Power quality assessment method of wind power grid connection based on least squares support vector machine

Jingjing Yang¹, Song Wang²*, Ruining He³, Changfeng Liu², Huaming Cao¹, Zitao Xu¹ and Xiangshan Hao³

¹Weifang Power Supply Company, State Grid Shandong Electric Power Company, Weifang, China
²School of Mechanical, Electrical & Information Engineering, Shandong University, Weihai, China
³Shandong Shanda Electric Technology Co.Ltd, Jinan, China

*Corresponding author e-mail: wszbs3@sdu.edu.cn

Abstract. In this paper, according to the 'Technical rule for connecting wind farm to power system' promulgated by the state, the least squares support vector machine (LSSVM) algorithm is used to build a mathematical model, and six indicators, namely, voltage deviation, voltage fluctuation, voltage flicker, grid harmonics, frequency deviation and three-phase voltage imbalance, are comprehensively considered and a certain quality level is calculated, which is used to evaluate the power quality of wind power stations and substation access points. LSSVM is a classification method based on the principle of minimizing institutional risk and combined VC dimension theory. Compared with traditional classification method, LSSVM has the advantages of low complexity, short training time, high accuracy and the like. According to the electric energy data provided by Weifang power grid, the accuracy and practicability of this evaluation method are verified.

1. Introduction

Wind energy is clean, efficient and renewable. And it has become China's third largest power generation method after coal power and hydropower. But due to the randomness and volatility of the wind speed, the wind farm output is uncontrollable, and the public connection point of wind power grid connection has a certain degree of impact on the power quality of the grid [1]. According to the national power quality standards, this paper aims to synthesize the six indexes, voltage deviation, voltage fluctuation, voltage flicker, grid harmonics, frequency deviation and three-phase voltage unbalance, and evaluate the quality of electric energy.

Support vector machine (SVM), a branch of machine learning theory, is a learning model which is widely used in the analysis of statistics, classification and regression. SVM is based on the structural risk minimization (SRM) principle, and solves problems of small sample, nonlinear, high dimension and local minimum points. The least squares support vector machine (LSSVM) is an extension of SVM. It uses equality constraint to replace the inequality constraint of the SVM. The quadratic programming problem is transformed into a linear system of equations, which reduces the computational complexity and improves the solution speed [2].

In this paper, a classification method based on LSSVM is proposed. Using the libsvm toolbox in MATLAB, the training sample data is converted into the required data format and perform a simple
scaling operation. Then select the appropriate kernel function and optimal parameters, and map the sample vectors to high-dimensional space to construct the optimal classifier faces and train a model. Finally, use the model to test and predict.

2. Characteristics of wind power grid connected generation

2.1. Influence of power quality fluctuation on the power grid
When the actual operating voltage does not match the rated voltage, the performance and safety of the electrical equipment will be affected; The voltage fluctuation will cause the quality of industrial products unqualified, lighting flashing, the uneven speed of the motor and the abnormal of electronic equipments; The damage of voltage flicker is that lighting may be hazardous to health and affect normal work; The harmonic will cause the waveform severely distorted and generate the EMI; The lower frequency of power system will have an impact on the safe operation of power plants and systems. The occurrence of frequency or voltage collapse will collapse the whole system and cause large-scale power failure; The unbalance of three-phase voltage will cause misoperation of relay protection, additional vibration torque of motor and heat generation [3]. It will also have a serious impact on the wind farm.

2.2. National power quality standard
GB/T 12325 Power quality - Deviation of supply voltage stipulates that the sum of absolute values of positive and negative deviation of 35 KV supply voltage and above shall not exceed 10 % of rated voltage. Allowable three-phase supply voltage deviation of 10 KV and below is ±7 % of rated voltage, while 220V single-phase power supply voltage is 7 % ~ 10 % of the rated voltage.

GB/T 12326 Power quality - Voltage fluctuation and flicker stipulates permissible value of voltage fluctuation at public power supply points: 2.5% for 10 KV and below, 2% for 35 ~ 110 KV and 1.6% for 220 KV and above.

GB/T 14549 Quality of electric energy supply-Harmonics in public supply network stipulates the total harmonic distortion rate of public power grids voltage (phase voltage): 5.0 % for 0.38 KV, 4.0 % for 6 ~ 10 KV, 3.0 % for 35 ~ 66 KV, and 2% for 110 KV.

GB/T 15945 Power quality-Frequency deviation for power system stipulates that the frequency deviation limit of power system under normal operation is ±0.2 Hz. When the system capacity is small, the limit can be relaxed to ± 0.5 Hz.

