Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Photovoltaic power generation forecasting model based on IFCM

Yanjuan Ma

myj_2016@163.com

Criminal Investigation Police University of China

Xueao Jiang

Criminal Investigation Police University of China

Mai Jiang

Criminal Investigation Police University of China

Article

Keywords: Photovoltaic power prediction, IForest algorithm, Least squares support vector machine, Particle swarm optimization

Posted Date: March 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8614486/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Photovoltaic power generation forecasting model based on IFCM

Yanjuan Ma[1*],Xueao, Jiang[1] and Mai Jiang[1, 2]
[1]Criminal Investigation and Counter-Terrorism College, Criminal Investigation Police University of China, Shenyang 110854, China.
[2]Key Laboratory of Trace Inspection and Identification Technology, Ministry of Public Security, Shenyang, China.
Word Count:

Article submitted to:

Version:

Corresponding author: Criminal lmvestiggtion and Counler-Terrrism Colege, Criminal lnvestigation Police University of China, Shenyang, 110854, China
E-**mail:** myj_2016@163.com

Abstract

The forecasting of photovoltaic power generation is of important meaning to both security and stability runs for the electrified wire netting. In this paper, a photovoltaic energy generation forecasting model is proposed called IFCM, which is combined with isolated forest (IForest), fuzzy c-means, and support vector machine (SVM). Firstly, the data is cleaned through the isolation forest anomaly detection algorithm; Secondly, the historical data of similar days is used for Training. With the isolation forest anomaly detection algorithm, the proposed IFCM has achieved great results. In the experimental analysis, the proposed method is compared with the prediction result of the traditional IForest-SVM and IForest-BP models. The experimental results demonstrate that the proposed IFCM algorithm exhibits superior performance.

Key words: Photovoltaic power prediction; IForest algorithm; Least squares support vector machine; Particle swarm optimization

1. INTRODUCTION

With the increasing growth of social processes, the utilization of electricity has risen significantly. Considering both the concept of green environmental protection and economic growth, the use of green energy has become the current development situation. For example, utilizing solar power generation can provide electrical support for police centers or police equipment. Among them, solar energy has the advantage of being inexhaustible and easy to obtain, which makes forecasting of photovoltaic power generation favored. Still, it also has the characteristics of fluctuation and instability, and electrified wire netting has trouble in the security and operational aspects. Therefore, promoting the precision of the forecasting of photovoltaic power generation is of large influence on the operational aspect of the electrified wire netting[1].So far, there have been a lot of studies on photovoltaic power generation forecasting; the main methods are direct prediction[2] and indirect prediction[3]. Among them, indirect prediction is based on a physical model, which includes all-day image predictions, the way of satellite imagery prediction, etc. If the weather conditions are great, the physical model can achieve better prediction results. However, its process is much more complicated in real conditions. The direct prediction method directly relies on historical data for prediction and uses artificial intelligence algorithms to build a neural network model for direct prediction, and its prediction cost is also low, so it's more widely used in the field of forecasting of photovoltaic power generation. It is mainly classified as time series prediction method[4], the BP neural networks[5], the support vector machine(SVM), etc.[6].

The machine learning method provides an effective way to accurately analyze the influencing factors of the forecasting of photovoltaic power generation and improve the forecasting precision of the forecasting of photovoltaic power generation. In the process of prediction, the precision of historical data often affects the precision of the forecasting result, so the data preprocessing can reduce the impact of predicting outlier results to some degree. In recent years, the ways of data outlier processing include the optimal variance algorithm[7], the combined quartile method[8], and Isolation Forest (IForest)[9]. The algorithm can distinguish abnormal operation of wind turbines for the data identification of wind turbines. However, its recognition efficiency is not high. The quartile method first counts the historical data and then analyzes it. Although it can effectively detect abnormal data, the processing is complicated. Among many anomaly detection algorithms, the iForest algorithm has the strength of operability, a great detection workpiece ratio, and high accuracy, so the isolation forest

algorithm is more widely used. Algorithms for clustering and analysis of data include the gray correlation coefficient method, the K-means algorithm[10], fuzzy clustering method[11], etc. Among them, the gray correlation coefficient method is too subjective, and it is difficult to reflect different weather factors' influence objectively and accurately on photovoltaic power. The selection of the K value of the K-means algorithm is random and cannot accurately classify the data. The fuzzy C-means clustering method adopts the idea of optimizing the objective function to automatically classify the data [12].

