

Load forecasting considering multiple influencing factors

Xin Ning^{1,2*}, Liang Jin¹

¹School of electrical engineering and automation, Tiangong University, Tianjin, 300387, China

²State Grid Tianjin electric power company, Tianjin, 300000, China

*727362732@qq.com

Abstract. With the improvement of power conservation awareness and the development of the power market, load forecasting is playing an important role gradually. Effective load forecasting can help power system operators to develop appropriate scheduling strategies and help users plan their power consumption rationally. Considering the important position of load forecasting in the future power system field, this paper focus on this field. Based on the measured data from Australia, this paper considers the environmental and social factors affecting the power consumption in the region. The LSSVM algorithm is used for short-term load forecasting. Firstly, the invalid data is eliminated by data preprocessing, the missing data is completed, and the data of the content is included. The normalization process is carried out, and then the processed data is used for prediction. Finally, the effectiveness of the Least squares support vector machine algorithm load prediction after considering various factors is verified by comparison with the traditional neural network algorithm.

1. Introduction

As the awareness of environmental protection and energy conservation increases, smart electricity strategies are beginning to emerge. Intelligent power consumption strategy can effectively help users and grid companies to make reasonable decisions, thus effectively achieving power savings. While accurate load forecasting is the basis for implementing smart power strategy, so it is necessary to improve the accuracy of load forecasting. Generally speaking, regional load changes do not have specific laws, but their changes are closely related to external factors. External environmental factors such as temperature and humidity have a greater impact on the load. Load forecasting is to achieve the purpose of forecasting by looking for the relationship between these factors and load consumption.

Load forecasting is generally conducted based on historical data through intelligent algorithms or time series methods. Commonly used intelligent algorithms include: neural network algorithm [1], random forest algorithm [2], and Bayesian algorithm [3]. Time series method regards time series data as random variables, using averaging or weighted average to get predictions. The artificial neural network method is used to predict the load of residential electricity energy consumption, and a better prediction performance is achieved in [4]. An online energy prediction platform is proposed in [5], which uses intelligent algorithms such as artificial neural networks to predict the electrical energy of buildings. Combining the particle swarm optimization algorithm to determine the number of decision trees, the improved prediction method based on random forest is realized in [6]. A short-term load forecasting model based on clustering analysis and Bayesian algorithm is proposed in [7] and the accuracy of short-term load forecasting is improved. Time series analysis refers to the use of historical data as input to predict, Based on the time series model, [8] achieves the prediction of wind speed. But



the time series model can only predict the linear relationship between input and output, it is difficult to predict when there is a nonlinear relationship between input and output, so there is a limit when predicting loads.

The support vector machine (SVM) was first proposed by Vapnik [9], which mainly solves the problem of pattern classification and nonlinear regression in practical applications. The least squares support vector machine (LSSVM) changes the constraint of the support vector machine to the equality constraint, which simplifies the solution process and improves the prediction accuracy compared with the basic vector machine. The LSSVM model can effectively deal with the nonlinear relationship existing in the prediction process. At the same time, this method can effectively overcome the long training time and reduce the randomness in the training results, so it has received extensive attention in the prediction field.

Considering the advantages of least squares support vector machine in the accuracy and efficiency of prediction, this paper based on the actual data, through the least squares support vector machine algorithm, and fully consider the factors that fully affect the temperature, humidity and electricity price of load changes for Australia. The load prediction is carried out in a certain place. The simulation results show that the load forecasting method based on least squares support vector machine considering multiple factors can achieve better prediction results.

2. LSSVM method and data pre-processing

2.1. LSSVM theory

Here the original sampling is (x_i, y_i) , $i=1, 2, \dots, n$. The principle of SVM model is to map the samples into the feature space through a nonlinear function $\varphi(x)$ and the mapped results in feature space can be expressed as follows:

$$\Psi(x) = (\varphi(x_1), \varphi(x_2), \dots, \varphi(x_n)) \quad (1)$$

where $\Psi(x)$ is the vector of mapped results.

