# Short-Term Photovoltaic Power Generation Forecasting Based on Environmental Factors and GA-SVM

### Jidong Wang[†], Ran Ran*, Zhilin Song* and Jiawen Sun*

**Abstract** – Considering the volatility, intermittent and random of photovoltaic (PV) generation systems, accurate forecasting of PV power output is important for the grid scheduling and energy management. In order to improve the accuracy of short-term power forecasting of PV systems, this paper proposes a prediction model based on environmental factors and support vector machine optimized by genetic algorithm (GA-SVM). In order to improve the prediction accuracy of this model, weather conditions are divided into three types, and the gray correlation coefficient algorithm is used to find out a similar day of the predicted day. To avoid parameters optimization into local optima, this paper uses genetic algorithm to optimize SVM parameters. Example verification shows that the prediction accuracy in three types of weather will remain at between 10% -15% and the short-term PV power forecasting model proposed is effective and promising.

**Keywords**: Photovoltaic, Short-term power forecasting, Environmental factors, Genetic algorithm, Support vector machine

## 1. Introduction

PV is used widely because of such advantages as less pollution, and the flexibility of scale [1]. However, it is difficult for PV systems to implement energy management and grid scheduling effectively due to the uncertainty, the volatility and intermittent of the PV system affected by environmental factors. As a result, the operational risks of the power grid increase. Therefore, the accuracy of 24 hours-ahead PV power forecasting is of great significance.

Current PV power forecasting method can be classified into physical method and statistical method. The physical method is to calculate photovoltaic power generation based on the environmental information of the photovoltaic cell (including the type of weather, solar radiation intensity, temperature, etc.), which is more complex. In comparison, the statistical method is to calculate the power generation based on historical output power data of PV systems, which is more widely used. And the statistical method includes direct and indirect prediction method. The indirect method is to calculate the output power of PV system based on the predicted value of solar radiation intensity and environmental parameters of the PV system. The direct method is to predict the output power directly based on historical output power data as input, which is widely used. Currently, short-term power forecasting method for PV systems mostly is the direct prediction. And depending on the mathematical model used, the direct method can be divided into time series prediction methods [2, 3], the regression model prediction method [4], neural network prediction method [5, 6], Markov Chain prediction method [7-9] and the like. Among these methods, the artificial intelligence algorithm, including the neural network algorithm and the probability prediction algorithm, is widely used currently. Papers [10-13] establish BP neural network model to forecast the PV power output by analyzing the influence of solar radiation intensity, temperature, cloud cover and other meteorological factors on PV output power. The paper [14] establishes dynamic Bayesian network model to forecast the probability distribution of the short-term power output of PV system, after analyzing relationships among the various factors affecting the PV system. The paper [15] uses support vector machine algorithm to build a predictive model, through analyzing the influence of solar radiation intensity and temperature for photovoltaic power generation system, analyzes and compares prediction accuracy of two different weather conditions in Italy. The paper [16] uses quantile regression algorithm and random forest algorithm for modeling, taking historical generation data, weather measurement data as input to get the numerical prediction and the probability prediction of photovoltaic power generation system.

After analyzing the effects of different weather types and other environmental factors on the PV system, this paper classifies the weather categories into three types, and determines the similar day of the forecasting day according to the gray correlation coefficient algorithm. Then this paper takes power data of the similar day and the weather data of the forecasting day as inputs of the model, and use improved support vector machine algorithm (GA-SVM) to

† Corresponding Author: Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, China. (jidongwang@tju.edu.cn)
* Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, China. (ranran1110@tju.edu.cn, songzhilin111@163.com, jiawensun89@gmail.com)

forecast short-term power output of the PV system. At last, the simulation example shows the effectiveness of this model. Specifically, the contribution of this paper is as follows:

(1) Firstly, we classified the weather categories into three types, and used gray correlation coefficient algorithm to find the similar day, which was the base of accurate forecast of PV power.
(2) Secondly, the GA algorithm was applied to optimize parameters of classical SVM, which can avoid parameters optimization into local optima and increased the accuracy of classical SVM algorithm to a large extent.