GB/T 15543 Power quality-Three-phase voltage unbalance stipulates that the allowable value of normal voltage unbalance degree is 2 %, which shall not exceed 4 % in a short time.

3. Principle of least squares support vector machine

3.1. Binary classification algorithm
There are two types of linearly separable sample sets as

\( (x_i, y_i), i = 1, \ldots, n, x \in R^n, y_j \in \{+1,-1\} \).

The general form of linear discriminant function in n - dimensional space is

\( f(x) = \omega \cdot x + b \) .

The corresponding classification surface equation is

\( \omega \cdot x + b = 0 \) .

The goal of the binary classification algorithm is to enable the separation hyperplane \( f(x) \) to correctly classify \( x \) and normalize the discriminant function, so that \( f(x) \geq 1 \) is met for all samples in the two categories. The following conditions need to be met is

\( y_i [\omega \cdot x_i + b] - 1 \geq 0, i = 1, \ldots, n \).

The distance from any sample point \( x_i \) to the interface \( f(x) \) is
\[ d_j = \frac{|f(x_j)|}{\|\omega\|} . \quad (4) \]

Maximize the minimum interval, and the smaller the \( \|\omega\| \) is, the better the classify interface is.

After introducing the relaxation factor \( \xi \geq 0 \), it can be expressed as the following constraint optimization problem.

\[
\begin{aligned}
\min_{\omega, \xi} & \quad \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{n} \xi_i \\
\text{s.t.} & \quad y_i (\omega^T \varphi(x_i) + b) = 1 - \xi_i
\end{aligned} \quad (5)
\]

The lagrange function of this problem is

\[
L(\omega, b, \xi, \alpha) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{n} \xi_i^2 - \sum_{i=1}^{n} \alpha_i \left[ y_i (\omega^T \varphi(x_i) + b) - 1 + \xi_i \right]
\quad (6)
\]

Among them, \( \alpha_i \geq 0 \) is Lagrange multiplier. According to KKT conditions.

\[
\begin{aligned}
\frac{\partial L}{\partial \omega} = 0 & \Rightarrow \omega = \sum_{i=1}^{n} \alpha_i y_i \varphi(x_i) \\
\frac{\partial L}{\partial b} = 0 & \Rightarrow \sum_{i=1}^{n} \alpha_i y_i = 0 \\
\frac{\partial L}{\partial \xi_i} = 0 & \Rightarrow \alpha_i = C \xi_i \\
\frac{\partial L}{\partial \alpha_i} = 0 & \Rightarrow y_i [\omega^T \varphi(x_i) + b] - 1 + \xi_i = 0
\end{aligned} \quad (7)
\]

The above formula can be written as the following linear equations

\[
\begin{bmatrix}
1 & 0 & 0 & -Z^T \\
0 & 0 & 0 & -Y^T \\
0 & 0 & C I & -I \\
Z & Y & I & 0
\end{bmatrix}
\begin{bmatrix}
\omega \\
b \\
\xi \\
\alpha
\end{bmatrix} =
\begin{bmatrix}
0 \\
0 \\
0 \\
1
\end{bmatrix} \quad (8)
\]

where \( Z = [\varphi(x_1)^T, \ldots, \varphi(x_n)^T]^T Y = [y_1, \ldots, y_n], I = [1; \ldots; 1], \xi = [\xi_1; \ldots; \xi_n], \alpha = [\alpha_1; \ldots; \alpha_n] \)

Eliminate \( \omega, \xi \) and get the following simplified expression as

\[
\begin{bmatrix}
0 \\
Y^T \\
Y Z Z^T + C^{-1} I
\end{bmatrix}
\begin{bmatrix}
b \\
\alpha
\end{bmatrix} =
\begin{bmatrix}
0 \\
0 \\
0
\end{bmatrix} \quad (9)
\]

Expand into a matrix of linear equations and we can obtain

\[
\begin{bmatrix}
0 & y_1 & \cdots & y_n \\
y_1 & y_1 y_1 K(x_1, x_1) + \frac{1}{C} & \cdots & y_1 y_n K(x_1, x_n) \\
\vdots & \vdots & \ddots & \vdots \\
y_n & y_n y_1 K(x_n, x_1) & \cdots & y_n y_n K(x_n, x_n) + \frac{1}{C}
\end{bmatrix}
\begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_n
\end{bmatrix} =
\begin{bmatrix}
0 \\
1 \\
\vdots \\
1
\end{bmatrix} \quad (10)
\]

