Journal of Frontiers in Engineering Technology
Clausius Scientific Press, Canada

DOI: 10.23977/jfet.2021.010108
Volume 1, Number 1, 2021

Ultra Short Term Load Forecasting Based on Optimized Weight Cubature Kalman Filter and Support Vector Machine Combination Model

Bei Huang[1,2], Weidong Kang [1], Hao Gu[1*],Boyang Zhou [1,2], Shi Chen[1,2]

1. Anhui Nanrui Jiyuan Power Grid Technology Co., Ltd., China

2. Anhui Huahong Information Technology Co., Ltd., China

**Corresponding Author*

Keywords: Cubature kalman filter, Least squares support vector machine, Grey neural network, Combinatorial model

Abstract: In this paper, a combined ultra short term load forecasting model is proposed to solve the problem of less feature dimension and unclear relationship in ultra short term load forecasting for industrial power users. The model combines the cubature Kalman filter (CKF) prediction method which is better in nonlinear dynamic system and the least squares support vector machine (LS-SVM) prediction method which is better in small-scale data prediction. It combines the advantages of the two algorithms by using the combination of grey neural network, and avoids a single algorithm falling into local optimum. It combines horizontal prediction with vertical prediction Finally, the combination model is better than the single prediction.

1. Introduction

Load forecasting is the premise of economic dispatch, real-time control, operation planning and development planning of power system. It provides important data for the planning and operation of power system, and is also one of the symbols of modern management[1-4]. According to the length of forecasting time, load forecasting can be divided into ultra short term load forecasting, short-term load forecasting, medium-term load forecasting and long-term load forecasting. The ultra short term load forecasting mainly refers to the forecasting of the load value in the next 10 ~ 15 minutes, so as to realize the real-time control of the system. In the industrial field, the industrial load is different from other power loads, which is less affected by temperature, population, income, holidays and other aspects, but is mainly determined by the production plan of industrial power users. With the gradual introduction of competition mechanism into the power market, the accuracy of load forecasting will directly affect the power trading strategy of industrial users, and further affect the economic benefits of power users. Therefore, ultra short term load forecasting for industrial power users is of great significance.

Particle swarm optimization algorithm and traditional method have become a common method. This method can improve the prediction accuracy by optimizing parameters. In literature [5], a

method of using chaos optimization particle swarm least squares support vector machine (LSSVM) is proposed. The method adopts the global search ability of particle swarm optimization and the traversal characteristics of chaos algorithm, and improves the parameter selection method to improve the selection of parameters And the local optimal situation of particle swarm optimization is easy to appear, but the problem of the method is not solved effectively in the global search, such as the selection of iteration times and the determination of initial value. In this paper, a limit learning machine prediction model is proposed to improve particle swarm optimization. The optimal input weight and hidden layer deviation are obtained by PSO algorithm, so as to reduce the random parameter error. In addition, chaos adaptive strategy is introduced to strengthen the diversity of PSO, and prevent the occurrence of local convergence too quickly. This model can reduce the random parameter error The method only improves the parameter selection without optimizing the algorithm structure itself. The single algorithm will still have local optimal situation for the load forecasting in complex situations. Kalman filter is usually used in navigation, and its strong real-time and adaptive ability can also play a great role in load forecasting. Therefore, in recent years, Kalman filter algorithm is more and more used in the field of load forecasting. The results of short-term load forecasting method based on the modified factor are better than the traditional Kalman filter algorithm, and the convergence speed and time are shorter. However, for the task of ultra short-term load forecasting, the first load model is a nonlinear model, and Kalman filter is more used in linear environment, so that the parameters above the second order term are discarded for the most The final prediction results have a great influence; secondly, Kalman filter has a certain lag, so in the short-term load forecasting, the load jump is very large, which will lead to a large error in the prediction results. Emre Akarslan [6-9] proposes a new short-term load forecasting method based on adaptive neural fuzzy inference system (ANFIS). This method is to use hourly load measurement to record the forecast model of future ANFIS. The model only needs to calculate the range of consumption parameters. In this range, the first derivative of load quantity, date and hour parameter can be input into ANFIS model, and then the following can be obtained This method needs a lot of historical data to analyze the load value for an hour. However, in the ultra short-term load, it is necessary to consider that a large amount of data can not be used as training data. If the production plan is not regular, it will lead to the interference of different production plans and increase the error. Aleksandr s. gritsay[10] proposes a short-term forecasting method of power load combining sine function and artificial neural network. The method is convenient and fast, and is suitable for load forecasting with less load jump. Sumit Kumar [11] proposed a load forecasting method based on the improved long and short-term memory network (LSTM) deep learning method based on the gate control cycle unit (Gru). The method uses LSTM and Gru to complete the load forecasting by adjusting the parameters, and finally the minimum average can be obtained The error parameters are set effectively by this method, and the calculation speed is fast, and the prediction results are good.