The traditional photovoltaic prediction method only uses optimization algorithms to improve the model to different degrees, but does not consider the importance of different samples to the model during model training, resulting in low model accuracy. And the time interval for photovoltaic power prediction is usually 1h, which cannot meet the higher scheduling requirements. Aiming at these problems, the article presents award support vector machine algorithm model based on the fuzzy c-means learning algorithm. In order to take full account of the similarity between historical power generation data and forecast data.

2. METHODS

2.1. Data Cleaning Based on Forest Algorithm

Abnormal operation of the distribution network, poor communication, and system failure will lead to abnormal measurement data, so the data needs to be pre-processed before making predictions. As an anomaly detection algorithm, the isolation forest is sensitive to global sparse points and can better filter outlier data. The test process of the isolated forest is shown in Figure 1. The abnormal score is related to the depth of the IForest tree. When the depth of the sample is smaller, the abnormal score is higher, and vice versa. The IForest algorithm is divided into two stages. It needs to build an isolated forest in the first stage, which consists of the trees. The steps for building an itree are as follows:

Step1. Randomly select *v* sample points from the training data set to form a sub-sample set *x*, and then put all the sub-sample sets into the root node, the current tree height is *a*, and the height is limited to 1.

Step2. Select a dimension *q* as an eigenvalue in the selected sub-sample set, and randomly determine a split point *p* between the maximum and minimum dimensions of the dimension, and divide the sub-data set *x* into two parts by comparing the sizes of *p* and *q* A space, the eigenvalue *q* less than *p* is stored in the left half space, and the eigenvalue *q* greater than *p* is stored in the right half space, so a hyperplane is also obtained.

Step3. It is iteratively constructed in child nodes to construct new child nodes until the data can no longer be split or the height of the tree reaches 1.

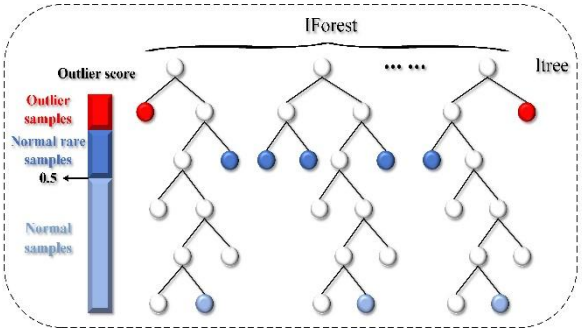


Figure 1 Output of abnormal values in iForest

The second stage is the judgment of the degree of abnormality, calculating the abnormality score of the detected sample. The definition of the abnormality score is as shown in the following formula:

$$s(x) = 2^{-\frac{E[h(x)]}{c(\varphi)}} \tag{1}$$

In the above formula, *h(x)* represents the path length of the detected sample in the iTree, *E( )* represents the average of *t* iTrees, and *c(v)* represents the average of the binary tree constituded of *v* points in the training data to search path. Among them, the formula of *c(v)* is as follows:

$$c(v) = 2h(v-1) - \left(\frac{2(v-1)}{v}\right) \tag{2}$$

Where, $h(x) = \ln(x) + \zeta$, ζ is Euler's constant. In From formula (1), there are three conditions:

1) $S(x) \to 1$, all examples are abnormal points at this time;

2) $S(x) \to 0$, there are no outliers in the sample;

3) $S(x) \to 0.5$, outliers in the sample are not obvious.

2.2. Select the similar day

The Pearson correlation coefficient method, it is used to react the degree of linear relevance with two variables, usually expressed by *r*, and its expression is as follows[13]:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{3}$$

In the above formula, *n* represents the number of a certain weather type; *x* represents the weather factor; *y* is the solar energy generation; $\bar{x}$ and $\bar{y}$ represent the two variables.

The γ is between [-1, 1], the size of the γ reflects the strength of the linear correlation with the two variables, and larger the absolute values. The correlation coefficient between photovoltaic power and weather factors as shown in Table 1.

