THE GREY ERROR CORRECTION FORECAST METHOD BASED ON SVR

PEI-GUANG WANG[1], YANG LI[1], XIAO-PING ZONG[1], FU-FEN ZHAO[2], CHUN-XIAO YAN[2]

[1] College of Electronic and Information Engineering, Hebei University, Baoding 071002, China
[2] Dispatch Center, Power Supply Corporation of Cangzhou, Cangzhou 061000, China
E-MAIL: liyang198488@yahoo.com.cn, pgwang@mail.hbu.edu.cn

Abstract:

The advantages and disadvantages of grey forecast method are analyzed respectively. The grey error forecast method based on support vector regression (SVR) is proposed in this article. The new method remedy the disadvantages of grey forecast model and weakens the stochastic undulation, avoids the theoretical defects existing in the grey forecast model. The forecast effect is improved for non-linear specimen.

Key words:

SVM; SVR; Grey forecast; Error correction; Load forecast

1. Introduction

The foundations of Support Vector Machines (SVM) have been developed by Vapnik based on statistical learning theory. SVM is a very effective method of machine learning, which can get better effect to solve the tradition of learning algorithms that exist in small samples, non-linear, high-dimension, the local minimum point, and other practical issues, and has a strong ability of generalization. Traditional gray forecast can carry out more precise forecast result when information is incomplete. Gray forecast has a better fit and extrapolate characteristics, and less modeling data. However, projections show that the practice of using GM(1,1)model to forecast, sometimes has a good effect, and sometimes gets a greater bias, or even failure[1], because studies have shown that there are some theoretical defects in grey forecast method, in recent years[2].

2. SVR and gray forecast model

2.1. Support vector regression (SVR) algorithm

There is a set of sample, $\{ (\boldsymbol{x}_i, \boldsymbol{y}_i)\}(i=1,2,\cdots,m)$, m is sample size, $\boldsymbol{x}_i$ is input vector, $\boldsymbol{y}_i$ is the corresponding target output data. Taking into account most of the samples showed non-linear relationship, estimate function f can be determined through the following methods: each sample point is mapped to the high-dimensional feature space with a non-linear function f , and then in high-dimensional feature space for linear regression, which achieve the effect of non-linear regression in the original space. This method for the function f is

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

ω is weight vector; b is bias; $(\omega, \Phi(x))$is inner product. SVR use of structural risk minimization principle, the risk here is measured by an ε-insensitive loss function. The ε-insensitive loss function is defined as

$$L_\varepsilon(y-f(x,\omega))=\left|y-f(x,\omega)\right|_\varepsilon=\begin{cases}0 & \left|y-f(x,\omega)\right|\le\varepsilon \\ \left|y-f(x,\omega)\right|-\varepsilon & \text{other}\end{cases} \tag{2}$$

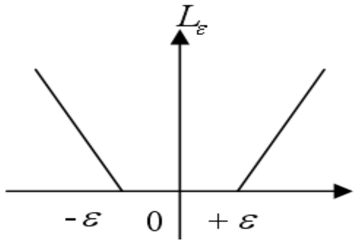

Figure 1. ε-insensitive loss function

Training for the ω, need to minimize the following functional

$$R(c,\varepsilon) = \frac{1}{m}c\sum_{i=1}^{m} L_\varepsilon(y_i - f(x_i,\omega)) + \frac{1}{2}\omega^T\omega \tag{3}$$

where c is the punishment factor.