In this paper, the least squares support vector machine prediction model and cubature Kalman filter prediction model are used to predict the industrial power load. In view of the problem that the accuracy of the two algorithms is not high, this paper proposes to use the gray neural network to combine the two algorithms, so that their advantages complement each other. The example results show that the combination algorithm can effectively improve the prediction accuracy.

2. Prediction Model of Least Squares Support Vector Machine

Least squares support vector machine (LSSVM) is one of the support vector machines. The model contains fewer parameters to be selected, so it does not need a lot of training data to adjust the parameters in the prediction. At the same time, it uses equality constraints instead of inequality

constraints to further reduce the uncertainty of the model.

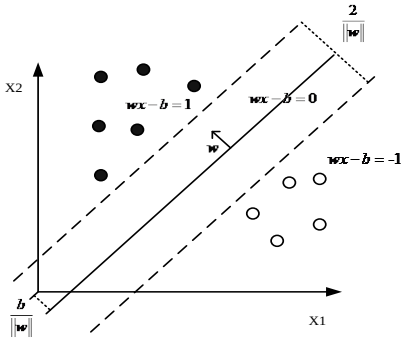

Fig.1 Classification Plane of Support Vector Machine

Suppose the training data set $\{(x_1, y_1), (x_2, y_2).....(x_n, y_n)\}$, which represents the input data, represents the output data, and represents the number of data. The basic idea of least squares support vector machine regression theory is to find a nonlinear function to complete the mapping from input data to output data, and then map the input data to a high-dimensional space in which to complete the linear regression of the estimated function. The flow chart of support vector machine algorithm is shown in Figure 1

Kernel function is a non-linear mapping to project data into high-dimensional feature space. In this paper, RBF kernel function is used. Because the nonlinear equations are transformed into linear equations for solving, the calculation amount of LS-SVM will be far less than that of SVM, and the calculation speed will be improved to a certain extent, but the error in accuracy is slightly greater than that of SVM, and the data processing ability is weak. This paper selects other algorithms combined with LS-SVM to complete the prediction work.

3. Prediction Model of Cubature Kalman Filter

In this paper, the time series analysis method is used to establish the system state equation and measurement equation.

3.1 Time Series Analysis Model

Time series analysis models mainly include three models: Auto Regressive Model (AR), moving average model (MA) and auto regressive moving average model (ARMA). Among them, the autoregressive moving average model is widely used:

$$X_t - \phi_1 X_{t-1} - ..... - \phi_n X_{t-n} = a_t - \theta_1 a_{t-1} - .... - \theta_m a_{t-m} \tag{1}$$

In a system, if its response at a certain time t is Xt, not only the response is related to its own value at the previous time, but also to the disturbance entering the system at the previous time, then the system is the autoregressive moving average model.

Time series analysis model is divided into traditional time series analysis and modern time series analysis. The traditional time series analysis also includes box and Jenkins method and Wu Mingxian and Pandit method. The main difference between the two methods is that the former uses guessing method and the latter uses F-test method in model order determination; the former uses ML method and the latter uses nonlinear LS method in parameter estimation. The problem of the two methods lies in the large amount of calculation. In modern time series analysis, F-test is also used in model order determination, but linear recursive least squares (RLS) method is used in parameter estimation. The advantage of RLS method is that adaptive process is added in the

operation process, and the amount of calculation is small.

