Energy Proceedings
Vol 31, 2023
ISSN 2004-2965

Long-short term full-process forecasting of solar power and inelastic load

Bingtao Zhang[1], Hongchuan Qin[1], Xi Li[1,2], Zhe Cheng[1], Renjie Zhou[1], Jian Li[3] and Jianhua Jiang[1,2,*]

1 School of Artificial Intelligence and Automation, Key Laboratory of Education Ministry for Image Processing and Intelligent Control, Huazhong University of Science and Technology. Wuhan, Hubei, China

2 Shenzhen Research Institute of Huazhong University of Science and Technology. Shenzhen, Guangdong, China

3 School of Materials Science and Engineering, Huazhong University of Science and Technology. Wuhan, Hubei, China

* Corresponding author: Jianhua Jiang Tel: (+86) 15827271233 Fax: (+86) 027-87557273 Email address: jiangjh@hust.edu.cn

ABSTRACT

The forecasting of photovoltaic (PV) power generation and inelastic load is of great significance for the stable and efficient power supply of a microgrid power system. However, most of the PV prediction research in literature is based on known solar radiation which is difficult to obtained. In order to relieve the uncertainty in a microgrid, this work proposes full-process forecasting methods based on solar energy and load periodic characteristics analysis. For solar energy, a combination of Gaussian process regression (GPR) and physical model methods is utilized for the short-term accurate forecasting. The long-term trend forecasting is realized based on a cascade online TS fuzzy model. For inelastic load, an improved online long-short term memory (LSTM) rolling forecasting method is proposed. Simulation results show that the GPR & physical model methods can reach or even exceed the existing accuracy, while the cascade online TS fuzzy model method can achieve 5.5% higher accuracy than existing algorithms. Compared with the current offline LSTM method, the accuracy of the online method can be improved by up to 4.92%.

Keywords: PV power generation, inelastic load, GPR, online TS fuzzy model, LSTM

NONMENCLATURE

Abbreviations	
PV	Photovoltaic
GPR	Gaussian process regression
LSTM	Long-short term memory

1. INTRODUCTION

Microgrid technology is one of the most potential applications of distributed energy supply systems in the future. However, compared with the traditional thermal power generation, the uncertainty of the PV power generation unit and the inelastic load unit bring great challenges to the energy dispatching management of microgrid.

As a kind of clean energy, the prediction research of PV power generation has always been favored by many scholars. From the prediction method, it can be divided into physical, statistical and hybrid method [1]. The physical method is mainly to use mathematical equations to speed up the physical and motion state of meteorological conditions [2]. The statistical method is mainly to use the historical data training model to predict the future power value [3]. The main purpose of the hybrid method is to integrate the advantages of multiple algorithms and improve the effect of prediction [4]. However, most PV power prediction strategies use PV power historical data and solar irradiance, temperature and other weather factors to directly predict the future PV power through the algorithm. These strategies are not complete in engineering sense, because it only studies the relationship between PV power and solar irradiance, temperature and other historical data, and does not give a complete prediction process.

From the time scale, the prediction of inelastic load in microgrid can be divided into long-term, medium-term and short-term forecasting. However, short-term load forecasting (STLF) is more important, so most studies focus on the accuracy of STLF [5-7].

In a word, the prediction of PV power and inelastic load have important research significance in the field of microgrid. For PV power, because it is mainly affected by solar irradiance, the overall strategy in this paper is divided into two steps: the first step is to predict the total solar irradiance, and the second step is to predict the PV power. For the prediction of solar irradiance, GPR and TS fuzzy model are used in this paper. While for the prediction of PV power, the physical modeling method and the prediction method based on TS fuzzy model are used in this paper, and the two methods are compared with the existing literature. For inelastic load forecasting, this paper puts forward an improved online LSTM rolling forecasting method to improve accuracy.

The main structure of this paper is as follows: The second section is methodology, introducing some basic algorithms. The third section is the specific steps of forecasting, including PV power forecasting and inelastic load forecasting. The fourth section is the result and discussion. The fifth section is conclusion.