Bring in kernel function $K(x, x_i)$, make the use of Wolfe dual skills and change the issues into the following dual problem

978-1-4244-3703-0/09/$25.00 ©2009 IEEE

$$\max_{\{\alpha_i\},\{\alpha_i^*\}}\{-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)K(x_i,x_j)$$
$$-\omega\sum_{i=1}^{m}(\alpha_i+\alpha_i^*)+\sum_{i=1}^{m}y_i(\alpha_i-\alpha_i^*)\}$$

$$s.t.\begin{cases}\sum_{i=1}^{m}(\alpha_i-\alpha_i^*)=0\\ \alpha_i,\alpha_i^*\in[0,C]\end{cases} \tag{4}$$

The regression function expressed by (1) can be written as

$$f(x,\alpha_i,\alpha_i^*)=\sum_{i=1}^{m}(\alpha_i-\alpha_i^*)K(x_i,x_j)+b \tag{5}$$

$K(x_i,x_j)$ is kernel function, such as polynomial kernel, RBF kernel and so on. Function in (5) is determined by α_i and α_i^*. According to support vector machine regression function, there are only a small number of non-zero α_i and α_i^*, these non-zero α_i and α_i^* are called support vector machine [3,4,5].

2.2. Gray forecast GM(1,1) model

Take an accumulated generating operation (AGO) to a raw data series $x^{(0)}=\{x^{(0)}(1),x^{(0)}(2),...,x^{(0)}(n)\}$, which has n variables, $x^{(1)}$ is generation series, $x^{(1)}=\{x^{(1)}(1),x^{(1)}(2),...,x^{(1)}(n)\}$,and $x^{(1)}(k)=\sum_{i=1}^{k}x^{(0)}(i)$.

Then $x^{(1)}$ is called one-time generation series of $x^{(0)}$. Because generation series $\{x^{(1)}(k)\}$ has the law of exponential growth, this series can be set up in the whitening form of differential equation

$$\frac{dx^{(1)}}{dt}+ax^{(1)}=u \tag{6}$$

a and u are identification parameters, which should be identified. a is called development parameter, which reflects the series of growth rate. u is called gray action. Then the forecast formula of $x^{(1)}$ is

$$\hat{x}^{(1)}(k+1)=[x^{(0)}(1)-u/a]e^{-ak}+u/a \tag{7}$$

3. The improvement of forecast algorithm

3.1. The defects of gray forecast

The solving process of $\hat{x}^{(1)}$ in forecast formula (7) as follows.

Solve equation (6), we have equation

$$x^{(1)}(t)=Ce^{-at}+u/a \tag{8}$$

After discretizing, equation (8) becomes

$$\hat{x}^{(1)}(k+1)=Ce^{-ak}+u/a \tag{9}$$

In order to solve constant C, a definite condition is necessary. Let $\hat{x}^{(1)}(1)=\hat{x}^{(0)}(1)=x^{(1)}(1)$ in GM(1,1) model, and $x^{(1)}(1)=x^{(0)}(1)$, we have equation

$$C=x^{(0)}(1)-u/a \tag{10}$$

Thus equation (9) becomes

$$\hat{x}^{(1)}(k+1)=[x^{(0)}(1)-u/a]e^{-ak}+u/a \tag{11}$$

Thus, the essence of GM(1,1) forecast is fitting $x^{(1)}$ with the solved $\hat{x}^{(1)}$ based on least square principle, and solving $\hat{x}^{(0)}$ by inverse accumulating $\hat{x}^{(1)}$.

Establishment of GM(1,1) model is divided into two steps. First solve parameters a,u with least square principle. Second determine the coefficient C of forecast formula, and then solve $\hat{x}^{(1)},\hat{x}^{(0)}$. The analysis indicats that it lacks strict basis of theory that GM(1,1) model forces forecast formula to thread through $\boldsymbol{x}^{(1)}(1)$. Thus the solved forecast formula is not the best one [6,7,8,9].

3.2. Grey error correction forecast method based on SVR

According to the analyses above, we propose the grey error correction forecast method based on SVR. We can suppose there are n data samples, would forecast m data samples from n+1 to n+m , the improved method is divided into three steps.

① Establish gray forecast model with known data and forecast n+m data samples from 1 to n+m with gray model.
② Use SVR to train n error values which are solved by subtracting n real values from 1 to n from n gray values from 1 to n, and forecast m error values from n+1 to n+m .
③ Correct m gray forecast values from n+1 to n+m by m

forecast error values from②.

