Photovoltaic (PV) Power Prediction Based on ABC - SVM

To cite this article: Han Mo et al 2018 IOP Conf. Ser.: Earth Environ. Sci. 199 052031

View the article online for updates and enhancements.
Photovoltaic (PV) Power Prediction Based on ABC - SVM

Han Mo¹, Yu Zhang², Zhaokun Xian¹ and Hangping Wang¹

¹Hezhou University, GuangXi 542800, China;
²Guilin University of Technology, GuangXi 541004, China

Abstract. Photovoltaic(PV) generation forecasting technology is an important part of the development of PV grid-connected system, which has a great impact on the reliability, stability and economy of the grid. Due to uncertain meteorological factors such as light intensity, temperature, humidity and wind speed, the output of PV system is random, intermittent and uncontrollable, which makes the prediction accuracy of PV power generation not high. In order to improve the accuracy of PV power generation forecasting, artificial bee colony (ABC) algorithm is used to optimize the support vector machine (SVM) model and predict the PV power generation: Firstly, the ABC algorithm is used to optimize the penalty factor $C$ and the kernel function $g$ of the SVM prediction model. Secondly, the optimized SVM algorithm is trained and tested. Finally, the ABC-SVM algorithm is used to predict the PV generation. The simulation results show that compared with the traditional SVM algorithm, the ABC-SVM algorithm has less control parameters, strong optimization ability, higher prediction accuracy and more stable system, which provides a certain scientific research value for PV power generation forecasting.

1. Introduction

The report of the 19th National Congress of the CPC pointed out that China's energy industry must establish and practice the concept of "Lucid waters and lush mountains are invaluable assets" during the 13th Five-Year Plan period. As a kind of renewable energy with zero pollution and wide sources, solar energy has attracted more and more attention in its research and application. However, affected by meteorological factors, solar power generation is intermittent, random and unstable[1]. With the increasing capacity of PV integrated into the grid, the coordination of conventional power and PV power generation has become increasingly prominent. Predicting PV power accurately in advance can effectively reduce the impact of grid-connected PV power on the power system, help the electric power management department to plan, dispatch and control the operation mode of the grid, improve the power quality of the grid and PV energy absorption capacity.

SVM’s research and application is very active in various fields because of the characteristics of high fitting accuracy, few selection parameters, strong generalization ability and insensitivity to dimension. In recent years, the application of SVM in PV power prediction has also been paid attention to. However, the accuracy and stability of SVM prediction model are deeply affected by the penalty coefficient parameter $C$ and kernel function parameter $g$. The parameter optimization has some shortcomings such as long time and easy to fall into local optimum. When the single SVM prediction method is used to predict the PV power generation, the error is relatively large, so it is usually combined with the advantages of each prediction method to improve the prediction accuracy.
Therefore, we usually combine each prediction method to complement each other in order to improve the prediction accuracy.

In this paper, an ABC algorithm is proposed to optimize the penalty coefficient $C$ and kernel function parameter $g$ of SVM PV prediction model. The special labor division and cooperation mechanism of ABC algorithm enables the bees to cooperate with each other according to different search strategies to complete the optimization work together, so that the optimized SVM algorithm has strong global optimization ability and learning ability\cite{2}. Finally, the historical data of PV power station are trained and predicted as the input of the prediction model.

2. PV cell output characteristics

PV cells are made of semiconductor materials, which are the basic units of the PV power generation system. The equivalent circuit is shown in Figure 1.

\[\text{Figure 1. The equivalent circuit diagram of PV cells}\]

Here, $I_{ph}$, $I_d$, $R_{sh}$, $I_{sh}$, $R_s$, $I_o$ and $U_o$ represent respectively photo-generated current, diode current, internal equivalent resistance, internal equivalent current, series resistance, output current and output voltage.

