Engineering*, 2010, pp. 461-463.

[18] V. N. Vapnik, "The nature of statistical learning theory," Springer, 1995.

[19] Z. Li, and M. Shen, "Analysis of dominant factors in Chinese power load growth based on fuzzy association rules," in *Proc. The 2nd International Workshop on Intelligent Systems and Applications*, May 2010, pp. 1-4.

[20] B. E. Bose, I. M. Guyon, and V. N. Vapnik. "A training algorithm for optimal margin classifiers". In *Proc. The 5th Annual ACM Workshop on Computational Learning Theory*, 1992.