Table 1 Correlation coefficient between photovoltaic power and weather factors

Factor	Correlation coefficient
Light intensity	0.97
Temperature	0.41
Wind speed	0.22
Rainfall	-0.05

3. CONSTRUCTION OF FCM-SVM PHOTOCOLTAIC POWER PREDITICTION MODEL

3.1. Fuzzy C-Means

The fuzzy c-means clustering method (FCM) is described as follows [14]. The clustering loss function is defined as below.

$$J(U,V)=\sum_{i=1}^{n}\sum_{j=1}^{c}u_{ij}^{m}d_{ij}^{2} \tag{4}$$

In the above formula, u_{ij} is the membership degree of the ith sample belonging to the jth class, U is the membership degree matrix, V is the cluster center, m is the membership degree factor $1\le m<\infty$, d_{ij} represents the sample to cluster Distance from class center. The central idea of fuzzy C-means is to continuously iteratively calculate U and V to minimize $J_{(U,V)}$. The specific steps are as follows:

Step 1. Number of initialized clusters c and the membership matrix $U^{(0)}$. Let $l=1$, which means the first iteration.

Step 2. Calculate the 1_{th} cluster center $V^{(l)}$:

$$v_j^{(l)}=\frac{\sum_{i=1}^{n}\left(u_{ij}^{(l-1)}\right)^{m}x_i}{\sum_{i=1}^{n}\left(u_{ij}^{(l-1)}\right)^{m}},\ j=1,2,\cdots,c \tag{5}$$

Step 3. Update the membership matrix $U^{(l)}$, and Figureure out the clustering loss function $J^{(l)}$.

$$u_{ij}^{(l)}=\frac{1}{\sum_{k=1}^{c}\left(\frac{d_{ij}^{(l)}}{d_{ik}^{(l)}}\right)^{\frac{2}{m-1}}},i=1,\cdots c;\ j=1,\cdots n \tag{6}$$

$$J^{(l)}\left(U^{(l)},V^{(l)}\right)=\sum_{i=1}^{n}\sum_{j=1}^{c}\left(x_{ij}^{l}\right)^{m}\left(d_{ij}^{l}\right)^{2} \tag{7}$$

Where, $d_{ij}^{(l)}=\left\|x_i-v_j^{(l)}\right\|$.

Step 4. Given the membership planting threshold $\varepsilon_u>0$, the loss function terminates the threshold $\varepsilon_J>0$, and stops the iteration when $\max\left\{\left|x_{ij}^{(l)}-u_{ij}^{(l-1)}\right|\right\}<\varepsilon_u$ or $\left|J^{(l)}-J^{(l-1)}\right|<\varepsilon_J$, otherwise returns to step 2.

Through the above steps, the most membership degree matrix *U* and cluster center matrix *V* can be obtained. According to *U*, the category to which the sample belongs can be determined. When the sample is classified as the j_{th} category.

In this paper, FCM is used to divide the historical daily weather data and weight the training samples based on the membership matrix. In order to make the clustered categories have similar weather

characteristics and output power characteristics, vertical power data is used as the clustering feature. Perform statistical analysis on historical output data and perform maximum-minimum normalization on the calculation results.

3.2. SVM algorithm

Support vector machines (SVM) is a binary classification model. The specific calculation process is as follows:

The $dataset = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$, $x_i \in R^d$ is input value, $y_i \in R$ is target value, *i=1, 2, ..., n,* through the linear regression function.

$$f(x) = w^T x + b \tag{8}$$

To fit the sample (x_i, y_i) , using the loss function ε

$$|y - f(x)|_\varepsilon = \begin{cases} 0 & |y - f(x)| \le \varepsilon \\ |y - f(x)| - \varepsilon & |y - f(x)| > \varepsilon \end{cases} \tag{9}$$

The penalty factor *C* and relaxation factor ξ_i , ξ_i^* is used to solve the optimization problem is

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \tag{10}$$

$$s.t.\ f(x_i) - y_i \le \varepsilon + \xi_i$$

$$y_i - f(x_i) \le \varepsilon + \xi_i^*$$

$$\xi_i \ge 0, \xi_i^* \ge 0, i = 1, 2, \cdots, n$$

Lagrangian multipliers α_i , α_i^* , μ_i , μ_i^* is used to obtain the Lagrangian function of the optimization problem as.

$$L = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) - \sum_{i=1}^{n}\mu_i\xi_i + \sum_{i=1}^{n}\alpha_i^*(y_i - f(x_i) - \varepsilon - \xi_i^*) \tag{11}$$

Take the partial derivatives of L with respect to *w*, *b*, ξ_i and ξ_i^* , so that each partial derivative is 0, we can get

$$w = \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)x_i \tag{12}$$

$$0 = \sum_{i=1}^{n}(\alpha_i^* - \alpha_i) \tag{13}$$

$$C = \alpha_i + \mu_i \tag{14}$$

$$C = \alpha_i^* + \mu_i^* \tag{15}$$

Taking the above equation (9) into Eq. (11), the dual problem of equation (10) can be obtained.