2. METHODOLOGY

2.1 Select the appropriate input

Before making a prediction, we first need to determine the appropriate input vector. For PV power, this paper mainly cites seven factors such as humidity, temperature, etc. Then the Pearson correlation coefficient is used to calculate the correlation between each influence factor and temperature, and the factors with strong correlation are selected as the model prediction input. Its formula is as formula (1).

$$p_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma X \sigma Y} \tag{1}$$

2.2 Data clustering

There are many kinds of clustering algorithms, because the data sets in this paper are numerical data, and the samples are not large, so the typical partition clustering can meet the requirements. K-Means algorithm is a simple and efficient partition clustering method, which has low complexity and fast operation speed, and meets the needs of this paper. Because PV power generation is affected by meteorological conditions, at present, based on the idea of "similarity", most researchers use clustering algorithm to divide the weather into sunny, cloudy and rainy days, and establish forecasting models respectively. However, according to the usual habits of sunny, cloudy and rainy days, the probability does not accord with the historical data set, that is, according to the actual local weather conditions, the best clustering center of data clustering may not be three. Therefore, based on this background, this paper chooses the K-Means clustering algorithm with elbow evaluation method for data preprocessing. For specific algorithms, please refer to [8].

2.3 GPR & physical model

For the prediction of total solar irradiance, because the internal of the physical model is more complex and the generalization ability is poor, the data-driven model is selected in the prediction. As one of the data-driven algorithms, GPR can solve the problem that the model is complex and difficult to establish an accurate model. GPR is not only a nonparametric Bayesian method, which can obtain the ability-confidence interval with uncertain prediction results, but also has nonparametric characteristics, which is to allow the parameters of the model to be calibrated according to the requirements of the data. Therefore, GPR is used to predict the total solar irradiance in this paper. For specific GPR, please refer to [9].

At present, the physical model of PV power generation is mainly divided into three-parameter, four-parameter, five-parameter, six-parameter and seven-parameter models [10]. Because the three-parameter and four-parameter model is relatively simple, the accuracy is lower than that of the high-parameter model. Although the accuracy of the six-parameter and seven-parameter model is high, the modeling and correction process is complex and the convergence speed is slow. Therefore, this paper eclectically chooses the five-parameter model for physical modeling, and then integrates it to establish the PV array model.

2.4 Online TS fuzzy model

The main idea of TS fuzzy model is to use linear model weighted combination to approximate nonlinear model. Its advantage is that few rules can be used to achieve a higher non-linearized model. In this paper, the reason why TS fuzzy model is introduced to establish the total solar irradiance model and PV power model is that most scholars use the idea of "similarity" when studying PV power generation prediction, cluster the historical data and establish models for the resulting clusters respectively. Therefore, from the theoretical analysis, the use of TS fuzzy model method is beneficial to improve the accuracy of PV prediction. For specific algorithms, please refer to [11].

2.5 LSTM

One of the important features of LSTM network is that its holding unit can keep memory information for a long time, integrates RNN framework, eliminates the problem of gradient disappearance, and has the ability to screen memory information [7]. For load forecasting, because of the long-term sequence and periodicity of its data, it is very suitable to use LSTM algorithm to learn and predict the future load demand.

3. PV AND LOAD FORECASTING

3.1 Data feature extraction

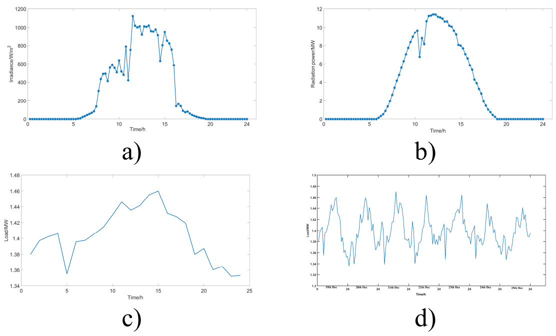


Fig. 1. PV and load data curve: a) solar irradiance for a day, b) PV power for a day, c) inelastic load for a day, d) inelastic load for a week.