## 2. Impact of Environmental Factors on PV Systems

The power output of the PV system is greatly influenced by the natural environment, including the fixed location, irradiation angle, solar radiation intensity, temperature, humidity, cloud cover and other environmental variables, as well as the conversion efficiency associated with the device characteristics. Although the generating capacity of the PV system is a non-stationary random process for the established photovoltaic array, the time series of the historical power has the same source. The data itself contains the fixed location, and other factors of the irradiation angle. So the time series has a high degree of autocorrelation. While the power output of the PV system is closely related to the weather type, temperature, solar radiation intensity, humidity and other environmental factors related, it will be more useful to predict the short-term output of the PV system accurately, after taking the historical power output data and environmental factors related as inputs of the forecasting model of the PV system.

### 2.1 The classification of weather category

To establish the corresponding prediction model, it is necessary to divide traditional weather category for the forecasting model of PV systems. Current weather category can be classified into sunny day, cloudy day, overcast day, rainy day and foggy day. It is difficult to train the model with too many weather types, because it can cause feature vector data sets of weather category too scattered. Since main natural factors influenced PV system is solar radiation intensity and ambient temperature, it will be rational for us to classify rainy day and the overcast day with small fluctuations, low average in solar radiation into one type. According to the fluctuation degree and the average value of natural environment, this paper classifies the weather category into three types. The result is shown in Table 1 [17].

**Table 1.** The classification of weather category

| Reclassification of weather category | Type A | Type B | Type C |
|---|---|---|---|
| Traditional weather category | Sunny day | Cloudy day | Rainy day, Overcast day, Snowy day |

### 2.2 The selection of the similar day

Determine the forecasting day's weather type according to this day's weather data. Then select its corresponding similar day based on gray correlation coefficient algorithm, and take this day's power output data as the input of the forecasting model.

Because the solar radiation intensity, temperature and humidity are the key to the determination of the class of weather type, and the papers [18-19] analyze the fluctuation of different weather factors to the power output of the PV system. We analyze the relevance of various weather factors and photovoltaic power, and then select the maximum solar radiation intensity, the minimum solar radiation intensity, the highest temperature, the lowest temperature, and humidity as feature amounts of similar day.

The meteorological eigenvector selected is shown as follows:

$$x_i = \left[ x_i(1), x_i(2), x_i(3), x_i(4), x_i(5) \right] = \left[ l_{hi}, l_{li}, t_{hi}, t_{li}, h_i \right] \tag{1}$$

where, $l_{hi}$ indicates the value of the maximum solar radiation intensity in the $i^{th}$ day; $l_{li}$ indicates the value of the minimum solar radiation intensity in the $i^{th}$ day; $t_{hi}$ indicates the value of the highest temperature in the $i^{th}$ day; $t_{li}$ indicates the value of the lowest temperature in the $i^{th}$ day; $h_i$ indicates the value of the humidity in the $i^{th}$ day.

The formula of the correlation coefficient is shown as follows:

$$\delta_i(j) = \frac{(\min \min |x_0(j) - x_i(j)| + \rho \max \max |x_0(j) - x_i(j)|)}{(|x_0(j) - x_i(j)|) + \rho \max \max |x_0(j) - x_i(j)|)} \tag{2}$$

where, $x_0(j)$ is the $j^{th}$ feature vector of the predicted day; $x_i(j)$ is the $j^{th}$ feature vector of the $i^{th}$ day. $\rho$ is taken as 0.5.

The formula of the similarity between the predicted day and the $i^{th}$ history day is shown as follows:

$$\lambda_i = \frac{1}{5} \sum_{j=1}^{5} \delta_i(j) \tag{3}$$

The day with the max similarity is taken as the similar day of the forecast day [18-19].

## 3. The Algorithm of Support Vector Machine

The network structure of support vector machine (SVM) is shown in Fig. 1.

The input is represented by $x_i$. $k(x_i, x)$ represents the kernel function. $a_i$ represents the weight. And the output is represented by $y$. Compared with the traditional artificial neural network, support vector machines is a multilayer perception including a hidden layer. And hidden nodes and network weights are calculated automatically by the algorithm [20-21].

To a given sample $\{(x_i, y_i)\}(i = 1, 2, \cdots, n)$, $x_i$ is the Input vector; $y_i$ is the output vector; $n$ is the sample size. The algorithm of SVM utilizes the linear function below to implement the regression forecasting.

$$Y_N = f(x) = (\omega, \varphi(x)) + b \qquad (4)$$

where, $\omega$ is the weight vector; $\varphi(x)$ is the corresponding high-dimensional feature space; $b$ is the offset.