$$\max_{\alpha,\alpha^*} \sum_{i=1}^{n} y_i\left(\alpha_i^* - \alpha_i\right) - \varepsilon\left(\alpha_i^* + \alpha_i\right) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\alpha_i^* - \alpha_i\right)\left(\alpha_j^* - \alpha_j\right)x_i^T x_j \tag{16}$$

$$s.t. \quad \sum_{i=1}^{n}\left(\alpha_i^* - \alpha_i\right) = 0,$$

$$0 \le \alpha_i, \alpha_i^* \le C$$

Set Eq. (12) into Eq. (8) to obtain a regression model, such as Eq. (17) shows.

$$f(x) = \sum_{i=1}^{n}\left(\alpha_i^* - \alpha_i\right)x_i^T x + b \tag{17}$$

If the kernel function is used instead of the inner product calculation, the regression model can be expressed as

$$f(x) = \sum_{i=1}^{n}\left(\alpha_i^* - \alpha_i\right)k\left(x, x_i\right) + b \tag{18}$$

3.3. IFCM model construction

In this paper, the proposed Method IFCM combines the algorithm of SVM to upgrade the forecasting precision from the perspective of refining data. The isolated forest method is used for the historical data set to filter out abnormal data in the sample; Considering the effect of weather elements on the forecasting of photovoltaic power generation prediction, using the FCM to classify the weather into three common weather types: sunny, cloudy and rainy; Determine the weather conditions with forecast days as a reference and select the homologous meteorological data as a training set sample; Use the SVM algorithm sample for accurate prediction, as shown in Figure.2 The concrete prediction process is shown in Figure 3, and the specific steps are shown below.

1. From the perspective of data mining, first process the abnormal data in the historical data: use the isolated forest algorithm to clean the historical data, and calculate the abnormal index to filter out the abnormal points;
2. Photovoltaic power generation varies due to different weather conditions. The FCM algorithm is used to classify the meteorological data, which are classified as 3 weather types: sunny days, cloudy days, and rainy days.
3. Taking the weather conditions with forecast days as a reference, the data samples that react to the weather type are selected as training examples;

Using the SVM algorithm to train the samples, which highlights the importance of similar day samples in model training, so that the model can learn more useful knowledge from similar day samples during training.