In the feature space, the optimal decision function is defined as follows:

$$y = \omega^T \cdot \Psi(x) + b \quad (2)$$

where the ω and b is the parameters need to be determined. The unknown parameters are always determined by the structure risk minimum principle. Here the structure risk can be calculated by loss function and regularization parameter. The structure risk can be calculated as follows:

$$R = c \cdot R_{emp} + \frac{1}{2} \|\omega\|^2 \quad (3)$$

where c is regularization parameter and R_{emp} is the loss function. According to the different loss function, the SVM can be divided as different model. The LSSVM uses the square loss function and the square loss function is illustrated as follows:

$$R_{emp} = \sum_{i=1}^n \xi_i^2 \quad (4)$$

where ξ_i is the prediction error.

The process of calculating the unknown parameters can be expressed as:

$$\min R = c \sum_{i=1}^n \xi_i^2 + \frac{1}{2} \|\omega\|^2 \quad (5)$$

$$\text{s.t. } y_i = \omega^T \varphi(x_i) + b, i = 1, 2, \dots, n$$

Using Lagrange multiplier method, the aforementioned process can be illustrated as:

$$L(\omega, b, \xi_i, \alpha) = c \sum_{i=1}^n \xi_i^2 + \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^n (\alpha_i (\omega^T \varphi(x_i) + b - y_i)) \quad (6)$$

where α_i is the Lagrange multiplier. According to the optimization condition, the optimal solution occurs when (7) is solved.

$$\frac{\partial L}{\partial \omega} = 0, \frac{\partial L}{\partial b} = 0, \frac{\partial L}{\partial \xi_i} = 0, \frac{\partial L}{\partial \alpha_i} = 0 \quad (7)$$

After (7) is solved the unknown parameter and optimal decision function can be calculated as:

$$\begin{aligned} \omega &= \sum_{k=1}^n \alpha_k \varphi(x_k) \\ \text{s.t. } \sum_{k=1}^n \alpha_k &= 0 \\ 2c\xi_i &= \alpha_i \\ y_i &= \omega^T \varphi(x_i) + b + \xi_i \end{aligned} \quad (8)$$

Furtherly the optimal decision function can be expressed as:

$$y_i = \sum_{j=1}^n (\alpha^j (\varphi(x_j), \varphi(x_i))) + b + \frac{1}{2c} \alpha_i \quad (9)$$

Define the kernel function as $K(x_i, x_j) = (\varphi(x_j), \varphi(x_i))$, the optimal decision function can be expressed as:

$$y_i = \sum_{j=1}^n (\alpha^j K(x_i, x_j)) + b + \frac{1}{2c} \alpha_i \quad (10)$$

The common used kernel function is the radial basis function and the radial basis function is given by (11).

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (11)$$

where σ is the width parameter.

Considering the constraints in (8), the unknown parameter and Lagrange multipliers can be calculated as follows.

$$Y \cdot \begin{bmatrix} b \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (12)$$

where Y can be calculated as follows.

$$Y = \begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & K(x_1, x_1 + \frac{1}{2c}) & K(x_1, x_2) & \dots & K(x_1, x_n) \\ 1 & K(x_2, x_1) & K(x_2, x_2 + \frac{1}{2c}) & \dots & K(x_2, x_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & K(x_n, x_1) & K(x_n, x_2) & \dots & K(x_n, x_n + \frac{1}{2c}) \end{bmatrix} \quad (13)$$

According to the (12) and (13), the optimal decision function can be determined and the optimal function can be used to do the prediction.

2.2. Data pre-processing

Since the data adopted in this paper includes load history data of a certain area and information such as meteorology and electricity price, in order to ensure the accuracy of the prediction result, the data needs to be preprocessed, including data filling and elimination, and normalization [10].

Due to the lack of data in the data collection process, considering the smoothness of the load curve, the data filling is achieved by the existing data. The data filling is given as follows.

$$y_i = w_1 y_{i-1} + w_2 y_{i+1} \quad (14)$$

where y_i is the missing data; y_{i-1} and y_{i+1} are the earlier data point and latter data point; w_1 and w_2 is the weights and the weights usually equal to 0.5.

The purpose of data elimination is to eliminate the abnormal data point. Here a limit is defined to determine whether the data point is the abnormal data. Considering the data we used in this paper is selected in a cycle of 30 minutes the load cannot be fluctuated in a large range so the value of earlier data point is defined as the threshold in this paper. That is the difference between the former data point and present data point should satisfy (15) if the present data point is right data.

$$|y_i - y_{i-1}| < y_{i-1} \quad (15)$$

Like the data filling, after the abnormal point is culling the culling data point need to be replaced by the approximated value.