3.2 Cubature Kalman Filter Model

Cubature Kalman filters (CKF) is a nonlinear Gaussian filtering method based on the third-order integral principle. The equation of state and measurement equation can be set in the following forms:

$$\begin{cases} X_k = F(X_{k-1}) + w_{k-1} \\ Z_k = H(X_k) + v_k \end{cases} \tag{2}$$

Where x is the state quantity, Z is the measurement, F and H are the nonlinear functions, W is the process noise, and V is the measurement noise. The steps of CKF algorithm are as follows:

(1) Time update $P_{k-1|k-1}$ $S_{k-1|k-1}$

1) The Cholesky decomposition of yields , i.e

z2)Volume point estimation $x^l_{k-1|k-1}$

$$x^l_{k-1|k-1} = S_{k-1|k-1}\xi_l + \hat{x}_{k-1|k-1} \tag{3}$$

Among them: $\xi_t = \sqrt{\frac{2}{L}}[\delta]_{l,}\ l = 1, 2.....L, [\delta]_l \in R^{m\times 1}$

Representation matrix $\left[I^{m\times m}, -I^{m\times m}\right] \in R^{m\times L}$

Column l element

$I^{m\times m}$ is the m-dimensional identity matrix

3)Calculation volume point propagation $x^{*,l}_{k|k-1}$

$$x^{*,l}_{k|k-1} = f_k(x^l_{k-1|k-1}) \tag{4}$$

4)One step prediction of state $\hat{x}_{k|k-1}$

$$\hat{x}_{k|k-1} = \sum_{i=1}^{L} x_{k|k-1} / L \tag{5}$$

5)The covariance matrix of state one-step prediction error is calculated $P_{k|k-1}$

$$P_{k|k-1} = \sum_{l=1}^{L} x^{*,l}_{k|k-1}(x^{*,l}_{k|k-1}) / L - \hat{x}_{k|k-1}(\hat{x}_{k|k-1})^T + \Gamma_{k-1}Q_{k-1}\Gamma^T_{k-1} \tag{6}$$

(2)Measurement update

1) Cholesky decomposition of $P_{k|k-1}$

$$P_{k|k-1} = S_{k|k-1}(S_{k|k-1})^T \tag{7}$$

2)Volume point estimation

$$x^l_{k|k-1} = S_{k|k-1}\xi_l + \hat{x}_{k|k-1} \tag{8}$$

3)Volume point propagation

$$z^l_{k|k-1} = h(x^l_{k|k-1}) \tag{9}$$

4)Calculate the one-step predicted value of measurement

$$\hat{Z}_{k|k-1} = h(x^{j}_{k|k-1}) / L \tag{10}$$

5)Calculate the variance $P^{zz}_{k|k-1}$ and the cross covariance matrix $C^{zz}_{k|k-1}$ of state and measurement:

$$v_k = z_k - z_{k|k-1} \tag{11}$$

$$S_k = P^{zz}_{k|k-1} = \sum_{l=1}^{L} z^{l}_{k|k-1}(z^{l}_{k|k-1})^T / L - \hat{z}_{k|k-1}(\hat{z}_{k|k-1})^T + R_k \tag{12}$$

$$C^{zz}_{k|k-1} = \sum_{l=1}^{L} x^{l}_{k|k-1}(z^{l}_{k|k-1})^T / L - \hat{x}_{k|k-1}(\hat{z}_{k|k-1})^T \tag{13}$$

6)Calculate the filter gain at k time K_k

$$K_k = C^{zz}_{k|k-1}(P^{zz}_{k|k-1})^{-1} \tag{14}$$

7)State estimation at computing time $\hat{x}_{k|k}$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K(z_k - \hat{z}_{k|k-1}) \tag{15}$$

8)K-time state estimation error covariance matrix $P_{k|k}$

$$P_{k|k} = P_{k|k-1} - K_k P^{zz}_{k|k-1}(K_k)^T \tag{16}$$

4. Grey Neural Network Model

The grey problem refers to the problem of predicting the development and change of the behavior eigenvalues of the grey uncertain system. After the accumulation of the original sequence, a new sequence can be obtained. The sequence shows only the growth law, so a continuous function or differential equation can be used for data fitting. Then this process is the working process of the grey neural network. The structure of the grey upgrading network is shown in Figure 2 It is shown that.