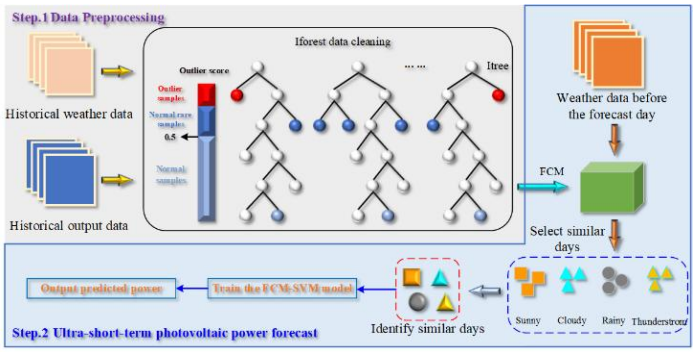


Figure 2 Short-term photovoltaic power generation power forecast block diagram

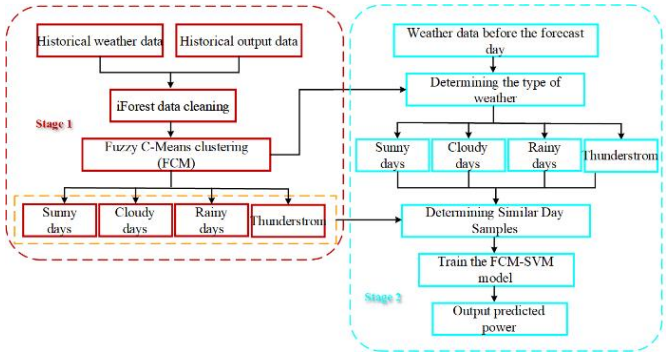


Figure 3 Short-term photovoltaic power forecast flow chart

4. Simulation Examples

Depending on the type of weather, this paper predicts sunny days, cloudy days, and rainy days. The historical data of June and July are selected in a certain place as the experimental data. Three days of sunny, cloudy, and rainy weather types were selected as the forecast days, and three sets of measurement data of the city's photovoltaic panel temperature, light intensity, and photovoltaic power generation were also selected. The sampling period is 5 minutes, and the forecast period is 24 hours, with 288 sampling points per day. And select the data of the most similar 5 days from the corresponding weather type as its similar days. The selected similar days are 15 days in total, a total of 2160 training data. Select 3 days as the forecast day, 432 sample data. First, the isolated forest algorithm is used to preprocess similar daily data, and finally, 2093 data points are obtained. Figure 4 shows the photovoltaic output of a photovoltaic power station under different weather types. Table 2 specifically shows the selection of similar days and forecast days. And the potential weather information is obtained from historical data.

4.1. Data Normalization and Model Evaluation Metrics

Before training the SVM support vector machine, it is necessary to normalize the data of similar days and prediction days. This paper adopts the normalization method of the following formula.

$$z_i = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}} \tag{19}$$

Where y_i is the original data, $y_{\min}$ and $y_{\max}$ is the minimum and maximum value and z_i is the initialization data.

The evaluation index analysis model adopts mean absolute percentage error (MAPE) and root mean square error (RMSE). The formula is as follows:

$$MAPE = \frac{1}{K}\sum_{i=1}^{K}\left(\left|\frac{P_{ti} - P_{prei}}{P_{ti}}\right| \times 100\%\right) \tag{20}$$

$$RMSE = \sqrt{\frac{1}{K}\sum_{i=1}^{K}\left(P_{ti} - P_{prei}\right)^2} \tag{21}$$

Where P_{ti} is the actual value of power; P_{prei} is the predicted value of power; K is the amount of data.

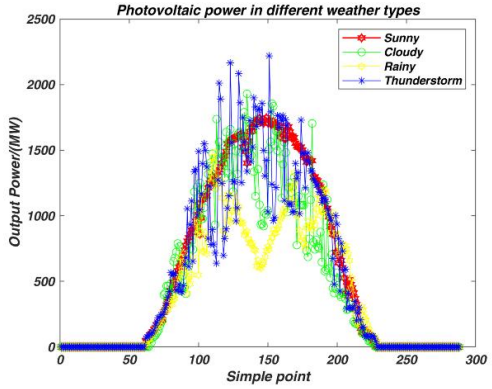


Figure 4 Photovoltaic output diagram under different weather conditions

Table 2 Similar days and forecast days

Features	Sunny	Cloudy	Rainy
Similar days	June 3	June 6	June 25
	June 4	June 8	June 26
	June 9	June 12	July 9
	June 10	June 18	July 15
	June 11	July 2	July 18
Forecast days	June 5	June 15	July 1

4.2. Example analysis and comparison

As can be seen from the previous analysis, similar days and forecast days are classified as 4 types in accordance with the weather type: sunny, rainy, and cloudy. After processing the abnormal data, using the FCM, classify the data. Under these three weather conditions, an SVM prediction model (model A)

is established from the clustered data, and the meteorological data of similar days and forecast days are used as the input of the model to make predictions, and the one-day forecast is synthesized according to the time point data. To verify validity of the model put award in the article, a non-clustered SVM (model B) prediction comparison model was added, and then a comparison experiment was performed with the BP neural network prediction model (model C).LSTM(model D)

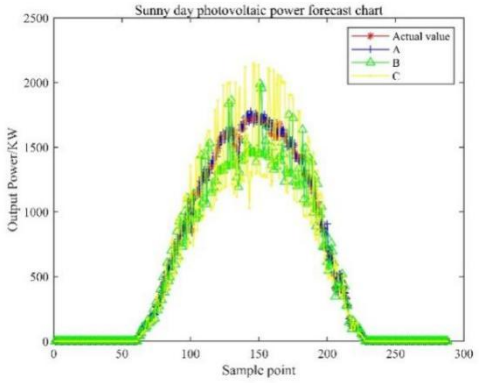


Figure 5 Sunny day photovoltaic power forecast chart

Figure 5 shows the sunny day photovoltaic power forecast. It can be seen from the Figure that the curve of the predicted value and the actual value is relatively close, and the prediction effect is better in this weather condition. The changes in the forecasting of photovoltaic power generation also have certain rules, and the prediction model was proposed in the three models. Figure 6 shows the cloudy day photovoltaic power forecast, and Figure 7 shows the rainy day photovoltaic power forecast. From the results, it can be seen in the figure that the forecasting accuracy of the model mentioned in the text is higher than the other two models, and there are some points that are insufficient for accurate prediction, possibly due to unstable weather conditions and limited samples. Using the two error indicators in the previous section, Table 3 shows the forecast error indicators for the forecast days. Figure 8 shows the prediction results of four models for thunderstorm weather types. It can be seen from the figure that the model proposed in this paper achieves the best performance, followed by LSTM, then SVM, and BP ranks last.