Fig. 1 a) data from the Solar Radiation Research Laboratory of National Renewable Energy Laboratory (NREL) in the United States. What is mainly provided is the 2012 total solar irradiance historical data set in Denver, Colorado, which shows the 24-hour (96 data points) total solar irradiance curve on August 1st. Fig. 1 b) data comes from a PV project in Huzhou, Zhejiang Province, China. It is mainly from the PV power map for the whole day of August 1, 2016. The data is collected every 15 minutes, with a total of 96 data points. The data in Fig. 1 c) and d), from a district in Ningxia, China, contains data for the whole year of 2016, which is collected every other hour. Fig. 1 c) shows load data from Monday, December 19, 2016, and Fig. 1 d) shows data from the week from December 19, 2016 to December 25, 2016.

As can be seen from the above four Figs, these curves all have a small range of slope events and have certain volatility. And it has obvious regularity and periodicity, which will be of great help to the follow-up forecasting work.

3.2 PV forecasting

As mentioned earlier, two strategies are used to predict PV, each of which is divided into two steps. The first step is to predict the total solar irradiance, and the second step is to predict the PV power.

3.2.1 Short-term forecasting: GPR & physical model

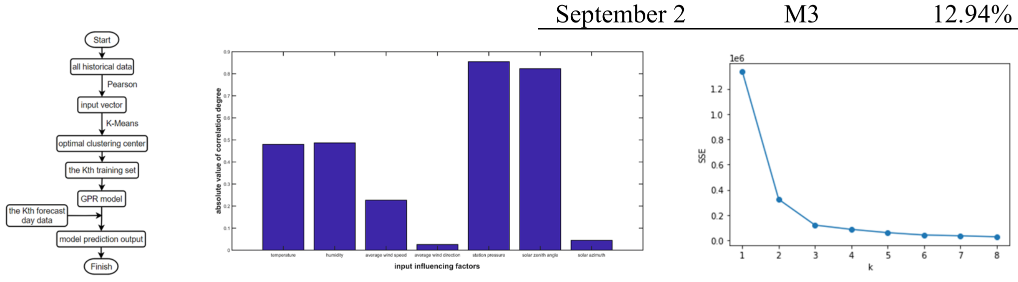
The first step is to predict the total solar irradiance based on the GPR data-driven method. The specific flow chart is shown in Fig. 2 a).

The experimental data here are the same as Fig. 1 a). The data include seven influencing factors: temperature, humidity, average wind speed, average wind direction, pressure in the station, solar zenith angle and solar azimuth. The data sample is selected from 7:00 to 18:00 every day, every 15min measurement, a total of 44×92=4048 data in the training set, and September 1 and September 2, 2012 are selected as forecast samples.

By using Pearson correlation coefficient method, the correlation between the seven influencing factors and solar irradiance is shown in Fig. 2 b). So, this paper selects the station pressure, solar zenith angle, temperature and relative humidity as the input vector of the model. Next, using the K-Means clustering algorithm with elbow evaluation method, we get the elbow fig is shown in Fig. 2 c). So, the best clustering center is K=2, and the dataset is divided into two categories. Then, for two types of data sets, GPR algorithm is used to train the model. MAPE and RMSE are selected as the evaluation index here, and the results are shown in Table 1.

Table 1. Short-term irradiance forecast results

Date	Kernel Function	MAPE	RMSE
September 1	M3	15.50%	86.4W/m^2
September 2	M3	12.94%	100.66W/m^2



a) b) c)

Fig. 2. Short-term forecasting fig: a) the flow chart of the solar irradiance forecasting, b) the correlation between the influencing factors and the total irradiance, c) irradiance clustering elbow diagram.

The second step is PV power prediction based on the physical model.

PV power data are the same as Fig. 1 b). Select August as the calibration data, September 2, 4 and 17 as the prediction verification samples (In order to compare with the existing algorithms, and these three days are sunny, cloudy and rainy days respectively).

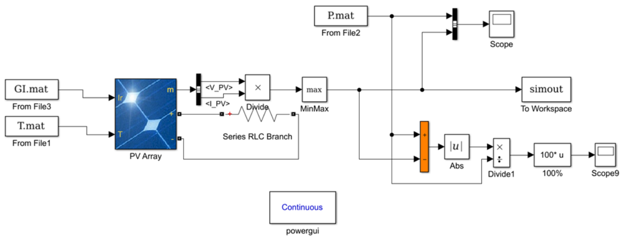


Fig. 3. Physical model of PV

The physical model of PV power generation is established as shown in Fig. 3. And the input vector is temperature and total irradiance (that is, the total solar irradiance predicted by GPR algorithm). Then, calibrate the physical model is necessary. Last, use the physical model to predict. MAPE is selected as the evaluation index here, and the results are shown in Table 2.

