LOWESS Smoothing and Random Forest Based GRU Model：A Short-term Photovoltaic Power Generation Forecasting Method

Yeming Dai[1*], Yanxin Wang[1], Mingming Leng[2], Xinyu Yang[1], Qiong Zhou[1]

1. *School of Business, Qingdao University, Qingdao 200071, China*

2. *Faculty of Business, Lingnan University, Hong Kong.*

Abstract: Accurate prediction of photovoltaic power generation is vital to guarantee smooth operation of photovoltaic power stations and ensure the electricity consumption of end users. As a good forecasting tool, Gated Recurrent Unit method has been widely used in different forecasting areas. However, the existing studies ignore the impact of data fluctuations on prediction accuracy since photovoltaic power generation is intermittent and uncertain, then the prediction results of Gated Recurrent Unit are facing challenges. To fill the gaps and enhance prediction accuracy, this paper develops an improved Gated Recurrent Unit photovoltaic generation prediction method. Several different data smoothing techniques are introduced and compared to reduce fluctuations, Random Forest method is used for feature selection, and RepeatVector layer extended by attribute dimensions and TimeDistributed layer with full connectivity are utilized to optimize the Gated Recurrent Unit model. A real-world case from the photovoltaic power plant in Xuhui District, Shanghai, China, is adopted to evaluate the performance of proposed method. After comparing with the prediction results of Recurrent Neural Networks and Long Short-Term Memory, and the actual data as well, it is found that the proposed prediction method can effectively improve the prediction accuracy of photovoltaic power generation.

Keywords: Photovoltaic Power Generation; Prediction; Locally Weighted Scatterplot Smoothing; Random Forest; Gated Recurrent Unit

1 Introduction

Energy saving and carbon emission reduction is a topic of common concern all over the world. In China, the government has drawn the historic lessons from the failure of energy saving and carbon emission and sought a climate governance road with Chinese characteristics. Moreover, the clean air action plan, carbon peaking and carbon neutral strategies to achieve energy conservation and environment protection goals were claimed in 2021 [1]. Other countries are also making contributions to environment protection now. For instance, UK announced a policy for its Net Zero Strategy to support the transition of British businesses and consumers to clean energy and green technologies [2].

In particular, as a complementary and multi-system coordinated energy supply and consumption mode, renewable energy (wind energy, water energy, solar energy, geothermal energy, etc.) in the electricity market has become an important means to improve the efficiency of energy utilization and make great contributions to energy saving and carbon emission reduction in recent

[*] *Corresponding author: E-mail: yemingdai@163.com.*

Nomenclature			
		kNN	k-Nearest Neighbor algorithm
		MAE	Mean Absolute Error
$(x, \hat{y})$	Center value of regression line	MAPE	Mean Absolute Percentage Error
i	Smoothing window	ML	Machine Learning
h_i	Smoothing coefficient	MLP	Multilayer Perceptron
$w(i, j)$	Weight value	MSE	Mean Square Error
$f(i, j)$	Original data	LOESS	Locally Estimated Scatterplot Smoothing
$g(x, y)$	Smoothed data	LOWESS	Locally Weighted Scatterplot Smoothing
x^t	Current input	LR	Linear Regression
h^{t-1}, h^t	Hidden state	LSTM	Long-Short Term Memory
y^t	Output	NN	Neural Network
$y_i, y_i^{'}$	real value, predicted value	NRMSE	Normalized Root Mean Squared Error
$y_{\max} - y_{\min}$	Full distance of real value	OOB	Out Of Bag
AAv	Adjacent Average method	PCA	Principal Component Analysis
AI	Artificial Intelligence	PF	Percentile Filtering
ARD	Automatic Relevance Determination method	R^2	R-Squared
CNN	Convolutional Neural Network	RMSE	Root Mean Square Error
DEM	Digital Elevation Model	RMSPE	Root-Mean-Square Percentage Error
EANN	Evolutionary Artificial Netural Network	RNN	Recurrent Neural Networks
ES	Exponential Smoothing	SG	Savitzky-Golay Smoothing
GRA	Grey Relational Analysis	SVM	Support Vector Machine
GRU	Gated Recurrent Unit	Var	Variance value

years. Compared with other renewable energy sources, solar energy has the advantages of high flexibility, superior adaptability, and low development cost, which has broaden the social development potentials and prospects [3]. Photovoltaic power generation then has been a priority since it can convert solar energy into electricity. With the popularity of photovoltaic power generation, more and more countries and regions have been implementing their initiatives of integrating photovoltaic power generation into power grids, which has resulted in an increase in daily power supply and a reduction in carbon emissions. Nevertheless, the development of renewable energy is a combination of opportunities and challenges. One of the major challenges is that renewable energy is affected by natural factors and can be unstable when connected to the power grid [4]. The photovoltaic power generation is greatly affected by weather factors. This results in its intermittent defects, which, in turn, is not conducive to the stable operation of power grid. Therefore, accurate prediction results of photovoltaic power generation can make appropriate operations and scheduling efforts and alleviate the instability issues.