The purpose of improving forecast algorithm is to avoid the accumulation of errors caused by GM(1,1) model forcing forecast formula to thread through $\boldsymbol{x}^{(1)}(1)$.

Improved forecast algorithm flow chart is Figure2.

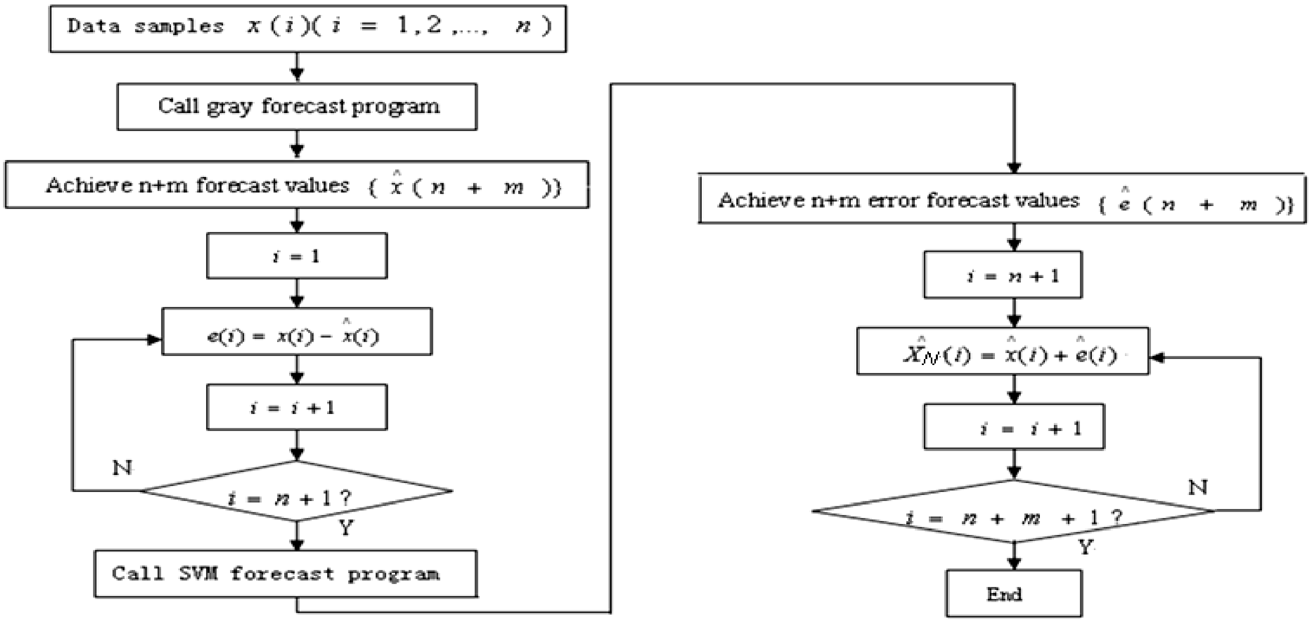

Figure 2. The flow chart of improved forecast algorithm

4. Example

Forecast data are each month power load data which come from one city of Hebei Province of December 1996 to April 1998, which are 17 in total, train the data from December 1996 to December 1997, and forecast the data from January 1998 to April 1998.

①Establish gray forecast model with the data from December1996 to December 1997, the solved identification parameters are a=-0.0212, u=74.5410.Gray forecast values and real values are shown in Figure3.

②Train error data with SVR which are solved by subtracting real values from gray values, and forecast the error data from January 1998 to April 1998.SVR forecast choose the RBF kernel function, $k(x,x_i)=\exp\{-\frac{|x-x_i|^2}{\delta^2}\}$, when ε choose 0.001, RBF kernel function parameter δ set 1.532, parameter C set 2000. In gray forecast model

$\hat{x}^{(1)}(1)=\hat{x}^{(0)}(1)=x^{(1)}(1)=x^{(0)}(1)$, so $e(1)=x^{(0)}(1)-\hat{x}^{(0)}(1)=0$, then the data of December 1996 is not trained by SVR, and 16 error data are trained by SVR. Forecast error curve is Figure4.