Since the collected data contains different types, and the numerical differences are very large, so normalization is adopted to get an accurate result. Data not located in the range of [0,1] such as temperature, humidity, and electricity price are converted in a linear normalized form to make sure the converted value is located in [0,1].

$$e_i = \frac{x_i - x_{\min}}{x_i - x_{\max}} \quad (16)$$

where x_i is the original outside data such as the temperature and humidity; x_{\min} and x_{\max} is the minimum and maximum value of these outside data; e_i is the normalization value.

3. Overall Process of load forecasting

Based on the above theory, this paper analyses the load information of regional power grid based on data, studies its change trend, and uses the measured data to consider the multiple complex factors affecting the load of power system to effectively predict the load data of a certain area. The LSSVM load prediction steps considering multiple influencing factors are as follows:

Step 1: Data pre-processing, including padding and culling of data, and normalization of data.

Step 2: Train the LSSVM with a large amount of historical data and verify the validity of the model to ensure the accuracy and reliability of the predicted values.

Step 3: Predict the test set data and compare the predicted results with other algorithm results.

4. Case study

This paper uses LSSVM and BPNN to predict the load of a region separately. The data used in this paper is Australian electricity load and price forecast data, which can be used for electricity price forecasting and load forecasting, with an interval of half an hour, including data from January 1, 2006 to January 1, 2011. The variables in the data set that are closely related to the load are dry bulb temperature x_1 , dew point temperature x_2 , wet bulb temperature x_3 , humidity x_4 and electricity price x_5 . Part of the original data content is shown in the Table 1.

Table 1. Part of the data content in December 31, 2010

Time	Hour	Dry bulb temperature	Dew point temperature	Wet bulb temperature	Humidity	Electricity price	Load
2010/12/31	0.5	22.1	20.15	20.85	88.5	20.7	7384.33
2010/12/31	1	22	20.1	20.8	89	18.76	7161.66
2010/12/31	1.5	21.95	20	20.7	88.5	19.94	6842.79
2010/12/31	2	21.9	19.9	20.6	88	19.85	6594.82
2010/12/31	2.5	21.75	19.9	20.55	89	19.16	6407.01
2010/12/31	3	21.6	19.9	20.5	90	18.17	6317.23
2010/12/31	3.5	21.5	19.95	20.5	91	17.74	6287.11
2010/12/31	4	21.4	20	20.5	92	17	6295.38

We choose two weeks data as train data which is from December 1, 2010 to December 14, 2010 to predict the load of December 15, 2010. We measure the accuracy of the algorithm by three indicators. They are Root Mean Square Error (*RMSE*), Mean Absolute Error (*MAE*), and Coefficient of Determination (*CD*). Their calculation formulas are as follows:

$$MAE(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (17)$$

$$RMSE(\mathbf{y}, \hat{\mathbf{y}}) = \left[\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 \right]^{\frac{1}{2}} \quad (18)$$

$$R^2(\mathbf{y}, \hat{\mathbf{y}}) = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2} \quad (19)$$

where \mathbf{y} is the actual value, $\hat{\mathbf{y}}$ is the predicted value, and \bar{y} is the average of actual values. The smaller the *RMSE* and the *MAE*, the better the prediction performance is, and the larger the *CD* is, the better the prediction performance is. The final result is shown below.

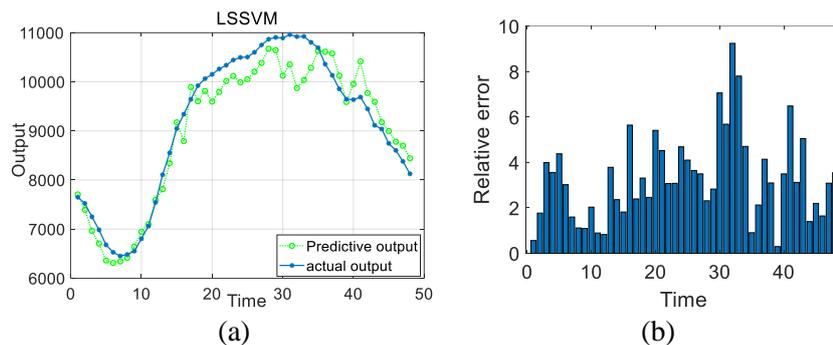


Figure 1. (a) The predictive result of LSSVM (b) The absolute error of prediction