Table 2. Short-term PV forecast results

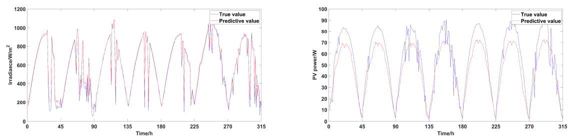
Date	MAPE
September 2	3.01%
September 4	3.22%
September 17	4.28%

3.2.2 Long-term forecasting: online TS fuzzy model

Corresponding to 3.2.1, it is also divided into two steps. The first step is to use the online TS fuzzy model to predict the total solar irradiance. The same data set as 3.2.1 is selected here, with data from June to August as the training set and August 25 – August 31 in the last week of August as the prediction set. The prediction steps are also similar to 3.2.1. In the online TS fuzzy model, the number of rules is equal to the number of clustering centers in 3.2.1, so the best number of rules is 2, that is, the data set is divided into two categories. Firstly, it is judged that the forecast day data belong to which category, and then the prediction set is input into the trained model to get the prediction output, as shown in Fig. 4 a). At this point, the overall error of the week is MAPE=8.26%, RMSE=27.602 W/m^2.

The second step is to use online TS fuzzy model to predict PV power. However, when establishing and verifying the PV power of TS fuzzy model, the data used is the same data set as section 3.2.1, but this section expands the historical training set. The training set in this section uses PV data from June to August. The data are measured every 15 minutes from 7:00 to 18:00, with a total of 44×92=4048 data samples. The prediction and verification data are selected for August 25 – August 31.

Total irradiance and ambient temperature are selected as input vector. And the best clustering center K=3, so the optimal number of rules of the online TS fuzzy model is also 3. That is, the existing historical data sets are optimally divided into three categories. Firstly, it is judged that the forecast day data belong to which category, and then the prediction set is input into the trained model to get the prediction output, as shown in Fig. 4 b). At this point, the overall error of the week is MAPE=20.23%, RMSE=113.823W.



a) b)

Fig. 4. Long-term forecasting fig: a) irradiance predictive result, b) PV power predictive result.

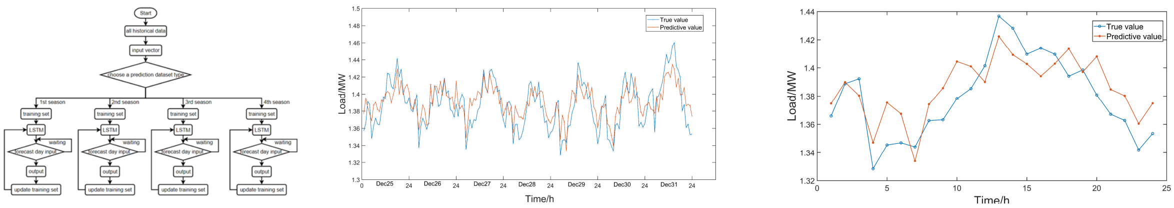
3.3 *Inelastic load forecasting*

Compared with the traditional offline LSTM method, here we propose an improved online LSTM rolling forecasting method.

Offline LSTM cannot be updated in real time because of its fixed training set. With the increase of time, the prediction performance of offline LSTM prediction method will be greatly reduced. Therefore, the short-term load rolling forecasting algorithm of online LSTM is introduced, such as Fig. 5 a). Compared with offline LSTM, online LSTM mainly puts forward two improvements: one is online training, that is, each prediction, update the historical training set to improve the prediction accuracy. The other is rolling prediction, that is, each model only selects one group of data in the prediction input set, and when the prediction is completed, update the historical data set, actively input the next set of prediction input data set to achieve rolling prediction until the prediction is completed.