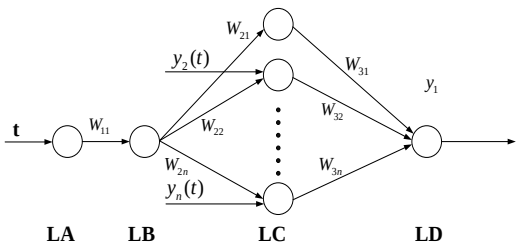

Fig.2 Topological Structure of Grey Neural Network

In the figure, $\omega_{11}, \omega_{21}, \omega_{22}, \omega_{31}.....$ represents the weights of neural network, $y_2(t), y_3(t), ...y_n(t)$ represents the input parameters, LA, LB, LC, LD represents the four layer grey neural network, and y_1 represents the predicted value.

Next, let's order:

$$\frac{2b_1}{a} = u_1, \frac{2b_2}{a} = u_2, ..., \frac{2b_{n-1}}{a} = u_{n-1}$$

Then the initial weights of the network can be expressed as:

$\omega_{11} = a, \omega_{21} = -y_1(0), \omega_{22} = u_1 .....\omega_{2n} = u_{n-1}$

$\omega_{31} = \omega_{32} = ... = \omega_{3n} = 1 + e^{at}$

The output node threshold of LD layer can be expressed as:

$\theta = (1 - e^{at})(d - y_1(0))$

The learning process of grey neural network includes five parts

1) According to the characteristics of training data, the network structure and parameters a, B are initialized and μ is calculated;

2) Calculate network weight $\omega_{11}, \omega_{21}, \omega_{22}, ...\omega_{2n}, \omega_{31}, \omega_{32}, ...\omega_{3n}$;

3) Input parameters and calculate each layer of neural network;

La layer: the calculation of parameter a can be expressed as

$$a = \omega_{11} t \tag{17}$$

LB layer: the calculation of parameter B can be expressed as

$$b = f(\omega_{11} t) = \frac{1}{1 + e^{w_{11} t}} \tag{18}$$

LC layer: the calculation of parameter C can be expressed as

$$c_1 = b\omega_{21}, c_2 = y_2(t) b\omega_{22}, ...c_n = y_n(t) b w_{2n} \tag{19}$$

LD layer: the calculation of parameter D can be expressed as

$$d = \omega_{31} c_1 + \omega_{32} c_2 + ... + \omega_{3n} c_n - \theta_{y1} \tag{20}$$

4) The prediction error is calculated and the weight is adjusted according to the error;

5) Judge whether the training is over, if not, return to the third step;

5. Examples and Analysis

The data source of this project comes from the load data collected by industrial users in a city with a frequency of 15 minutes a week, a total of 672 groups of data. The least squares support vector machine prediction model, cubature Kalman filter prediction model and combination prediction model will be used to predict the group of data respectively, and the prediction error results are compared, and the experimental conclusion is drawn. The average absolute error (m) is selected as the error index Ean absolute deviation, also known as mean absolute deviation, is the average of the absolute values of the deviations of all single observations from the arithmetic mean. The average absolute error can avoid the problem that the errors cancel each other, so it can accurately reflect the actual prediction error. It can be expressed as follows:

$$marerr = \frac{1}{N} \sum_{i=1}^{n} \left(\frac{|x_i - X_i|}{X_i} \right) \times 100\% \tag{21}$$

Where x_i represents the predicted value, , X_i represents the actual value,and N represents the amount of data.

Figure 3 shows the prediction results of least squares support vector machine. This paper uses the data of the first six days as the training data, and the data of the last day as the test data, and uses the test results and the actual results to calculate the average absolute error. The final error rate is 14.46%.