To achieve the goal of structural risk minimization, the hazard function is defined as follows:

$$L(y - f(x), x) = \begin{cases} |y - f(x, \omega)| - \varepsilon & (|y - f(x, \omega)| > \varepsilon) \\ 0 & else \end{cases} \qquad (5)$$

where, $\varepsilon$ is called loss parameter.

To train parameters $b$ and $\omega$, minimize the following function:

$$Rsvm(C) = \frac{1}{2}\|\omega\|^2 + \frac{C}{n}\sum_{i=1}^{n} L\varepsilon(d_i, y_i) \qquad (6)$$

where, $\frac{C}{n}\sum_{i=1}^{n} L\varepsilon(d_i, y_i)$ is experience error; C is called regularization parameter.
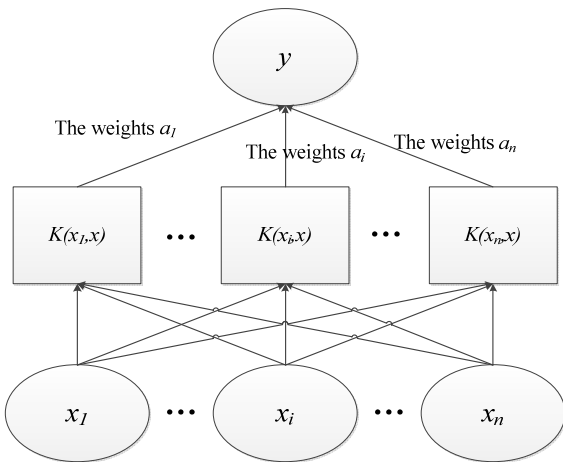


**Fig. 1.** Structure of SVM

To solve the optimization problem proposed upon, introduce slack variables $\zeta$ and $\zeta^*$. Then the objective function is shown as follows:

$$MinR_{svm}\left(\omega, b, \zeta_i, \zeta_i^*\right) = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}\left(\zeta_i, \zeta_i^*\right)$$
$$s.t. \begin{cases} y_i - f(x_i) \le \varepsilon + \zeta_i \\ f(x_i) - y_i \le \varepsilon + \zeta_i^* \\ \zeta_i, \zeta_i^* \ge 0 \end{cases} \qquad (7)$$

Use the Lagrange function to transform the objective function. The function is shown as follows:

$$\begin{aligned} L\left(\omega, b, \zeta_i, \zeta_i^*, \alpha_i, \alpha_i^*, r_i, r_i^*\right) &= \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}\left(\zeta_i + \zeta_i^*\right) \\ &- \sum_{i=1}^{n}\alpha_i\left[\zeta_i + \varepsilon - y_i + f(x_i)\right] \\ &- \sum_{i=1}^{n}\alpha_i^*\left[\zeta_i + \varepsilon + y_i - f(x_i)\right] \\ &- \sum_{i=1}^{n}\left(r_i\zeta_i + r_i^*\zeta_i^*\right) \end{aligned} \qquad (8)$$

where, $(i = 1, 2, 3...n)$; $n$ is the sample size; $\alpha_i, \alpha_i^*$, $r_i, r_i^* > 0$, and $\alpha_i, \alpha_i^*, r_i, r_i^*$ are all Lagrange multipliers.

Based on the extreme condition of the function $L$ upon and the equation $\phi(x_i)\phi(x) = k(x, x_i)$, the regression function is shown as follows:

$$f(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{m}(\alpha_i - \alpha_i^*) \cdot k(x, x_i) + b \qquad (9)$$

where, $k(x, x_i)$ is called kernel function.

In this paper, Gaussian function is selected as the kernel function. The kernel function is shown as follows:

$$k(x_i, x_j) = e^{\frac{\|x_i - x_j\|^2}{2\sigma^2}} \qquad (10)$$

where, $\sigma$ is Gauss parameter.

## 4. Short-term PV Power Generation Forecasting Model Based on GA-SVM

Genetic algorithm (GA) is a kind of optimization algorithm through simulating Darwin's biological evolution process. For this algorithm, any one group can be set as a starting point, and a new individual with better adaptability will be produced after operations of random selection, crossover and mutation. On the basis of the constraint, Groups will move toward a better search space to find out the individual with the highest adaptability through the survival of the fittest mechanism. This individual is the

optimal solution of the optimization problem [22-23]. It is faster to find the optimal solution of parameters after the optimization of the genetic algorithm.