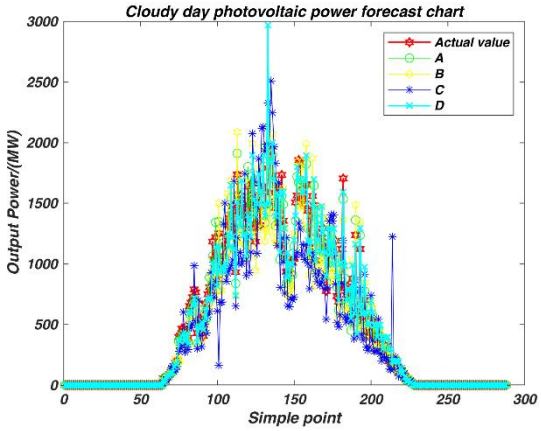


Figure 6 Cloudy photovoltaic power forecast chart

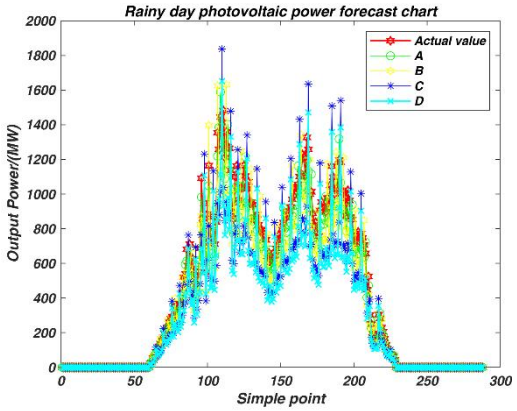


Figure 7 Rainy photovoltaic power prediction chart

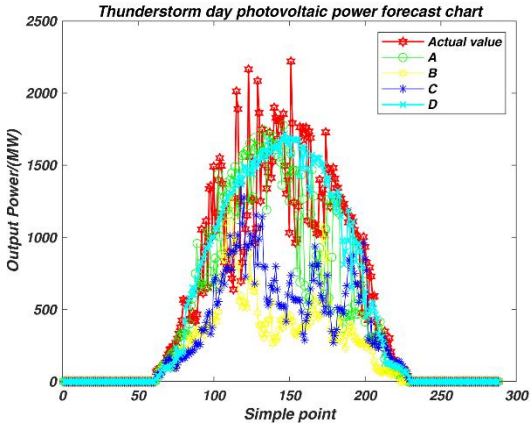


Figure 8 Thunderstrom photovoltaic power prediction chart

As it can be seen from Table.3 (the following model A represents the IFCM. The model B is the iForest-SVM prediction model, and model C is the iForest-FCM-BP prediction model), With the MAPE value of model A is 3.74% on sunny days , model B is 4.81%, model C is 11.9%; The EMSE values for Model A were 0.13kW, Model B was 0.24kW, and Model C was 0.32kW. Predictions for sunny weather are more accurate than predictions for cloudy and rainy weather. When it is cloudy, the MAPE value of Model A is 12.23%, that of Model B is 19.26%, and that of Model C is 27.35%; The EMSE values for Model A were 0.17 kW, 0.31 kW for Model B, and 0.41 kW for Model C. In rainy weather, the MAPE value of model A is 16.27%, that of model B is 24.33%, and that of model C is 34.46%; The EMSE values for Model A were 0.22kW, Model B was 0.28kW, and Model C was 0.36kW. From the above simulation and data analysis, it can be concluded that the prediction model of Model A gets the best prediction results.