After simplification the equation (10) becomes as

\[
\begin{aligned}
\sum_{i=1}^{n} \alpha_i y_i &= 0 \\
\sum_{i=1}^{n} \sum_{j=1}^{n} \left[ \alpha_i y_i y_j \varphi(x_i) \varphi(x_j) + b y_i + \frac{\alpha_i}{C} \right] &= 1
\end{aligned} \quad (11)
\]
Therefore, the LSSVM classifier is

\[
f(x) = \text{sgn} \left[ \sum_{i=1}^{n} \alpha_i y_i K(x, x_i) + b \right], \quad i = 1, 2, \ldots, n.
\] (12)

3.2. Multi-classification algorithm

Among the multi-classification methods, the basic idea of one-to-one algorithm is to divide m-kind-sample into pairs and construct m(m-1)/2 SVM classifiers [4]. For a single classifier, only two types of data need to be classified to solve the quadratic optimization problem of the classifier as shown in equation (13).

\[
\begin{align*}
\min & \quad \frac{1}{2} (\omega^{j,k})^T \omega^{j,k} + C \sum_{i=1}^{n} \xi^{j,k}_i \\
\text{subject to} & \quad (\omega^{j,k})^T \varphi(x_i) + b^{j,k} \geq 1 - \xi^{j,k}_i, \quad y_i \neq j \\
& \quad (\omega^{j,k})^T \varphi(x_i) + b^{j,k} \geq -1 + \xi^{j,k}_i, \quad y_i \neq k \\
& \quad \xi^{j,k}_i \geq 0
\end{align*}
\] (13)

m(m-1)/2 decision functions can be obtained by solving equation (13).

\[
(\omega^{1,1})^T \varphi(x) + b^{1,1}, \ldots, (\omega^{j,k})^T \varphi(x) + b^{j,k}
\] (14)

Finally, the classification decision function is

\[
f(x) = \text{sgn} \left[ (\omega^{j,k})^T \varphi(x) + b^{j,k} \right].
\] (15)

Compared with other classification methods, the one-to-one algorithm has the best performance in training time, classification accuracy, determinacy, recognition rate and classification speed. Therefore, the one-to-one algorithm is used in the evaluation of wind power quality level in this paper.

4. Evaluation of power quality level

4.1. Evaluation model

Based on the principle of LSSVM and using libsvm toolkit in MATLAB, the power quality evaluation model of Weifang wind power generation is established. Since there is no clear standard for power quality level in China, we refined the scope of each evaluation index in combination with the 'Technical rule for connecting wind farm to power system' promulgated by the state and the distribution characteristics of wind power generation and substations in Weifang city. Finally formulated the limit of six indexes under each level. Each power quality index is divided into 6 levels. Level 1 - 6 are defined as superfine, high quality, good, medium, qualified and unqualified. See Table 1 for the limit values of each index level.

| Level | frequency deviation /Hz | voltage deviation /% | voltage fluctuation /% | voltage flicker % | harmonic voltage /V | three-phase imbalance /% |
|-------|-------------------------|----------------------|------------------------|------------------|---------------------|------------------------|
| 1     | \leq 0.02               | \leq 1.5             | \leq 0.1               | \leq 0.1         | \leq 1.0            | \leq 1.0               |
| 2     | \leq 0.04               | \leq 2.5             | \leq 0.2               | \leq 0.15        | \leq 2.0            | \leq 1.1               |
| 3     | \leq 0.06               | \leq 3.5             | \leq 0.3               | \leq 0.2         | \leq 3.0            | \leq 1.2               |
| 4     | \leq 0.08               | \leq 4.5             | \leq 0.4               | \leq 0.25        | \leq 4.0            | \leq 1.5               |
| 5     | \leq 0.20               | \leq 7.0             | \leq 2.0               | \leq 1.0         | \leq 5.0            | \leq 2.0               |
| 6     | \leq 0.20               | \leq 7.0             | \leq 2.0               | \leq 1.0         | \leq 5.0            | \leq 2.0               |

Table 1. Limit values of levels
According to the purpose of the study, the input of the evaluation model is the actual value of six indexes, and the output is the comprehensive level of power quality we want to get [5]. The specific flowchart is shown in the Fig. 1. After using the model for comprehensive evaluation, the data in level 6 is unqualified data, which means, the power quality is unqualified and cannot be incorporated into the power grid node.

![Figure 1. LSSVM power quality comprehensive evaluation flowchart](image)

During the modeling process, according to the discriminant function (12), we choose the radial basis function (RBF) as the kernel function of the training model. Its form is

$$K(x, x_i) = \exp\left(-g \|x - x_i\|^2\right), \quad g > 0,$$

where parameter $g$ is an important parameter in kernel function, which affects the complexity of LSSVM classification algorithm. Therefore, we choose the grid search algorithm to find the optimal parameter $g$ and penalty parameter $C$. Grid search algorithm uses the K – CV method to select the group of $C$ and $g$ which has the highest classification accuracy as the best parameter. This method can find the global optimal solution in the optimum range which is large enough and the step distance which is small enough [6].