## 4.1 The input and output of the forecasting model

The input data of the forecasting model includes two types, namely, environmental factors data of the predicted day, including the temperature and weather humidity of the sampling point before the predicted point (Sampling interval is 15min) from numerical weather prediction (NWP), and historical power output data, including the data of the predicted point and the data of the two adjacent time points before and after the predicted point (Sampling interval is 15min). For the boundary points, the historical power output data just includes the power data of two points. For the point of 6:00, the historical power output data includes the data of the predicted point 6:00 and the data of the point after the predicted point 6:15. And for the point of 19:00, the historical power output data includes the data of the point 19:00 and the data of the point 18:45.

The output data of the forecasting model is the power output of the PV system at every predicted point of the forecasting day (Sampling interval is 15min).In this paper, there are 53 predicted points.

The data of input and output of the prediction model is shown in Fig. 2. The input data of the forecasting model includes two types. And the sampling interval time is 15 min. The interval time of the output power is also 15 min. The detailed progress of GA-SVM model is shown in Fig. 2.

Where, $X_i^P$ is the predicted point of the predicted day; $X_i^S$ is the point of the similar day which is at the same moment with the point $X_i^P$ in the predicted day; $X_{i-1}^S$ is the point just before the point $X_i^S$; $X_{i+1}^S$ is the point just after the point $X_i^S$; $i = \{1, 2, \cdots, 53\}$.
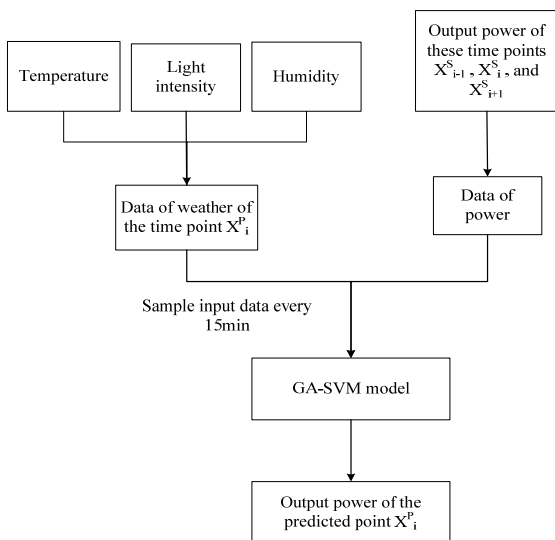


**Fig. 2.** The data of input and output of the prediction model

## 4.2 The main process of the forecasting model

Two parameters, namely regularization parameter and kernel parameter, need to be optimized in the short-term forecasting model based on SVM. Considering the advantages of GA, such as simple structure, strong generalization ability, and robustness, this paper uses GA to optimize model parameters, which can accelerate the convergence speed of the model. The main steps are shown as follows:

### 4.2.1 Determine the coding mode

To optimize regularization parameter and kernel parameter of the forecasting model is a complex continuous nonlinear optimization problem. For the binary encoding, the coding string is too long. It will decline the accuracy because of the reason that SVM is sensitive for initial parameters. While for the float encoding, the value of each individual gene is represented by a float within a certain range, namely the true value of the decision variable is used. This coding mode is suitable for the situation that the presented decision variable has a wide range. And the precision is also higher than that of binary encoding method. At the same time, this coding mode improves the computational complexity and operational efficiency of GA. So this paper selects the float encoding method as the coding mode of the forecasting model.

### 4.2.2 Generate the initial population

If the scale of the initial population is large, it is possible to maintain the genetic diversity of the population to some extent. And the predicted model will have a stronger global optimization capability. But the cost is that the iterative calculation time increases. While if the scale of the initial population is too small, it may lead to local minima problems, although it will have a fast convergence rate. In this paper, for the GA-SVM model, parameters to be optimized are regularization parameter $C$ and Gauss parameter $\sigma$. We should generate the initial population for them. After a series of simulation, we obtain the size of the generation. And then, the initial population is generated. The initial population of the parameter $C$ is randomly generated between 0 and 1, and the initial population of the parameter $\sigma$ is randomly generated between 0 and 100.