Existing studies have shown that the length of prediction period has an important impact on the prediction accuracy and application scenarios. According to the length of prediction period, forecasting methods can be divided into three types: very short-term [5], short-term [6], and medium and long-term [7]. For very short-term forecasting, it is accurate in seconds to minutes and is suitable for real-time dispatching different sizes grids, so as to reserve spare capacity for power grid in a timely manner [6]. In terms of short-term forecasting with the range from hour to day, it concerns economic dispatch and decision making of power grid to balance power market transactions whose meaning is adjusting the unit combination scheme and optimizing the generation plan [5]. In addition, medium and long-term prediction focuses on day to week/month/year based on the very short-term and short-term forecasting, which provides a long range plan for power grid and shows the ability for the equipment maintenance and the siting of new energy base stations. Many facts have shown that medium and long-term forecasting is filled with the big picture concept, which is also the development direction of forecasting in many economic and industrial areas [7]. In general, the three different prediction methods are adopted according to the characteristics of power grid in terms of time scale, application scope, and purpose. Since short-term prediction concerns the economic dispatch and decision making of power grid to balance power market transactions, our paper focuses more on short-term prediction of photovoltaic power generation in order to make better generation planning and more timely power plant offers to the dispatch center, as well as to improve the security and economy of power grid [8].

The short-term prediction process of photovoltaic power generation mainly includes data processing and model prediction [9]. Data processing is an important prerequisite for prediction, which is shown as data cleaning, data integration, data transformation, and data protocol [10-12]. However, existing literatures ignore the importance of data fluctuations, there are still fewer analysis of data fluctuation in the prediction of photovoltaic power generation [13]. Therefore, it is necessary to pay more attention to the research on these smoothing methods. To the best of our knowledge, Locally Weighted Scatterplot Smoothing (LOWESS), Locally Estimated Scatterplot Smoothing, LOESS) [14], Savitzky-Golay Smoothing (SG) [15], Adjacent Average method (AAv) [16], and Percentile Filtering (PF) [17] have been representative smoothing methods with convenient operation and rapid arithmetic in recent years and have not been applied to the prediction of photovoltaic power generation [10,18]. Accordingly, our paper applies LOWESS, LOESS, SG, AAv, and PF smoothing methods to the data processing in the prediction of photovoltaic power generation to improve the data quality. We also compare the above five smoothing methods based on different evaluation metrics, and then filter out the best method with the lowest prediction error, so as to achieve the goal of improving prediction accuracy.

In addition, we also use feature selection in the data processing. Feature selection refers to the process of selecting some effective features from existing features to reduce the data dimension, mainly including filter, embedded, and wrapper [19]. For example, in the filter method, Ref. [20] used the Automatic Relevance Determination method (ARD) to point out the most relevant input for the accurate monthly average daily solar radiation prediction, Ref. [21] used ridge regression algorithm in the embedded method. These feature selection methods improve the prediction accuracy of the model, but the performance and calculation speed are not as good as the wrapper method. Wrapper method mainly includes Random Forest, SVM(Support Vector Machine), and k-Nearest Neighbor algorithm (kNN) [19]. Compared with SVM and kNN algorithm, Random Forest can process high-dimensional data, deal with many problems such as classification, feature selection, and regression [22]. Existing research and experimental results on Random Forest have exposed that Random Forest feature selection can effectively improve the prediction accuracy [22]. Then, this paper also applies Random Forest for data processing to analyze the factors that affect power generation and obtain higher input data quality.

The next step after data processing is the prediction with a suitable model. At present, the prediction models of photovoltaic power generation are mainly divided into four categories: persistence forecast of "today equals tomorrow" [23], physical model based on terrain research [24], statistical techniques related to time series [25], and Artificial Intelligence (AI) prediction represented by Machine Learning (ML) [26]. The first three categories possess some prediction flaws because of an increase in time span and abnormal sudden changes in weather [27]. ML mainly includes Linear Regression (LR), SVM, and Neural Network (NN). Especially, NN can efficiently process a large amount of data, improve the prediction accuracy and solve the defects in persistence forecast, physical model, and statistical techniques [26]. Today, NN has become the primary choice of prediction methods in many fields.