The meteorological factors which affect PV power generation are solar irradiance, temperature, humidity and wind conditions\cite{3}. According to the influencing factors and the equivalent circuit diagram of PV cells, a mathematical model is established and the characteristics of PV cells are simulated in Matlab/Simulink. The results are shown in Figure 2.

\[\text{(a) The P-U characteristics of PV cells with constant light intensity}\]
The simulation results show that the maximum output power increases with the increase of illumination intensity when the temperature is constant and the output power decreases with the increase of temperature when the illumination intensity remains unchanged.

In actual working environment, the PV cells of light intensity and temperature both are real-time changes, not invariable. And the PV cells generally will not be used as a single power supply, but by a number of series, parallel PV arrays. This makes the output power to a certain degree of randomness, intermittence and volatility. It requires a good maximum power point tracking (MPPT) technique to capture the optimal operating voltage of the PV system.

3.ABC optimization SVM model

3.1 Nonlinear SVM model

Support Vector Machine (SVM) is an intelligent learning method with machine learning ability. By observing the sample data, the corresponding mathematical expressions are established to analyze the potential laws in the data, then the obtained law is used to analyze and predict the future data. Because there are many characteristics, which belong to nonlinear factors affecting PV power generation, nonlinear SVM is used to predict. The input variables are mapped into a high-dimensional feature space by kernel function, and then linear regression is performed in the feature space.

Suppose the given sample is \( \{X_i, Y_i\}_{i=1,2,...,N} \), \( X_i \) is the input vector, \( Y_i \) is the corresponding output value and \( N \) is the sample number. The mapping process is: \( X_i \rightarrow K(X_i, X_j) \rightarrow \phi(X_i) \cdot \phi(X_j) \), which \( K(X_i, X_j) \) is the kernel function, and the objective function is as follows:

\[
F(X, \omega) = \omega \cdot \phi(X) + b
\]

Here, \( \omega \) is the weight, and \( \phi(X) \) is the nonlinear mapping function of input, \( X \) is the sample input value, and \( b \) is the bias value. Then the linear regression is realized by using the insensitive loss function \( \varepsilon \) in high space, \( \varepsilon \) represents allowable training loss, its formula is:

\[
L_{\varepsilon}(Y, F(X, w)) = \begin{cases} 
0, & |y - F(X, w)| \leq \varepsilon \\
|y - F(X, w)| - \varepsilon, & \text{others}
\end{cases}
\]

Here, \( Y \) is the true value of the function. The structural risk minimization principle is used to reduce the complexity of the linear regression model in high-dimensional space. To avoid over-learning of function \( F(X) \), penalty coefficient \( C \) and relaxation variables \( S \) and \( S^* \) are introduced. The optimization formula is as follows:

\[
\text{min} \left\{ \frac{1}{2} \| \omega \|^2 + C \sum S_i \right\}
\]
By introducing Lagrange function, the duality principle of Wolf is transformed into dual form.

\[
\min \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} (S_i + S_i^*) \right\}
\]
\[
st. \begin{cases} 
Y_i - w \cdot \varphi(X_i) - b \leq \epsilon \\
w \cdot \varphi(X) + b - Y_i \leq \epsilon \quad i = 1, 2, \ldots, N \\
S_i, S_i^* \geq 0
\end{cases}
\]

Here, \( \alpha_i \) and \( \alpha_i^* \) are Lagrange multipliers.

3.2 The optimization principle of ABC algorithm

Artificial bee colony (ABC) algorithm was proposed by Karaboga group of Ergis University in Turkey in 2005, mainly to simulate the intelligent nectar gathering behavior of bee colony \(^{(5)}\). The standard ABC algorithm divides the artificial bee colony into three categories by simulating the honey-collecting mechanism of the actual bee: employed bee, onlooker bee, and scout bee. Those bees collect nectar according to their respective division of labor, and realize the sharing and exchange of nectar source information, so as to find the optimal solution to the problem \(^{(6)}\).