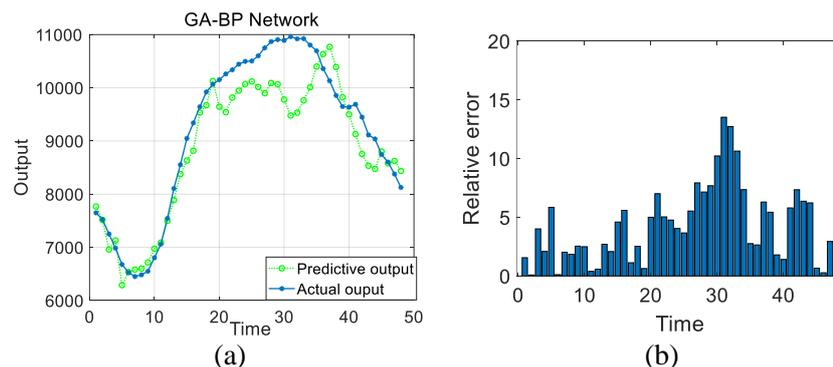


Figure 2. (a) The predictive result GA-BP (b) The absolute error of prediction

Table 2. Part of the data content in December 31, 2010

Index	SVM	GA-BP
<i>CD</i>	0.9393	0.8470
<i>MAE</i>	0.2644	0.3861
<i>RMSE</i>	0.4200	0.5859

It can be seen from the Figure 1 that the proposed method can precisely predict the load value. The relative error is small. Then we use BPNN to predict the same data set. The BPNN is optimized by Genetic algorithm. The performance indicators of the algorithm are shown in Table.2. From Table 2, we can see that LSSVM is superior to BPNN and can better track load changes. In terms of running time, the running time of the LSSVM is 0.904197s, and the BPNN is 1.504197s. So, compared with BPNN, LSSVM is more suitable for load forecasting.

5. Conclusion

Short-term load forecasting of power systems is of great significance to the reliability and economic operation of power systems. Especially with the development of the power market, load forecasting has received more and more attention. The LSSVM can effectively deal with the nonlinear relationship existing in the prediction process for which it has received wide attention. In this paper, the load is predicted by LSSVM considering many outside elements. The prediction results show that the prediction result of LSSVM is better than BPNN.

References

- [1] Schoonees, J.A. (1988) Parallel distributed processing: practical applications of neural networks in signal processing. In: Communications and Signal Processing. Pretoria, Southern Africa. pp. 76-80.
- [2] Breiman, L. (2001) Random forests. *Machine Learning*, 45 (1): 5-32.
- [3] Heckerman, D. Geiger, D. Chickering, D. (1995) Learning Bayesian Networks: the combination of knowledge and statistical data. *Machine Learning*, 20: 196-243.
- [4] Biswas, M.A. Rafe, Robinson, Melvin D, Fumo N. (2016) Prediction of residential building energy consumption: A neural network approach. *Energy*, 117: 84-92.
- [5] Taborda, M. Almeida, J. Oliveir-Lima, J.A. Martins, J.F. (2015) Towards a web-based energy consumption forecasting platform. In: International Conference-Workshop on Compatibility and Power Electronics, Costa da Caparica, Portugal. pp. 577-580.
- [6] Xing, S. Sun, W. Yan, Y. Zhang, Z. (2019) Short-term load forecasting model of power system based on improved random forest algorithm. *Journal of Qingdao University (E&T)*, 34: 8-10.
- [7] Wang, W. Tian, S. Wang, Z. Xie, W. Bu, F. Tian, Y. Su, Y. (2017) A Power Peak Load Forecasting Method Based on Bayesian Network. *Distribution & Utilization*, 57-64.
- [8] Pan, D. Liu, H. Li, Y. (2008) Optimization Algorithm of Short-term Multi-step Wind Speed Forecast. *Proceedings of the CSEE*, 28: 87-91.
- [9] Ceperic, E. Ceperic, V. Baric, A. (2013) A strategy for short-term load forecasting by support vector regression machines. *IEEE Trans on Power Systems*, 28(4): 4356-4364.
- [10] Lu, H. Luo, F. Yang, X. Luo, B. Li, Y. Hao, J. (2019) A Power Load Forecasting Method Based on Data Driven. *Proceedings of the CSU-EPSA*.