### 4.2.3 Select the fitness function

SVM desires the output of the forecasting model becoming excellent in the training process. This algorithm requires the sum of squared errors between the predicted output and the actual output of the system less than a preset accuracy, and the smaller the better. So the fitness function is shown as follows:
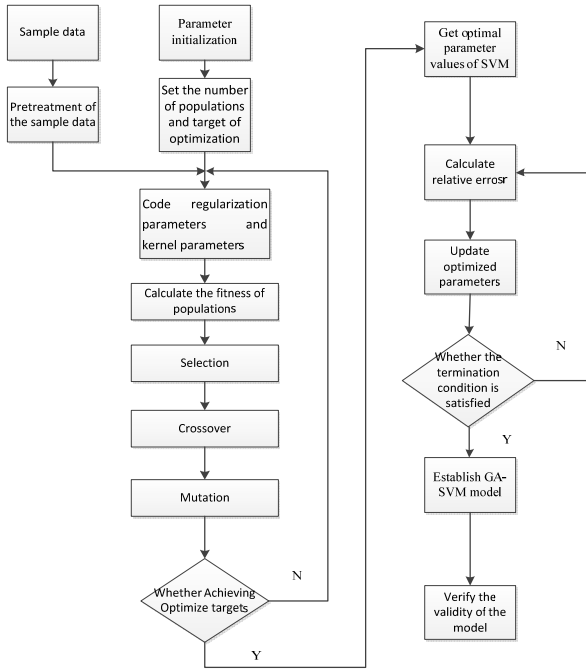
**Fig. 3.** The flow chart of the GA-SVM model

$$f = \frac{1}{K+1}$$
$$K = \frac{1}{2}\sum_{i=1}^{l} e(i)^2 \qquad (11)$$
$$e(i) = y(i) - y_m(i)$$

where, $l$ is the learning sample; $y(i)$ is the actual power output; $y_m(i)$ is the expected power output; $e(i)$ is the error between $y(i)$ and $y_m(i)$.

The flow chart of the GA-SVM model is shown in Fig. 3.

### 4.2.4 Optimize the model parameter

Optimize the regularization parameter and kernel parameter of the model by operations of selection, crossover and mutation. To avoid gene missing and improve the global convergence of the algorithm, this paper uses the fitness of individuals to perform selection operation. And the specific operation is as follows: To find out the highest degree of fitness of the individual and the lowest individual in the current population; if the current best individual's fitness is even higher than the overall best individual's fitness, replace the overall best individual as the current best individual and take the place of the worst individual in the current population as the overall best individual. The specific operation of the crossover operation is as follows: First, select a point as the intersection randomly in the chromosome. Then combine a new chromosome with the string in front of the first father chromosome's intersection and the string after the second father chromosome's intersection. And combine another

chromosome with the string before the second father chromosome's intersection and the string after the first father chromosome's intersection. The mutation operation changes the character value of some chromosome gene with a certain probability to maintain the diversity of species.

## 5. Example Verification

Based on the measured power data of Key Laboratory of Smart Grid of Ministry of Education in Tianjin University and the weather data from NWP, establish and verify the short-term power forecasting model of the PV system. In this paper, we establish the forecasting model to predict the 24 hours-ahead power output.

Through classifying the original weather category into three types, this paper establishes different forecasting models according to corresponding weather type.

$C_{best} = 0.8$, Gauss parameter $\sigma_{best} = 24.5$. Enter the parameter value into the forecasting model. Then calculate the power output of the PV system after sample training. Finally compare the predicted power with the actual power output and the power output of the forecasting model based on SVM without the optimization algorithm. The results are shown in Fig. 4, Fig. 6 and Fig. 8 corresponding to the three weather type. And relative errors of the two forecasting models are shown in Fig. 5; Fig. 7 and Fig. 9.