As one of NN, Gated Recurrent Unit (GRU) can solve gradient disappearance and explosion of Recurrent Neural Networks (RNN), simplify parameters of Long-Short Term Memory (LSTM) [7,29], which shows excellent performance in prediction and obtains some further improvement [30]. However, the impact of data smoothing on prediction accuracy is not considered, the input data in multiple dimensions is not analyzed, and the information is shared diversely. There is still much room for the improvement of GRU model. Therefore, we develop an improved GRU model by introducing RepeatVector layer and TimeDistributed layer to optimize the GRU model, which is different from other optimized GRU models without diverse and multidimensional improvements in term of model hierarchy. To provide an improved reflection made to the GRU model in this paper and indicate the differences with other NN-based literatures, we have made a comprehensive comparison in Table 1. Moreover, our paper introduces data smoothing techniques while none of the remaining references introduces them. For feature selection, we use Random Forest for feature analysis which is similar to [11,30,32] while the other literatures do not perform feature selection. In summary, compared to other NN-based references, the innovation of this paper mainly includes the application of data smoothing techniques and the optimization of GRU model.

Based on the above analysis, we first introduce LOWESS, LOESS, SG, PF, and AAv data smoothing methods and compare them to filter the best method with the least error. Secondly, considering the variety of natural factors that affect photovoltaic power generation, we use Random Forest for feature selection. Finally, we optimize the GRU model for prediction by using RepeatVector layer and TimeDistributed layer. The main contributions of this paper are as follows:

(1) We consider different data smoothing technologies to reduce the data fluctuation of daily photovoltaic power generation. We also compare these data smoothing techniques to find the best smoothing method that has the least prediction error.

(2) We use Random Forest to extract the characteristics of natural factors affecting daily photovoltaic power generation.

(3) We add RepeatVector layer and TimeDistributed layer into the GRU model to improve its prediction accuracy.

(4) We utilize the dataset from Shanghai, China and three prediction models to verify the accuracy and feasibility of our proposed method.

The remainder of this paper is arranged as follows: Section 2 introduces the basic concepts of the five smoothing methods mentioned in this paper, the generation process of Random Forest, and original GRU model. Section 3 describes the structure of optimized GRU with the addition of RepeatVector layer and TimeDistributed layer. Section 4 gives the specific forecasting steps. In Section 5, we provide details for our case study, which include a description of experimental data, the selection of evaluation metrics, the setting of model parameters, and our comprehensive analysis of experimental results. Section 6 presents a summary of our works in this paper.

Table 1. Comparison between this paper and other NN-based references.

Study	Year	Prediction period	Data source	Data processing	Prediction model	Optimization method	Prediction error
This paper	2021	Short-term	China	Data smoothing and Random Forest feature selection	GRU	Use RepeatVector layer and TimeDistributed layer to optimize GRU	RMSE: 2.352 MAE: 1.851 MAPE: 19.715 MSE: 5.530 R^2: 0.955 NRMSE: 0.102
[5]	2021	Very short-term	Basque	N-nearest-station model	MLP (Multilayer Perceptron)	Optimize the length of the input window	RMSE: 0.2515 R^2: 0.9985
[9]	2020	Short-term	Spain	Pearson correlation	EANN (Evolutionary Artificial Netural Network)	Evolutionary algorithm	MBE: 0.30 MAE: 33.46 RMSE: 0.9709
[26]	2020	Medium and long-term	Korea	DEM (Digital Elevation Model)	LSTM-RNN	Use LSTM layer to optimize stacked RNN	R^2: 0.724 RMSE: 14.003 NRMSE: 7.416 MAPE: 10.805
[28]	2019	Medium and long-term	China American	Copula function	LSTM	Joint prediction (wind and photovoltaic power generation)	MAPE: 6.65 RMSPE: 8.43
[30]	2021	Short-term	Australia	Remove outliers and feature normalization	Conv-GRU	Use convolutional layers to optimize GRU	R^2: 0.8938 RMSE: 2.630
[31]	2021	Very short-term	American	Data augmentation techniques	CNN (Convolutional Neural Network)	Adam algorithm	RMSE: 3.259
[32]	2021	Short-term	Vietnam	Pearson correlation and remove outliers	LSTM	Replace the historical weather data entered into the model with forecast weather data	MSE: 56.348 RMSE: 7.507 MAE: 4.743 MAPE: 9.881

2 Theoretical basis

This section presents the base methods used in data processing and prediction. The data processing mainly includes LOWESS, LOESS, SG, AAv and PF in data smoothing and Random Forest in feature selection. We specify the original GRU model as a base prediction model.