Table 3 Comparison table of model errors under different weather types

Weather type	A		B		C		D	
	E_{MAPE}	E_{RMSE} /kW	E_{MAPE}	E_{RMSE} /kW	E_{MAPE}	E_{RMSE} /kW	E_{MAPE}	E_{RMSE}

	/%		/%		/%		/%	/kW
Sunny	3.47	0.13	4.81	0.24	6.20	0.32	4.52	0.17
Cloudy	12.23	0.17	19.26	0.31	27.35	0.41	13.26	0.22
Rainy	16.27	0.22	24.33	0.28	34.46	0.36	15.22	0.26
Thunderstorm	13.23	0.15	20.23	0.26	25.26	0.37	17.26	0.27

5. Conclusion

In this paper, an algorithm called IFCM is proposed to solve the photovoltaic energy generation forecasting problem. And in the IFCM, the isolation forest is added to solve the anomaly detection problem, which has great performance. Since the IFCM combined fuzzy c-means with a support vector machine, it has obtained good results in the problem of photovoltaic energy generation forecasting. The experimental result shows that the IFCM is better than the traditional SVM and BP models.

Author Contribution declaration

In this paper, author 1（Yanjuan Ma） was responsible for proposing the research idea, designing the overall framework of the manuscript, and conducting the simulation experiments. Author 2（Xueao Jiang） contributed to the refinement and editing of the full text. Author 3(Mai Jiang) assisted in performing the simulation experiments.

Data availability

The data that support the findings of this study are available on request from the corresponding author [Yanjuan Ma] upon reasonable request.

Funding

This work was financially supported by Youth Project of the Fundamental Research Project for Universities of Liaoning Province (Grant No. JYTQN2023092).

REFERENCES

1. CUI Yang, ZHANG Huiquan, ZHONG Wuzhi, et al. Multi-source optimal scheduling of renewable energy high-permeability power system with CSP plants considering demand response[J]. High Voltage Engineering, 2020, 46(5): 1499-1507.
2. WANG Jituo, WANG Wancheng, CHEN Hongwei. Photovoltaic power generation forecasting based on regression-Markov chain[J]. Electrical Measurement & Instrumentation, 2019, 56(1): 76-81.
3. ZHOU Y, ZHOU N R, GONG L H, et al. Prediction of photovoltaic power output based on similar day analysis, genetic algorithm and extreme learning machine[J]. Energy, 2020, 204: 117894.
4. Sun H , Wang S , Jiang Q .FCM-Based Model Selection Algorithms for Determining the Number of Clusters[J].Pattern Recognition, 2004, 37(10):2027-2037.
5. Yu J , Cheng Q , Huang H .Analysis of the weighting exponent in the FCM.[J]. IEEE Transactions on Systems Man & Cybernetics Part B, 2004, 34(1):634-639.
6. Bezdek J C , Ehrlich R , Full W .FCM: The fuzzy c -means clustering algorithm[J]. Computers & Geosciences, 1984, 10(2–3):191-203.
7. A D W V D M , A M S , B A S ,et al.Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using Gaussian Processes[J]. Applied Energy, 2018, 213:195-207.
8. Souhe F G Y , Mbey C F , Kakeu V J F ,et al.Optimized forecasting of photovoltaic power generation using hybrid deep learning model based on GRU and SVM[J].Electrical Engineering, 2024, 106(6).DOI:10.1007/s00202-024-02492-8.
9. Fei M , Yi P , Kedong Z ,et al.A Hybrid Online Forecasting Model for Ultrashort-Term Photovoltaic Power Generation[J]. Sustainability, 2018, 10(3):820.
10. Zhang Yujin, Yang Lingfan, Ge Shuangzhi, et al. Short-term photovoltaic power forecasting based on Kmeans algorithm and support vector machine[J]. Power System Protection and Control, 2018, 46 (21): 118-124.
11. Guo Yifei, Gao Houlei, Tian Jia. Photovoltaic output modeling based on cluster analysis and its application in reliability evaluation[J]. Automation of Electric Power Systems, 2016, 40(23): 93-100.
12. Ji Ling, Niu Dongxiao, Wang Peng. Research on photovoltaic power generation prediction based on similar day clustering and Bayesian neural network[J]. Chinese Journal of Management Science, 2015, 23(3): 118-122.
13. SHANGGUAN Xin, QIN Wenping, XIA Fuliang, et al. Pole-to-ground fault protection scheme for MMC multi-terminal flexible DC distribution network based on pearson correlation of transient voltage[J]. High Voltage Engin-eering, 2020, 46(5): 1740-1749.
14. Zheng S , Li D , Li Y ,et al.An adaptive weighted stacking ensemble framework for photovoltaic power generation forecasting with joint optimization of features and hyperparameters[J]. Engineering Applications of Artificial Intelligence, 2025, 144(000). DOI:10.1016/j.engappai.2025.110075.