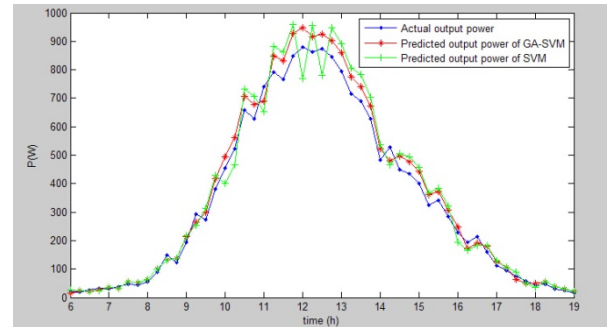The way to calculate relative error (*RE*) is shown as



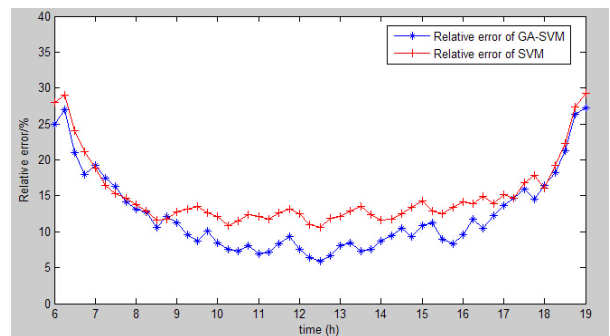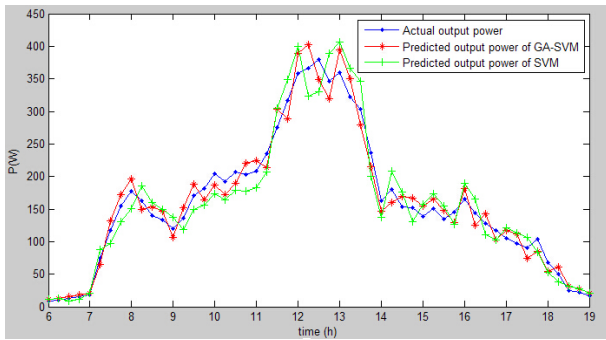**Fig. 4.** The actual power and forecasting power in the weather type A
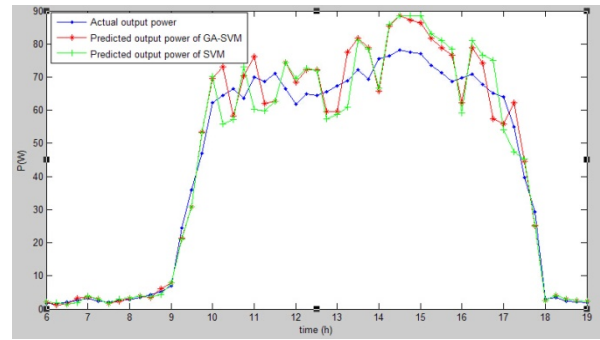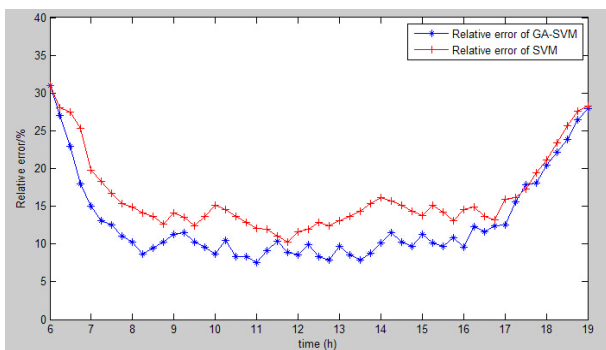


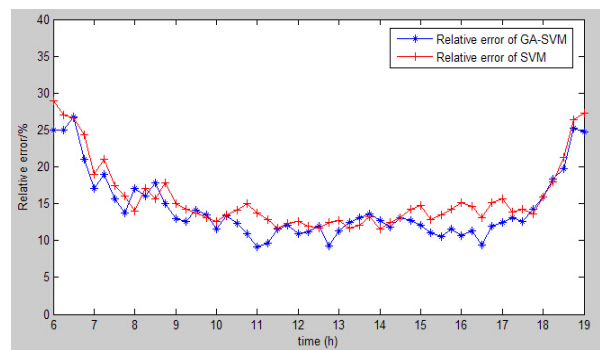**Fig. 5.** Relative errors in the weather type A

**Fig. 6.** The actual power and forecasting power in the weather type B



**Fig. 8.** The actual power and forecasting power in the weather type C



**Fig. 7.** Relative errors in the weather type B



**Fig. 9.** Relative errors in the weather type C

**Table 2.** The relative errors in different condition

|  | Type A (Sunny) | Type B (Cloudy) | Type C (Rainy) |
|---|---|---|---|
| GA-SVM | 9.28% | 11.04% | 13.23% |
| SVM | 12.25% | 16.73% | 15.21% |

follows:

$$RE = \frac{\left| P_{actual} - P_{priditated} \right|}{P_{\text{actual}}} \times 100\% \qquad (12)$$

Through analyzing Fig. 4 and Fig. 5, the forecasting result of the model based on GA-SVM is better in the condition of weather type A with the smooth change of the solar radiation intensity and temperature. The overall relative error is maintained within the range of 10%, while the error of the SVM model is within the range of 10% to 15%. It is obvious that the improved SVM model has better prediction accuracy.