2.1 Data smoothing

2.1.1 LOWESS smoothing

① Definition: take point x as the center, intercept a section of proportional data forward and backward respectively, make weighted linear regression with weight function W for this section of data, $(x, \hat{y})$ is the center value of the regression line, $\hat{y}$ represents the corresponding value after fitting the curve, all n data can make n weighted regression lines. The connection of the central value of each regression line is the LOWESS smooth curve of this data [14].

② Weight function W

The commonly used weight function is the cubic function $W(x)$.

$$W_{(x)} = \begin{cases} (1-|x|)^3, & if \ |x|<1; \\ 0, & if \ |x| \geq 1. \end{cases} \quad (1)$$

2.1.2 LOESS smoothing

LOESS smoothing divides the samples into multiple cells, performs polynomial fitting on the interval samples, repeats the fitting process continuously, and obtains multiple weighted regression curves, finally connects the center of the curve to obtain the smooth curve [14].

2.1.3 Savitzky-Golay smoothing

Savitzky-Golay smoothing is based on the least square principle and performs k -order polynomial fitting for data points in a certain length window [15]. In formula (2), i represents the i th smoothing window, h_i represents the smoothing coefficient, h_i/H is solved by the least square method.

$$x_{k,smooth} = \bar{x}_k = \frac{1}{H}\sum_{i=-w}^{+w} x_{k+i} \cdot h_i \quad (2)$$

2.1.4 AAv smoothing

AAv smoothing is a smoothing method for calculating the arithmetic mean of several adjacent data [16], use neighborhood average (formula (3)) or weighted average (formula (4)) for smoothing. $w(i, j)$ represents weight value, $f(i,j)$ represents the original data, M denotes the number of adjacent data, $g(x,y)$ is the smoothed data.

$$g(x,y) = \frac{1}{M}\sum_{i,j\in s} f(i,j) \quad (3)$$

$$g(x,y) = \frac{1}{M}\sum_{i,j\in s} w(i,j) f(i,j) \quad (4)$$

2.1.5 PF smoothing

PF smoothing is a non-linear smoothing method that calculates a specified quantile value for local data and replaces the original data with this quantile value, which is suitable for signal smoothing with pulse characteristics [17].

2.2 Random Forest

The generation process of Random Forest is to put back samples from the original training samples to obtain numerous subsets. These subsets train different base classifiers, and the optimal classification results are determined by the voting of the base classifier [33]. The evaluation of Random Forest performance mainly uses Out Of Bag error (OOB error). When the total number of samples=N, the importance of features is calculated by formula (5), $errOOB_1$ represents the Out Of Bag data error of each decision tree, $errOOB_2$ represents the Out Of Bag data error after adding noise interference.

$$\sum \frac{errOOB_2\text{-}errOOB_1}{N} \quad (5)$$

2.3 GRU prediction model

The realization process of GRU is as follows: combine the current input x^t and the hidden state h^{t-1} passed down from the previous node to obtain the output y^t of the current hidden node and the hidden state h^t passed to the next node [7]. The basic GRU model has only one layer, and there is room for optimization. Fig.1 shows the network structure of GRU model, and Fig.2 shows the module internal structure of GRU model.

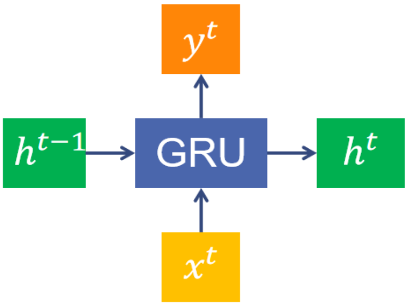

Fig.1. GRU model network structure.

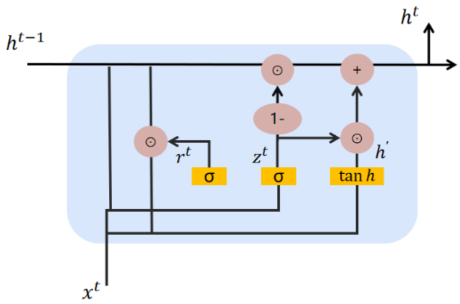

Fig.2. GRU model internal structure.

The procedure of GRU is as follows [29]:

Step 1: Calculate the gate: obtain the weight and parameter status of the update gate and reset gate through h^{t-1} and x^t, σ represents the ***sigmoid*** activation function.