③Train the power load data from January 1998 to April 1998 with SVR, when ε choose 0.001, RBF kernel function parameter δ set 16, parameter C set 500.SVR forecast curve is Figure5.

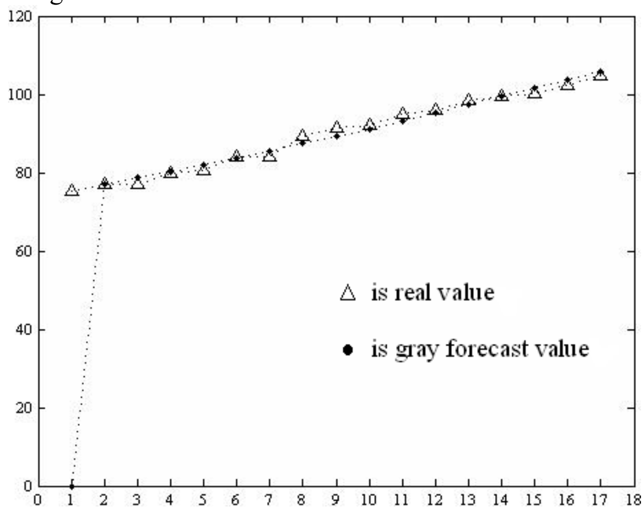

Figure 3. Gray forecast values curve and real values curve

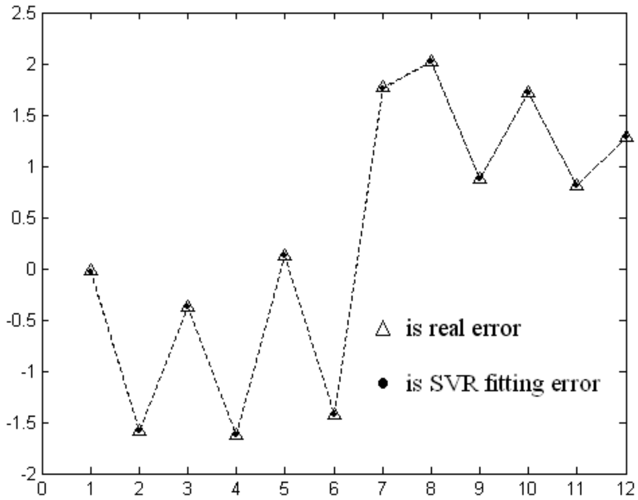

Figure 4. SVR forecast error curve and real error curve

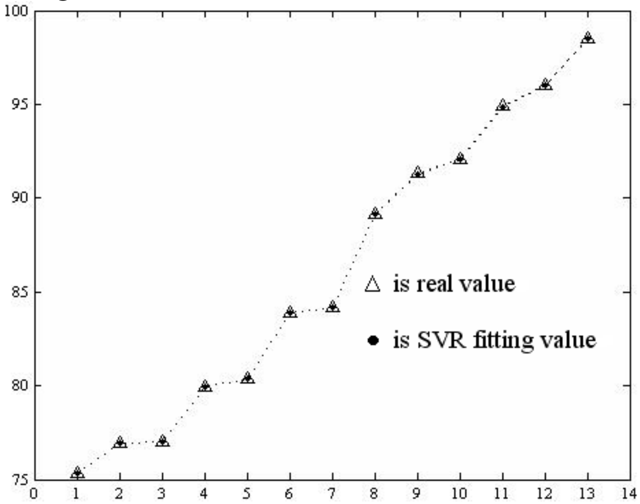

Figure 5. SVR forecast values curve and real values curve

The forecast results of improved forecast method, SVR and GM(1,1)Model are shown in Figure6 and Table1.