The optimization problem is assumed to be \( \{ \min F(x), x \in S \} \). In ABC algorithm, finding the location of honey source is understood as finding the point in space, each feasible solution of the problem corresponds to the location of one honey source. The process of the bees searching for the best honey source is the process of searching for the optimal solution. The nectar quantity contained in the honey source corresponds to the fitness value of the solution of the problem. Suppose ABC algorithm contains the initial population of \( NP \) solutions, each solution is \( x_i (i = 1, 2, \ldots, n) \).

The bees search for nectar sources around them first, then search for new sources according to the following formula \(^{(7)}\):

\[
v_{ij} = x_{ij} + \varphi_{ij} (x_{ij} - x_{kj})
\]

Here, \( x_{ij} \) is the location after search for \( i \) bee \( j \) dimension, \( K \in \{1, 2, \ldots, n\}, k \neq i, j \in \{1, 2, \ldots, D\} \), both numbers are randomly selected, and \( \varphi_{ij} \) is a random number between \([-1, 1]\), which controls the generation range of the \( x_{ij} \) domain. As the search approaches the optimal solution, the scope of the domain gradually decreases.

All employed bees were returned to the hive after completing the search task, and the honey source information was shared with onlooker bees. Then the nectar source was selected by onlooker bees which according to the nectar quantity of each honey source (i.e. the fitness function value of the solution) and formula \(^{(6)}\) \(^{(8)}\).

It can be seen from the formula \(^{(6)}\) that the more nectar sources, the more nectar sources are, the easier it is to be selected.

\[
P_i = \frac{\text{Fit}_i}{\sum_{n=1}^{N} \text{Fit}_n}
\]
\[
Fit_i = \begin{cases} 
\frac{1}{1 + F_i} & F_i \geq 0 \\
1 + |F_i| & F_i < 0 
\end{cases}
\]

Here, \(P_i\) represents the selection probability of the first honey source, \(Fit_i\) represents the fitness value of the second honey source, \(N\) is the number of honey sources, and \(F_i\) is the objective function of the optimization problem. Then onlooker bees search the neighborhood of selected honey sources and determine the location of a new honey source according to the formula (5), then retain the optimal solution according to the greedy selection strategy.

The ABC algorithm has a control parameter called Limit, which is used to record the number of times that a solution has not been updated. It shows that the solution falls into local optimum if a solution is not improved after adding a solution after Limit cycle\(^9\). Supposing the abandoned solution is \(x_i\) and \(j \in \{1, 2, \ldots, D\}\), a new solution is generated randomly by reconnaissance peak instead of \(x_i\) with following formula\(^10\):

\[
x_{ij} = x_{\text{min},j} + \text{rand}(0,1)(x_{\text{max},j} - x_{\text{min},j})
\]

3.3 ABC-SVM prediction model

The accuracy and generalization ability of SVM algorithm depend on the penalty factor \(C\) and kernel function parameter \(g\). Therefore, it is necessary to optimize the parameters \(C\) and \(g\) to obtain SVM with high prediction accuracy.

The ABC algorithm, which has the advantages of fewer parameters, simple principle, easy implementation and strong robustness in the global optimization algorithm\(^{11}\), is widely used in combination optimization, multi-objective optimization, artificial neural network training and so on.

The main steps of establishing ABC-SVM prediction model are as follows:

1) Input original sample data: The data of four factors affecting PV power generation, i.e. illumination intensity, temperature, humidity and wind speed, were monitored in the laboratory as input of the model. And the training set and test set were selected in advance.

2) Data normalization: The training set and the test set are normalized to the \([0, 1]\) interval.

3) Using ABC algorithm to optimize the penalty factor \(C\) and kernel function \(g\) of SVM prediction model as follows:

1) Initialization phase: It mainly initializes the number of peaks, the number of honey sources, the maximum number of iterations and the maximum number of searches\(^{12}\).

2) Computational fitness function: A new optimal solution is to be found by each bee source’ objective function according to the formula (1), and the mean square error (MSE) of the SVM model is used as the fitness function value of the solution.