① Update gate:

$$z_t = \sigma(W_z \cdot [h^{t-1}, x^t]) \qquad (6)$$

The parameter representation of the update gate:

$$[(x_{dim} + h_{dim}) * h_{dim} + h_{dim}] \qquad (7)$$

② Reset gate:

$$r_t = \sigma(W_r \cdot [h^{t-1}, x^t]) \qquad (8)$$

The parameter representation of the reset gate:

$$2 * [(x_{dim} + h_{dim}) * h_{dim} + h_{dim}] \qquad (9)$$

③ Total parameter formula expression: it is obtained by adding formula (7) and formula (9):

$$3 * [(x_{dim} + h_{dim}) * h_{dim} + h_{dim}] \qquad (10)$$

Step 2: Capture information: directly extract the local information $h^{t-1'}$ from the long-term information h^t through r_t.

$$h^{t-1'} = h^{t-1} \otimes r_t \qquad (11)$$

Step 3: Obtain the current information $h^{'}$: splice $h^{t-1'}$ and x^t, and use $\tan h$ to form compression.

$$h^{'} = \tan h(W \cdot [h^{t-1}, x^t]) \qquad (12)$$

Step 4: Generate a new output h^t: fuse h^{t-1} and $h^{'}$ by taking a part of each weight. Among them, $(1 - z_t) \otimes h^{t-1}$ is the selective "forgetting" of h^{t-1} unimportant information, $z_t \otimes h^{'}$ is the selective memory of $h^{'}$.

$$h^t = (1 - z_t) \otimes h^{t-1} + z_t \otimes h^{'} \qquad (13)$$

3 Multi-layer optimized GRU model with RepeatVector layer and TimeDistributed layer

Original GRU model has a GRU layer only, which possesses randomness and uncertainty. We add RepeatVector layer and TimeDistributed layer to make the GRU layer more diversified.

First, we add the RepeatVector layer to the GRU hierarchy to ensure the same vector in each time step, which specifically refers to increase the dimension of input data and add attribute dimension, then the model can be analyzed in all aspects from various dimensions [34]. The parameter is represented by *n*. For example, when *n*=3, it means that the dimension of input data increases to 2 dimensions;

Secondly, we add the TimeDistributed layer using time series for tensor operations to obtain a better weight information sharing, and the same fully connected layer can be applied to each time step, thus achieving full connectivity in the time dimension [34]. The structure layer and setting basis are shown in Table 2.

Table 2. GRU model structure layer and setting basis.

Model layer (in order)	Input and output	Number of parameters (201216+394752+1028=596996)	Calculation method (obtained according to formula (10))	Setting basis
GRU_1	Input:(None, 256) Output: (None, 256)	201216	3*[256*(256+5)+256]	Layer 1 GRU prediction model
RepeatVector	Input:(None, 256) Output:(None, 2, 256)	0	0	Repeated input of potential vectors can increase attribute dimensions, which is beneficial to multi-dimensional analysis of the model.
GRU_2	Input:(None, 256) Output:(None,2, 256)	394752	3*[256*(256+257)+256]	The second layer GRU prediction model, double-layer GRU to some extent improves the model prediction performance.
TimeDistributed	Input:(None, 2, 256) Output: (None, 2, 4)	1028	256*4+4 (Input*Output+Output)	The distributed temporal feature representation is mapped to the sample marker space for full connection in the temporal dimension.

4 Multi-layer optimized GRU prediction method based on LOWESS smoothing and Random Forest

The prediction method developed in this paper mainly includes four steps: data smoothing, feature selection, prediction, and outcome analysis.

Step1: **Data smoothing**. Five different smoothing methods of origin software, LOWESS, LOESS , PF, SG, and AAv smoothing, are respectively used to process the daily power generation, so as to compare the prediction results. Sort the root mean square error between the smoothed data and the actual data in descending order.

Step2: **Random Forest feature selection**. Random Forest is introduced for feature selection, Python is used to sort and screen the importance of factors affecting photovoltaic power generation.

Step3: **Model prediction**. The data set is divided into train set and test set, the results of Random Forest feature selection are combined with smoothed data to form a new set of high-quality data set, i.e., the results of Random Forest feature selection are used as the input features of the model. On the basis of GRU model, the RepeatVector and TimeDistributed layers are added for optimization. The optimized GRU model is then employed for prediction.

Step4: **Model evaluation and comparison**. The experimental results are split into vertical comparison and horizontal comparison for analysis.

(1) Vertical comparison: The smoothing method with the least error is first selected. We compare this smoothing method with the prediction results with no smoothing.

(2) Horizontal comparison: Under the GRU model, we conduct the comparison between smoothed and unsmoothed, the comparison between optimized and un-optimized, and the comparison between the accuracies of GRU, RNN, and LSTM prediction models.

The flow chart of prediction steps is shown in Fig.3.