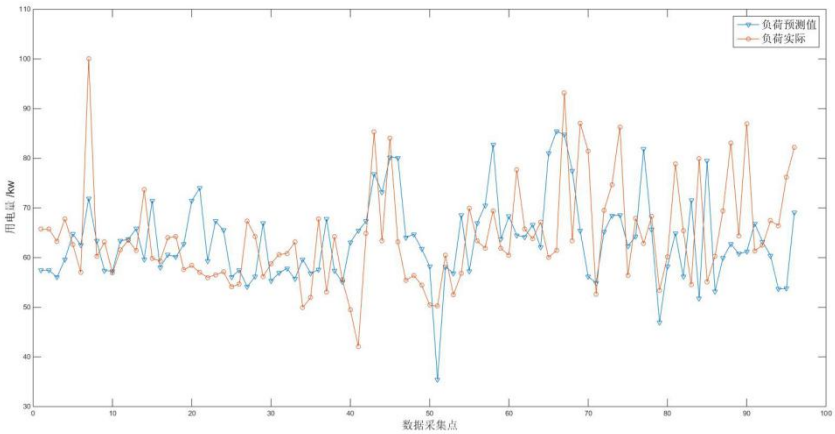

Fig.3 Prediction Results of Least Squares Support Vector Machine Model

As can be seen from Figure 3, according to the prediction results of LS-SVM model, the amplitude of fluctuation value is relatively large at some data acquisition points, and the prediction results of some data points that originally changed gently, such as the acquisition point near No. 20, have different degrees of jumping.

According to the tailing of autocorrelation coefficient and the third-order censoring of partial autocorrelation coefficient, the model is ARMA (3,0) model. Then the least square regression is used to estimate the parameters

$$X(k) = 0.1952 * X(k-1) + 0.1462 * X(k-2) + 0.0151 * X(k-3) + 40.7831 \quad (22)$$

The measurement equation is as follows:

$$Z(k) = X(k)^2 \quad (30)$$

Figure 6 shows the prediction result of cubature Kalman filter. The prediction result of the last day is also compared with the actual load data and the average absolute error is calculated. The final error rate is 13.06%.

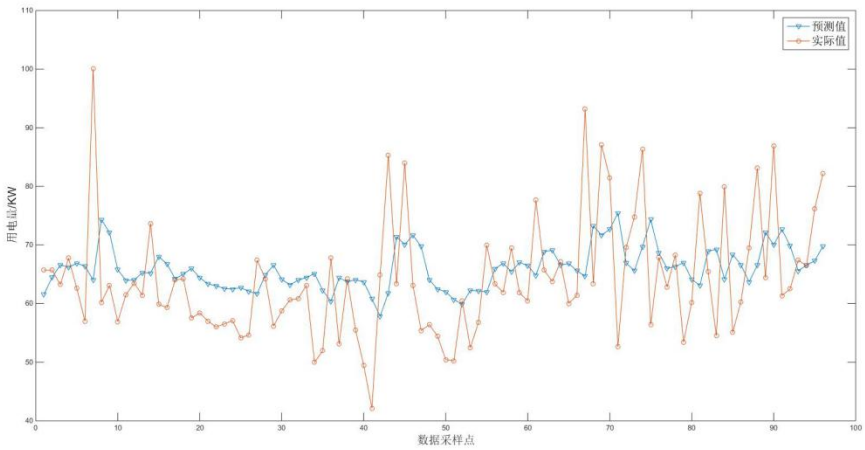

Fig.4 Prediction Results of Cubature Kalman Filter Model

It can be seen from Figure 4 that according to the prediction results of the cubature Kalman filter model, the fluctuation amplitude is small, and the deviation is large near the acquisition points with large jump ratio. Moreover, due to the adaptive characteristics of the Kalman filter, the prediction

value will be affected by the first three points. If there are jumps in these points, the prediction accuracy will be affected.