3) Employed bee phase: Employed bees search for each bee source and select parameters that need variation randomly.

4) Onlooker bee phase: If the honey quantity is better than the original solution, the new solution will be replaced by the new solution, otherwise the original solution will be retained.

5) Scout bee phase: Calculating the probabilities of bees selecting honey sources, and the greedy algorithm is used to select the source of honey. If the nectar quantity of the honey source does not change during scout bee phase, the solution should be abandoned and replaced by the honey source searched by the scout bee to preserve the optimal solution in scout bee period\(^{13}\).

6) Judgement of termination iteration condition: Determining whether the condition of terminating iteration is satisfied or not. If it is not satisfied, the iteration number is added to repeat the above process. If it is satisfied, the optimal \(C\) and \(g\) of SVM model are recorded.

4) ABC-SVM photovoltaic forecasting model: The parameters \(C\) and \(g\) obtained by ABC optimization are trained to SVM network, then ABC-SVM is used to predict the PV power generation in the next few days.
The algorithm flow chart of optimizing SVM model parameters based on ABC is shown in Figure 3.

![Figure 3. ABC-SVM algorithm flow chart](image)

4. ABC-SVM PV power generation forecast simulation

4.1 Error evaluation index

The root mean square error value (RMSE) and the determination coefficient $R^2$ are used to evaluate the prediction results. Their mathematical model formulae are as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (P_{ri} - P_{pi})^2}{n}}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (P_{ri} - P_{pi})^2}{\sum_{i=1}^{n} (P_{ri} - P_{ri})^2}$$

(8)

(9)

Here, $P_{ri}$, $P_{pi}$, and $P_{ri}$ are the average values of actual PV power, predicted PV power, and actual PV power respectively, and $n$ is the number of PV power data.

$RMSE$ reflects the degree that the measured data deviate from the true value. The smaller the $RMSE$, the higher the accuracy of measurement.

$R^2$ is used as a measure of the goodness of fit of the model. The closer $R^2$ is to 1, the better the goodness of fit. On the contrary, the smaller $R^2$ is, the worse the goodness of fit is.

4.2 Simulation analysis

In order to verify the validity of the ABC-SVM prediction model, the PV power data collected by GuangXi Key Laboratory of New Energy and Building Energy Saving of Guilin University of Technology are used as samples. These samples have 1650 PV power data which were collected every 10 minutes from 8:00 PM to 17:00 PM daily in December 2014 contain four characteristic values of irradiance, temperature, humidity and wind speed. Then, 100 data were randomly selected for testing. In order to evaluate the algorithm better, the ABC-SVM prediction model is compared with the traditional SVM prediction model. The prediction results and relative error comparisons of the two algorithms are shown in Figure 4 and Figure 5 respectively, and the error evaluation is shown in Table 1.
The predicted value of SVM
The predicted value of ABC-SVM
Figure 4. Comparison of ABC-SVM, SVM and actual values

The relative error curve of the ABC-SVM model of Figure 5 shows that the relative error curve of the ABC-SVM has much less fluctuation than SVM model, it means the prediction result of the ABC-SVM method is more accurate and more stable.

Table 1 Performance comparison of SVM, ABC-SVM models

| Sample number | $R^2$ | $RMSE$ | $MAD$ |
|---------------|-------|--------|-------|
| ABC-SVM       | 0.84  | 227.08 | 128.69|
| SVM           | 0.77  | 311.82 | 161.517|

As can be seen from Figure 4, the predicted curve of ABC-SVM model is closer to the actual value curve than SVM model, indicating that the ABC-SVM model is more capable of tracking PV power than SVM model.

The relative error curve of the ABC-SVM model of Figure 5 shows that the relative error curve of the ABC-SVM has much less fluctuation than SVM model, it means the prediction result of the ABC-SVM method is more accurate and more stable.