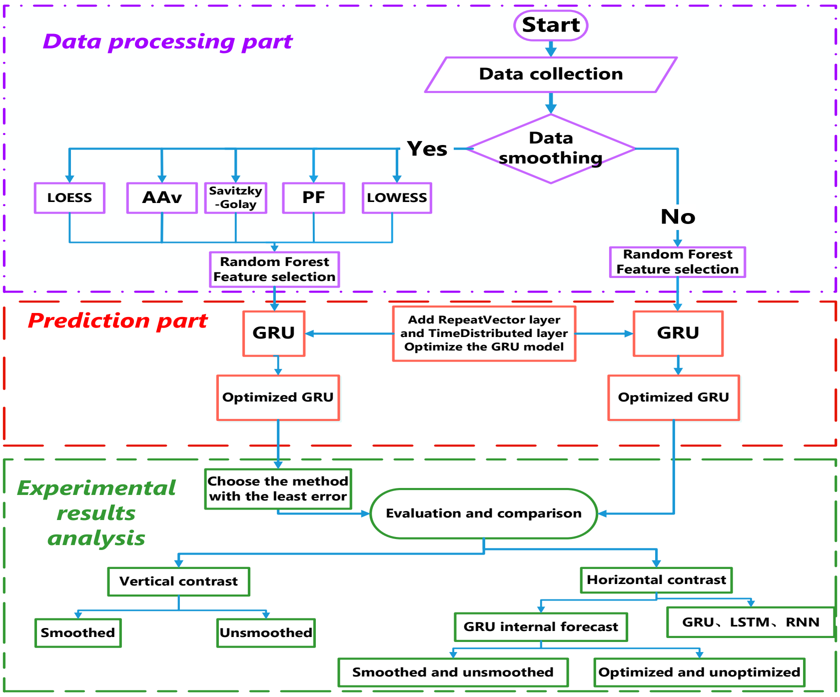

Fig.3. Flow chart of photovoltaic power generation prediction.

5 Case study

5.1 The experimental data

This experiment uses the daily power generation data of the photovoltaic power station of Xuhui District Government in Shanghai, China from January 1, 2015, to December 31, 2016 (train set) and January 2017 (test set) in smart PV website (https://www.lvsedianli.com/perHome.html) [35], with an interval of 24 hours.

The natural factors affecting the daily photovoltaic power generation are the data of Shanghai meteorological station from the national greenhouse system website, and the interval is also 24 hours. 11 natural factors affecting photovoltaic power generation are cumulative precipitation from 20 to 20 o'clock (mm) (hereafter, referred to as cumulative precipitation), average wind speed (m/s), maximum wind speed (m/s), average temperature (℃), daily maximum temperature (℃), daily minimum temperature (℃), sunshine hours (h), daily cumulative radiation (MJ/m^2), average relative humidity (%), minimum relative humidity (%), evaporation (mm).

5.2 Evaluation metrics

Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), NRMSE (Normalized Root Mean Squared Error) and R^2 (R-Squared) are selected as the metrics for evaluating the GRU model. Among these six evaluation metrics, the value of R^2 ranges from 0 to 1; and, the closer it is to 1, the better the fit of the model is. A smaller value of the remaining five evaluation criteria implies a higher prediction accuracy. Where y_i of formula (14-17) represents the real value, $y_i^{'}$ of formula (14-17) refers to the predicted value. In formula (18), $y_{\max} - y_{\min}$ means the full distance of the true value. Var denotes the variance value of formula (19), $n = 31$.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_i^{'}\right| \text{ ……………………………………….(14)}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i^{'}\right)^2 \text{ ……………………………………..(15)}$$

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-y_i^{'}\right)^2} \quad \text{.....................................}(16)$$

$$MAPE=\frac{1}{n}\sum_{i=1}^{n}\frac{\left|y_i-y_i^{'}\right|}{y_i}\times 100\% \quad \text{...............................}(17)$$

$$NRMSE=\frac{RMSE}{y_{\max}-y_{\min}} \quad \text{.....................................}(18)$$

$$R^2=1-\frac{RMSE}{Var} \quad \text{.......................................}(19)$$

5.3 Experimental setup

(1) Parameter settings of Random Forest

The accuracy of the Random Forest is mainly determined by the number of decision trees, the maximum depth of decision trees, the setting of random numbers, and the minimum sample number of leaf nodes. The parameter settings of Random Forest are shown in Table 3.

Table 3. Parameter settings of Random Forest.