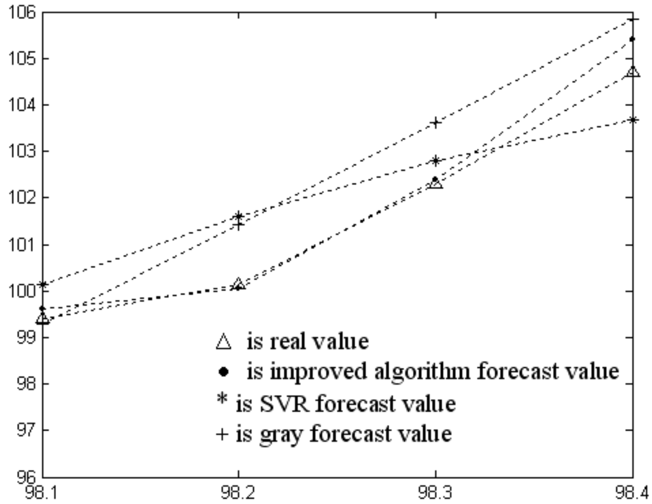

Figure6. Improved forecast method, SVR, GM(1,1)Model, forecast results curve

Table 1. Improved forecast method，SVR, GM(1,1)Model forecast results

date	Real data（kw）	Improved algorithm forecast data(kw)	SVR forecast data(kw)	Gray forecast data(kw)	Improved algorithm forecast relative error (%)	SVR forecast relative error(%)	Gray forecast relative error(%)
January 1998	99.42	99.61	100.15	99.30	-0.1911	-0.7343	0.1207
February 1998	100.15	100.04	101.60	101.43	0.1098	-1.4478	-1.2781
March 1998	102.30	102.38	102.79	103.61	-0.0782	-0.4790	-1.2805
April 1998	104.69	105.39	103.67	105.83	-0.6686	0.9743	-1.0889
Absolute average error(%)					0.2619	0.9089	0.9421

5. Conclusion

Examples show that improved forecast algorithm is better than simply using the gray model or the SVR. Because improved forecast algorithm avoids the accumulation of errors caused by simply using the gray model, improve the forecast accuracy. The improved forecast algorithm used to forecast the medium-term load forecast with more relevant factors is feasible and effective.

Reference

[1] Xiaojun Tong, Jinyun Chen, “Gray Logistic Model Based on Grade Difference Format”, Control and Decision, Vol 17, No. 5, pp. 554-558, 2002.

[2] Wanmei Tang, “New Forecasting Model Based on Grey Support Vector Machine”, Journal of Systems Engineering, Vol 21, No. 4, pp. 411-413, 2006.

[3] Feng Pan, Haozhong Cheng, Jingfei Yang, Cheng Zhang, Zhendong Pan, “Power System Short-Term Load Forecast”, Power System Technology,No. 28, pp. 39-40, 2004.

[4] Ruiming Fang, Theory and Applied Analysis of Support Vector Machine, China Electric Power Press, Beijing, 2007.

[5] Jian Zhang, Power System Load Models and Identification, China Electric Power Press, Beijing, 2007.

[6] Dahai Zhang, Shifang Jiang, Kaiquan Shi, “Theoretical Defect of Grey Prediction Formula and Its Improvement”, Systems Engineering —Theory & Practice, Vol 22, No. 8, pp. 140-142, 2000.

[7] Hai He, Jinyun Chen, “Prediction Formula’s Defect of GM (1, 1) and Its Improvement”, Journal of WUHAN University of Technology, Vol 26, No. 7, pp. 81-82, 2004.

[8] Peirong Ji, Weisong Huang, Xiangyong Hu, “A Study on the Properties of Grey Forecasting Model”, Systems Engineering —Theory & Practice, No.9, pp. 105-106, 2001.

[9] Dongxiao Niu, Shuhua Cao, Lei Zhao, Wenwen Zhang, Power Load Forecasting Technology and Its Applications, China Electric Power Press, Beijing, 1998.