Table 1 shows that the decision coefficient $R^2$ of ABC-SVM is improved about 7% compared with the traditional SVM algorithm, so the fitting degree of ABC-SVM is better. Meanwhile, the $RMSE$ and $MAD$(Mean Absolute Difference) of ABC-SVM are smaller than SVM, which indicating that the prediction accuracy of ABC-SVM is higher.

5. Conclusion
In this paper, the penalty factor $C$ of support vector and the kernel function parameter $g$ are optimized by the artificial peak swarm optimization algorithm, and the ABC-SVM PV power forecasting model is established. Four meteorological factors including solar radiation, temperature, humidity and wind speed were selected as the main influencing indexes. The ABC-SVM PV power forecasting model is
used to train and predict the samples. The prediction algorithm is evaluated by using the coefficient of determination and the root mean square error. The results show that compared with the traditional SVM, the optimized prediction model has higher fitting accuracy and robustness, it meets the actual needs of engineering.

Acknowledgment
1. The he Zhou Foundation for Research and Development of science and Technology.
2. Guangxi Undergraduate Training Programs for Innovation and Entrepreneurship.

References
[1] Caoli. Hu, Reliability Evaluation and Optimal Allocation of a Class of Power Generation System with Intermittent Energy[D], Hangzhou Dianzi University, 3(2016).
[2] Xiangming. GAO, Shifeng. YANG, Sanbo. PAN, A forecasting model for output power of grid-connected photovoltaic generation system based on EMD and ABC-SVM[J]. Power System Protection and Control, 43(21), 86-92(2015).
[3] Guang. Yang, Association Relationship Analysis and Research of PV Power Generation and meteorological Impact Factors. North China Electric Power University, 3(2014)
[4] A. Besalatpour, M. Hajabbasi, S. Ayoubi, A. Gharipour and A. Jazi. Prediction of soil physical properties by optimized support vector machines[J].
a) International Agrophysics, 26(2), 109-115(2012).
[5] Guangming. ZHAI, Guohe. LI, Weijiang. WU, Improved algorithm of artificial bee colony based on Spark[J]. Journal of Computer Applications, 37(7),
a) 1906-1910(2017).
[6] Jianxia, Yin. Research on Artificial Bee Colony Algorithm and its Application[D]. Xidian University,
a) 3(2012).
[7] Soheila Ghambari, Amin Rahati, An improved artificial bee colony algorithm and its application to reliability optimization problems. Applied Soft Computing, 62, 736-767, (2018).
[8] J.C. Bansal, A. Gopal, A.K. Nagar, Stability analysis of Artificial Bee Colony optimization algorithm, Swarm and Evolutionary Computation BASE DATA (2018), doi: 10.1016/j.swevo.2018.01.003.
[9] Yingyu. LU, Bing. WANG, Xuewen. LI, YongQIN[J]. Deep accelerating artificial bee colony algorithm, 29(12), 4445-4447(2012).
[10] Soheila Ghambari, Amin Rahati, An improved artificial bee colony algorithm and its application to reliability optimization problems. Applied Soft Computing, 62, 736-767(2018).
[11] Mingxiang. ZANG, Xuan. MA, Yiming. DUAN, Improved artificial bee colony algorithm[J]. Journal of Xidian University(Natural Science Edition), 42(2), 65-70(2015).
[12] Hongman. YAO, Liangxi. QIN, Pan. HU, Artificial Bee Colony Algorithm Based on Improved Search Strategy and Chaotic Mechanism[J], ISUANJI YU XIANDAIHUA, (6), 79-84(2016).
[13] Chaoqun.ZHANG, Jianguo. ZHENG, Taoshen. LI, Effect of Scout Bees on the Performance of Artificial Bee Colony Algorithm[J],Journal of Guangxi Normal University(Natural Science Edition), 21(3), 72-80(2013).
[14] Yu. ZHANG, Han. MO, Lieping. ZHANG, Photovoltaic Power Prediction Based on Fuzzy Support Vector Machine[J], 46(1),116-120(2017).