Algorithm	Parameter meaning	Parameter	The parameter value	Setting basis
Random Forest	the number of decision trees	*n*_estimators	200	Specify the number of classifiers. If the number is too small, it is not fitted, and the training rate is too much, it needs to be compromised.
	the maximum depth of decision trees	max_depth	3	The common value range is 10~100, which can be modified appropriately when there are many sample sizes and characteristic quantities.
	the setting of random numbers	random_state	42	It is used to ensure that the experiment is divided into the same training set and test set every time.
	the minimum sample number of leaf nodes	min_samples_leaf	2	It is related to decision tree pruning, which is generally set to 1, it can be increased under the condition of a large sample size.
	the number of unit layers	num_layers	2	The default is 1 layer, if there are 2 layers, two GRUs are stacked together to form a unit.

(2) Parameter settings of GRU

The accuracy of GRU prediction model mainly depends on the number of neurons, the number of unit layers, time step, hidden layer width, and iteration times. The parameter settings of GRU are shown in Table 4.

Table 4. Parameter setting of GRU.

Model	Parameter meaning	Parameter	The parameter value	Setting basis
GRU	the number of neurons	unit	256	It is a key parameter affecting the accuracy and cannot be increased indefinitely.
	the number of unit layers	num_layers	2	The default is 1 layer, if there are 2 layers, two GRUs are stacked together to form a unit.
	time step	time_step	2	The difference between the two-time points before and after, and this experiment is the prediction of daily power generation.
	hidden layer width	batch_size	256	The number of statements entered into the GRU at one time, there is no fixed value.
	iteration times	epochs	50	It is related to the computing capacity of the computer, and too many iterations are time-consuming and labor-intensive.

5.4 Experiment Results

In our experiments, the results are divided into vertical comparison and horizontal comparison. The vertical comparison refers to the comparison between the introduction of data smoothing and the non-introduction of data smoothing. For the prediction resulting from the involvement of data smoothing, the specific steps including LOWESS, LOESS, PF, SG, and AAv can smooth the actual data, then utilize Random Forest to select the features of smoothed data, and finally combine the smoothed data with the results of feature selection and substitute them into the GRU model for prediction. The non-introduction of data smoothing indicates that the actual data is directly featured through Random Forest and then is substituted in the GRU prediction model.

The horizontal comparison is divided into the internal comparison of GRU and the external comparison between GRU, RNN and LSTM. The internal comparison based on GRU model is the comparison between prediction results from smoothed and unsmoothed data under the same feature selection results. In addition, the internal comparison also includes the compared results of optimized GRU and un-optimized GRU models. Moreover, three prediction models with the same data processing results are compared in terms of prediction accuracy.

5.4.1 Vertical comparison

(1) The comparison with different data smoothing methods

Considering that the daily photovoltaic power generation is subject to solar radiation, temperature, and other factors, which have a large fluctuation ranges, the data smoothing technologies are introduced to reduce the noise and fluctuation range. We use LOWESS, LOESS, PF, SG, and AAv to process the daily power generation data. Then, we use Random Forest for feature selection. Since different data smoothing methods have different results of feature selection and prediction, it is necessary to compare and analyze the final prediction results, which are shown in Table 5. (Aim to find the smoothing method with the best prediction results by adopting four evaluation metrics)

Table 5. Comparison of prediction errors of five data smoothing methods.

Smoothing methods	RMSE	Random Forest feature selection results (in order of importance)	GRU model prediction results			
			RMSE	MSE	MAE	MAPE
Savitzky-Golay	6.271	① daily cumulative radiation ② daily maximum temperature ③ average relative humidity	7.263	52.750	6.589	120.148
LOWESS	**6.472**	**① daily maximum temperature** **② daily cumulative radiation** **③ daily minimum temperature**	**2.352**	**5.530**	**1.851**	**19.715**
LOESS	6.993	① daily maximum temperature ② daily cumulative radiation ③ daily minimum temperature ④ evaporation	5.163	26.658	4.182	Inf
AAv	7.339	① daily cumulative radiation ② daily maximum temperature ③ evaporation	4.486	20.120	3.630	29.321
PF	8.998	① daily cumulative radiation ② daily maximum temperature ③ evaporation	4.910	24.109	3.822	27.343

According to Table 5, we find that when LOWESS smoothing method and three features (daily maximum temperature, daily cumulative radiation, and daily minimum temperature) are selected by Random Forest, GRU prediction model can obtain the lowest experimental error. In fact, RMSE, MSE, MAE and MAPE represent the smoothed prediction error values. We find that MSE, MAE and MAPE are all lowest under LOWESS smoothing method. The RMSE values in the second column of Table 5 represent the error between the smoothed data and actual data. It is obvious that SG smoothing with the lowest RMSE has the largest prediction error. At this time, the prediction error of PF smoothing with highest RMSE is relatively low. Therefore, improving the accuracy of

GRU prediction model but not changing the attributes and authenticity of actual data as much as possible is the key to this experiment. Compared with other smoothing methods, the RMSE of LOWESS smoothing ranks second, which has less error with the actual data and the lowest prediction error, LOWESS smoothing method then reasonably become our experiment selection.