In addition, when using LSTM to predict load in this section, the year is divided into four quarters. The first quarter is from January to March, the second quarter is from April to June, the third quarter is from July to September, and the fourth quarter is from October to December. Four quarterly models were established to make predictions on the prediction training set to improve the prediction accuracy.

The load data are the same as Fig.1 c) and d). In the prediction verification, take the fourth quarter as an



a) b) c)

Fig. 5. Online LSTM fig: a) the flow chart of online LSTM prediction, b) weekly load forecasting curve, c) December 29th load forecasting curve.

example to verify the effectiveness of the model. The specific data selected the last week of December in the fourth quarter of 2016 (December 25-December 31) as the forecast set, with a total of 7×24=168 data. October 1-December 24, 2016 as a history training set, with a total of 24×85=2040 data.

The process of online LSTM direct prediction is described below in conjunction with Fig. 5 a). In this paper, when selecting the input data set of LSTM model, refer to [12]. Through the verification of the algorithm, a group of best load forecasting inputs (load history data, temperature, hour, daily type) are recommended. Then, we should train offline LSTM model, the final debugging of the specific parameters are shown in Table 3.

Table 3. Selection of training parameters for online LSTM

Parameter name	Parameter value
Training algorithm	MBGD
Optimization error algorithm	Adam
Learning rate	0.001
Number of iterations	1500
Number of hidden neurons	30

Last is prediction and verification. Firstly, it is judged that the forecast day data belong to which quarter, and then the prediction set is input into the trained model to get the prediction output, as shown in Fig. 5 b). At this point, the overall error of the week is MAPE=1.05%, RMSE=124.7W. In order to compare with the offline LSTM forecasting model, December 29th is selected as the comparison item, the load forecasting curve is shown in Fig. 5 c), and the forecasting performance of December 29th is MAPE=1.18%, RMSE=128.59W.

4. RESULT AND DISCUSSION

4.1 PV forecasting

Section 3.2 proposes two prediction strategies. One is short-term forecasting strategy, the other is long-term forecasting strategy.

For GPR & Physical model, if analyzed from the point of view of accuracy, the WPD-LSTM method proposed in reference [13] is compared with LSTM, GRU, RNN and MLP, and the superiority of WPD-LSTM method is proved. Here, we compare the PV power prediction results of the two strategies with the WPD-LSTM method from the point of view of three kinds of weather, as shown in Table 4. It can be seen from the table that the proposed strategy not only ensures the integrity of engineering meaning, but also ensures the prediction accuracy.

Table 4. Short-term PV forecasting of MAPE comparison (a day)

Day type	Short-term*	WPD-LSTM[13]**
sunny	3.01%	2.0367%
cloudy	3.22%	3.7923%
rainy	4.28%	4.3427%

* PV forecasting of this work based on irradiance prediciton
** PV forecasting based on given irradiance

For online TS fuzzy model, reference [14] propose two hybrid models (CNN-LSTM and ConvLSTM) to effectively predict the power production of a self-consumption PV plant. Here we compare our results with the reference [14], as shown in Table 5. It can be seen from the table that the proposed long-term forecasting strategy can improve the accuracy.

Table 5. Long-term PV forecasting of MAPE comparison (a week)

Long-term*	CNN-LSTM[14]**	ConvLSTM[14]**
20.23%	25.73%	30.74%

* PV forecasting of this work based on irradiance prediciton
** PV forecasting based on given irradiance

From the overall proposed prediction strategy, the method of GPR & physical model is more accurate, and the difference between them lies in PV power prediction. However, from the above analysis, for long-term continuous prediction, the physical model is not very suitable, because it needs real-time calibration. In contrast, if the short-term use, GPR & physical model prediction strategy is better, if the need for long-term use of secondary prediction model, then TS fuzzy model online prediction is better.

4.2 Inelastic load forecasting

The LSTM method is pioneered in reference [7], and the superiority of the LSTM method compared to

ARMA, SARIMA and ARMAX is proved. Here, we compare the results of the improved online LSTM method with the offline LSTM method used in [7], as shown in Table 6.