4.2. Example analysis

When the six indexes of a group of measured data are in different levels according to the above standards, it is difficult to determine the power quality level directly at that moment. Therefore, this paper proposes a method of using least square support vector machine to establish a multi-dimensional and multi-classification model, which can be used to judge the power quality level more intuitively and reliably.

First, in terms of parameter selection, the result of grid search algorithm is shown in Fig. 2. From the Fig. 2 we can see that the accuracy is 100% when $C$ is 0.57435 and $g$ is 3.0314, which means these are the optimal parameters.
In this paper, 660 samples are randomly generated according to the range of six indicators under each level, of which the first 600 samples are used as training samples to build one model. The last 60 samples are used as test samples to verify the accuracy of the model. The comparison of predicted and actual values is shown in the following Fig.3. According to this figure, the classification of the test data under each level is consistent with the actual category, which verifies the accuracy of the model.

Finally, use the model to predict the levels of 70 groups of continuous measured data provided by Weifang power supply bureau. The results are shown in the Fig.4. It can be seen that the comprehensive level of electric energy during this period is maintained between level 2 and level 3. The quality of electric energy is good and slightly superior, and the fluctuation is extremely small, which completely meets the requirements of wind power grid connection.

According to the operation characteristics of Weifang power grid, there are three key indicators: voltage deviation, voltage fluctuation and harmonic. Therefore, we compare the three indicators in pairs, and use the data distribution to concretely observe the classification of power quality levels.
When we choose voltage deviation and harmonic as the objects of observation, the distribution is shown in the Fig.5 (a). The green hollow points represent the data groups of level 2, and the blue ones represent the level 3. In the test set, voltage deviation range from 2.2 to 3.8, and by looking up Table 1 we can get that it should span three levels: level 2, level 3 and level 4. Similarly, harmonic range from 1.6 to 2.3, which belongs to level 2 and level 3. After evaluating the six indicators comprehensively by the model, the power quality is distributed in two levels.

Fig.5 (b) and Fig.5 (c) are the distribution of power quality levels under voltage deviation-voltage fluctuation and voltage fluctuation-harmonic with actual data. Voltage fluctuation range from 0.08 to 0.24, which belongs to level 1 to level 3 according to Table 1, and is also in the two levels finally.

Through such distribution results, we can find that when two or more indicators are at different levels, the model can comprehensively predict a certain level through learning and judgment. And with the model, it is easy to judge whether the quality of electric energy is qualified or not, as well as the excellent degree of qualified electric energy, which is of great significance to wind power grid connection technology.

5. Conclusions
The LSSVM is very suitable for power quality level evaluation, and even with fewer samples, it still obtains better classification results.

The evaluation model based on LSSVM method can quickly convert power data into level data, which can make us informed about the power quality more intuitively at any moment. It is of great significance in wind power grid connection.

According to the electric energy index data provided by Weifang city, after the model analysis in this paper, it can be concluded that the electric energy quality is stable and good within the corresponding time period, which meets the requirements of grid connection.

Acknowledgments
The authors wish to thank the corporate researchers in the common laboratory for the generous help of simulation program and experiment.

References
[1] Wang Shenzhe, Gao Shan, Li Haifeng, Zhang Long, “Evaluation of Power Quality in Grid Planning Scheme with Wind Power Integration,” Transactions of China Electrotechnical Society, pp.56-65, Aug. 2013.
[2] Zhang Quanming, LIU Huijin, “Application of LS-SVM in Classification of Power Quality Disturbances,” Proceedings of the CSEE, pp.106-110, Jan.2008.
[3] Liu Wei, “Evaluation and Control Measures on Power Quality of Wind Power Grid,” Beijing Jiaotong University.2014.
[4] Shan Yugang, Wang Hong, Dong Shuang, “Improved Multi-Classification Algorithm of One-against-on SVM,” Computer Engineering and Design, pp.1837-1841, May.2012.

[5] N. Hadjsaid, J.-F. Canard and F. Dumas, “Dispersed generation impact on distribution networks”, IEEE Computer Applications in Power, pp.22-28, Vol.12, 1999.

[6] Wang Jianfeng, Zhang Lei, Chen Guoxing, “A Parameter Optimization Method for an SVM Based on Improved Grid Search Algorithm,” Applied Science and Technology, pp.28-31, Jun.2012.