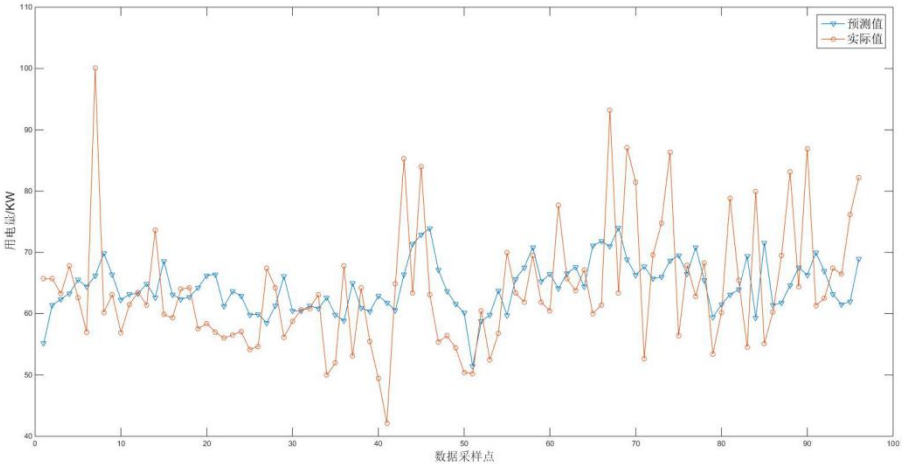

Fig.5 Prediction Results of Combined Model

As can be seen from Figure 5, according to the prediction results of the combined prediction model, the change amplitude is between the least squares support vector machine prediction model and the cubature Kalman filter prediction model, the prediction results are better than the former two, and the existing problems are also improved to a certain extent.

6. Conclusion

In this paper, the combination algorithm of least squares support vector machine prediction algorithm and cubature Kalman filter prediction algorithm is realized to predict the ultra short term load data of industrial users, and the final prediction error is 11.31%. Compared with the two algorithms, the error results of load forecasting are 14.46% and 13.06%, which are improved to a certain extent. The problem that a single algorithm is easy to fall into local optimum is effectively improved. The two algorithms are combined by grey neural network to avoid the problems of complex calculation and low efficiency existing in traditional linear weighted sum method. The example shows that the combined forecasting model is suitable for ultra short term prediction of industrial users. However, there is still room for improvement in the model. The combined model is not sensitive to load data jumping points, so the identification and prediction of jumping points need to be further improved.

References

[1] Chen Jun, Tao Wei, LV Yingfei, he Jianping. Short term load forecasting based on Kalman filter [J]. Electrical switch, 2014 (2): 43-45

[2] Chen Heng, Wang Ke, Chen Lihua. Research on holiday short term load forecasting based on filtering algorithm [J]. Electrical technology, 2014 (9): 12-15

[3] Li minggan, sun Jianli, Liu Pei. Short term load forecasting of power system based on Kalman filter [J]. Relay, 2004 (4): 9-12

[4] I.Arasaratnam, S.Haykin. Cubature Kalman Filters[J]. IEEE Trans. Automat.Control, 2009, 54(6):1254-1269.

[5] Y.X.Wu,D.W.Hu,M.P.Wu,X.P.Hu. Quasi-Gaussian Particle Filtering[J].Comput. Sci., 2006, (3991):689-696.

[6] Li Yanhong, Lei Jinhui. Application and research of power load time series forecasting [J]. Science, technology and engineering, 2011 (04): 860-864

[7] Ji Xiaojun, Yang Juan. Application of neural network based on wavelet transform in short-term load forecasting of

Shaanxi Power Grid [J]. Shaanxi electric power, 2013 (10): 66-70

[8] Zhao Zhiyong, Huang Wei, Yang Wei. A medium and long term power load forecasting method based on wavelet analysis and Elman dynamic neural network [J]. Shanxi electric power, 2013 (01): 1-5

[9] Li Xiao, Wang Xin, Zheng Yihui, Li Lixue, Sheng Xikui, Wu Hao. Short term wind power load forecasting based on improved least squares support vector machine and prediction error correction [J]. Power system protection and control. 2015 (11): 66-70

[10] HUANG F J,LECUN Y.Large-scale learning with SVM and convolutional for generic object categorization [C]. IEEE Computer Society Conference on Computer Vision and Pattern

[11] Recognition,2006: 284-291.

[12] Dong Hao, Cheng Peng, Li Lingling. Application of deep learning algorithm in short-term load forecasting of power system [J]. Electrical times, 2017 (02): 82-84