From Fig. 6 and Fig. 7, the forecasting result of the GA-SVM model is also better in the condition of type B with a certain volatility of the solar radiation intensity and temperature. The relative error of the GA-SVM model is maintained within the 10% - 15% range, while the error of the SVM prediction model in the range of 15% - 20 %. As a result, the prediction accuracy of the improved SVM model is higher.

As it can be seen from Fig.8 and Fig.9, in the condition of weather type C with low solar radiation intensity and temperature, the power output is low and the relative error is larger than that of the weather type A and B. From Fig.9, we can see the predicted fluctuation range is greater and the relative error is maintained in the range of 10% - 20%.

But the GA-SVM forecasting model also has better prediction accuracy than the SVM model.

Meanwhile, from Fig. 5, Fig. 7 and Fig. 9, the forecasting error of the model is very large when the actual power output is small, for example in periods of 6:00 - 8:00 and 17:00 - 19:00. The main reason is that environmental factors such as haze reduce atmospheric transparency. As a result, the prediction accuracy is relatively high with small effect in the period of 8:00 - 17:00.

To compare the accuracy of SVM and GA-SVM algorithms in different weather condition, we list the relative errors of two models in three types of weather in Table 2.

It can be seen from Table 2 that the forecast result of GA-SVM model is more accurate than that of SVM no matter what the weather is.

## 6. Conclusions

In order to effectively alleviate many of the adverse impacts of grid scheduling, energy management and

other aspects affected by the PV system, the short-term prediction of photovoltaic power generation is necessary. This paper establishes the forecasting model of PV system based on the improved support vector machine algorithm (GA-SVM). Through classifying the original weather category into three types, this paper establishes different forecasting models according to corresponding weather type. To improve the prediction accuracy of the forecasting model, this paper uses genetic algorithm to optimize the parameters of the model. Finally, through the comparison of forecasting results between the GA-SVM model and the SVM model, it is shown that the proposed algorithm and forecasting model are effective and promising.

## Acknowledgements

## References

[1] Jing, L. I., et al. "Dynamic Modeling and Simulation of the Grid-connected PV Power Station." *Automation of Electric Power Systems* vol. 32, no. 24, pp. 83-86, Dec. 2008.

[2] Kong, Lingzhi, X. Tang, and Z. Qi. "Study on modified EMAP model and its application in collaborative operation of hybrid distributed power generation system." *International Conference on Sustainable Power Generation & Supply IEEE*, Nanjing, China, 2009.

[3] Monai, T., et al. "Response characteristics and operating methods of new type dispersed power supply system using photovoltaic fuel cell and SMES." *Power Engineering Society Summer Meeting*, Chicago, USA, 2002.

[4] Guo XY, et al. "A time-interval based probabilistic production simulation of power system with grid-connected photovoltaic generation." *Power System Technology*, vol. 37, no. 6, pp. 1499-1505, Jun.2013.

[5] Zhang, Yan Xia, and J. Zhao. "Application of recurrent neural networks to generated power forecasting for photovoltaic system." *Power System Protection & Control*, vol. 39, no. 15, pp. 96-101, Aug. 2011.

[6] Hiyama, T., and K. Kitabayashi. "Neural network based estimation of maximum power generation from PV module using environmental information." *IEEE Transactions on Energy Conversion*, vol. 12, no. 3, pp. 241-247, Aug.1997.

[7] Ming, Ding, and X. U. Ningzhou. "A Method to Forecast Short-Term Output Power of Photovoltaic Generation System Based on Markov Chain." *Power System Technology*, vol. 35, no. 1, pp. 152-157, Jan. 2011.

[8] Li, Ying Zi, R. Luan, and J. C. Niu. "Forecast of power generation for grid-connected photovoltaic system based on grey model and Markov chain." *Power and Energy Engineering Conference*, Wuhan, China, 2009.