Table 6. Load forecasting of MAPE comparison

Strategy	1 Day	7 Days
Online LSTM	1.18%	1.05%
Offline LSTM [7]	1.522%	5.97%

It can be seen from the Table 6 that the accuracy of the proposed online LSTM method has been improved no matter in the prediction results of one day or one week. There are two reasons. The first is the improvement of the algorithm. Because of its fixed training set, offline LSTM can not be updated in real time. With the increase of time, the prediction performance of offline LSTM prediction method will be greatly reduced, while online LSTM can continuously update the training set and improve the accuracy because of the advantage of rolling prediction. Second, in the data processing, due to the different characteristics of the data in the four seasons of the year, this paper chooses to forecast the data in four quarters and establishes four models to improve the accuracy.

5. CONCLUSION

In this paper, forecasting strategies are proposed for PV power and load forecasting according to periodic characteristics. For PV power generation, the first is the GPR & physical model, and the second is the online TS fuzzy model. The results show that the short-term strategy can reach or even exceed the existing accuracy, while the long-term strategy can achieve 5.5% higher accuracy than existing algorithms. For inelastic load, compared with the current offline LSTM method, the accuracy of the online method can be improved by up to 4.92%. This work provides the basis for the future work of microgrid management.

ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China (62173154, 61773174, 61873323, U1910209), Science and Technology Planning Project of Shenzhen Municipality (JCYJ20210324141610026), and the Hubei Province Natural and Science Foundation (2017CFB4165). Thanks for the basic work from the workmates in my lab and my lab-Key Laboratory of Education Ministry for Image Processing and Intelligent Control, Huazhong University of Science and Technology.

REFERENCE

[1] Yadav H K, Pal Y, Tripathi M M. Photovoltaic power forecasting methods in smart power grid[C]//2015 annual IEEE India conference (INDICON). IEEE, 2015: 1-6.

[2] De Blas M A, Torres J L, Prieto E, et al. Selecting a suitable model for characterizing photovoltaic devices[J]. Renewable energy, 2002, 25(3): 371-380.

[3] Wang F, Zhen Z, Wang B, et al. Comparative study on KNN and SVM based weather classification models for day ahead short term solar PV power forecasting[J]. Applied Sciences, 2017, 8(1): 28.

[4] Ogliari E, Dolara A, Manzolini G, et al. Physical and hybrid methods comparison for the day ahead PV output power forecast[J]. Renewable Energy, 2017, 113: 11-21.

[5] Lu H, Azimi M, Iseley T. Short-term load forecasting of urban gas using a hybrid model based on improved fruit fly optimization algorithm and support vector machine[J]. Energy Reports, 2019, 5: 666-677.

[6] Barman M, Choudhury N B D, Sutradhar S. A regional hybrid GOA-SVM model based on similar day approach for short-term load forecasting in Assam, India[J]. Energy, 2018, 145: 710-720.

[7] Muzaffar S, Afshari A. Short-term load forecasts using LSTM networks[J]. Energy Procedia, 2019, 158: 2922-2927.

[8] Wang J, Ma X, Duan G. Improved K -means clustering k -value selection algorithm. Computer Engineering and Applications, 2019, 55(8): 27-33.

[9] He Z K, Liu G B, Zhao X J, et al. Overview of Gaussian process regression[J]. Control and Decision, 2013, 20(08): 1121-1129+1137.

[10] Wang C J. Versatile Model for Photovoltaic Cell Based on MATLAB[J]. Electric Power Science and Engineering, 2009, 25(04): 11-14.

[11] Takagi T, Sugeno M. Fuzzy identification of systems and its applications to modeling and control[J]. IEEE transactions on systems, man, and cybernetics, 1985 (1): 116-132.

[12] Chitalia G, Pipattanasomporn M, Garg V, et al. Robust short-term electrical load forecasting framework for commercial buildings using deep recurrent neural networks[J]. Applied Energy, 2020, 278: 115410.

[13] Li P, Zhou K, Lu X, et al. A hybrid deep learning model for short-term PV power forecasting[J]. Applied Energy, 2020, 259: 114216.

[14] Agga A, Abbou A, Labbadi M, et al. Short-term self consumption PV plant power production forecasts based on hybrid CNN-LSTM, ConvLSTM models[J]. Renewable Energy, 2021, 177: 101-112.