[9] Bessa, R. J., et al. "Solar power forecasting in smart grids using distributed information." *International Journal of Electrical Power & Energy Systems*, vol. 72, pp. 16-23, Aug.2015.

[10] Yona, Atsushi, et al. "Application of Neural Network to One-Day-Ahead 24 hours Generating Power Forecasting for Photovoltaic System." *International Conference on Intelligent Systems Applications To Power Systems*, Niigata, Japan 2007.

[11] Yona, A., T. Senjyu, and T. Funabashi. "Application of Recurrent Neural Network to Short-Term-Ahead Generating Power Forecasting for Photovoltaic System." *IEEE Power Engineering Society General Meeting*, Florida, USA, 2007.

[12] Yang, Ying, and L. Dong. "Short-Term PV Generation System Direct Power Prediction Model on Wavelet Neural Network and Weather Type Clustering." *International Conference on Intelligent Human-Machine Systems and Cybernetics IEEE,* Hangzhou, China, 2013.

[13] Wang, Shou Xiang, and Zhang Na. "A model to forecast short-term output power of photovoltaic system based on grey and neural theory." *Automation of Electric Power Systems*, vol. 36, no. 19, pp. 37-41, Oct.2012.

[14] Dong Lei, Zhou WP, Zhang P, et al. "Short-term output forecast based on dynamic Bayesian network theory." *Proceeding of the CSEE*, vol. 33, pp. 38-45, Jan.2013.

[15] Felice, Matteo De, M. Petitta, and P. M. Ruti. "Short-term predictability of photovoltaic production over Italy." *Renewable Energy*, vol. 80, pp. 197-204, Sep. 2015.

[16] Marcelo Pinho Almeida, Oscar Perpinan, Luis Narvarte. PV power forecast using a nonparametric PV model. *Solar Energy*, 2015, 115: 354-368

[17] Alonso-Montesinos, J., and F. J. Batlles. "Solar radiation forecasting in the short- and medium-term under all sky conditions." *Energy*, vol. 83, no. 4, pp. 387-393, Mar.2015.

[18] Hui, Lin, et al. "Short-term load forecasting for holidays based on the similar days' load modification." *Power System Protection and Control*, vol. 38, no. 7, pp. 47-51, Apr.2010.

[19] Wang, Xiaolan, and G. E. Pengjiang. "PV array output power forecasting based on similar day and RBFNN." *Electric Power Automation Equipment*, vol. 33, no. 1, pp. 100-103, Jan.2013.

[20] Lei, Huang, et al. "Photovoltaic Generation Forecast

Based on Multidimensional Time-series and Local Support Vector Regression in Microgrids." *Automation of Electric Power Systems*, vol. 38, no. 5, pp. 19-23, Mar.2014.

[21] Shi, Jie, et al. "Forecasting power output of photo-voltaic system based on weather classification and support vector machine." *IEEE Transactions on Industry Applications*, vol. 48, no. 3, pp. 1-6, May. 2012.

[22] Wang KQ, Yang SC, Dai TH, et al. "Method of optimizing parameter of least squares support vector machines by genetic algorithm." *Computer Application and Software*, vol. 26, no. 7, pp. 109-111, Jul. 2009.

[23] Golpîra, H., H. Bevrani, and H. Golpîra. "Application of GA optimization for automatic generation control design in an interconnected power system." *Energy Conversion & Management*, vol. 52, no. 5, pp. 2247-2255, May. 2011.

**Zhilin Song** He received his B.S. degree and M.S. degree from Hebei University and Tianjin University in 2013 and 2015 respectively. His research interest is photovoltaic generation forecast.



**Jiawen Sun** He received his B.S. and M.S. degree from Tianjin University in 2012 and 2014 respectively. His research interest is distributed generation forecast.



**Jidong Wang** He received his B.S. and M.S. degrees from Shandong University of Technology and Shandong University in 1999 and 2002 respectively. He received his Ph.D. degree from Tianjin University in 2005. He worked as a post-doctoral from 2005 to 2007 in Tianjin University. He is now an Associate Professor of School of Electrical Engineering and Automation in Tianjin University since 2007. His research interests include power quality, distributed generation system, microgrid and smart power consumption.



**Ran Ran** She received her B.S. degree from Tianjin University in 2015. She is now a postgraduate student in Tianjin University since 2015. Her research interest is photovoltaic generation forecast.