(2) Comparison of unsmoothed and smoothed experimental results

In this section, we discuss the comparison between smoothed and unsmoothed experimental results under three different contexts (Random Forest feature selection, prediction results, and prediction errors) to reflect the importance of data processing.

A. Comparison analysis in Random Forest feature selection

We use Random Forest to rank and filter the feature importance of actual data and smoothed data. The results of feature importance ranking are presented in Fig.4 and Fig.5.

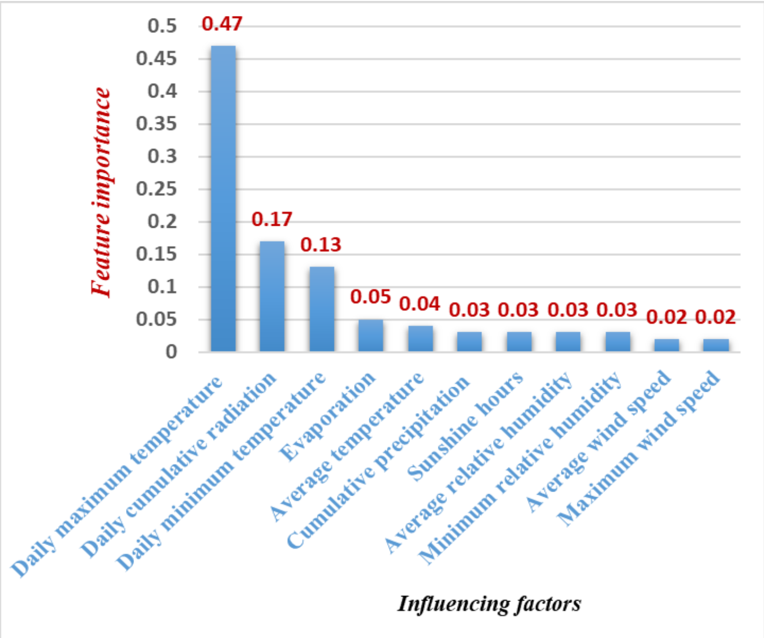

Fig.4. Random Forest feature selection results (Data after LOWESS smoothing).

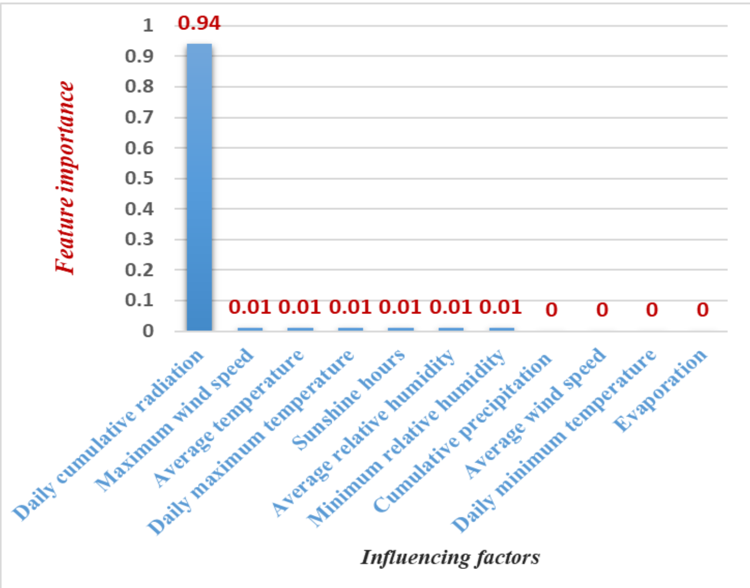

Fig.5. Random Forest feature selection results (Actual data).

As shown in Fig.4, for the LOWESS smoothed data, the results of Random Forest feature selection indicate a hierarchical gradient downward effect with a stepwise nature. According to the order of importance, the filtered features are daily maximum temperature, daily cumulative radiation and daily minimum temperature. As for the feature selection results considering actual data in Fig.5, the seven features filtered are daily cumulative radiation, maximum wind speed, average temperature, daily maximum temperature, sunshine hours, average relative humidity and minimum relative humidity, which filters out more features than smoothed data. The ratio of daily cumulative radiation is very high, and the remaining impact factors are extremely low, even to 0, with no stepwise. Therefore, the feature selection results considering actual data are obviously not as good as that considering smoothed data, which not only proves the effectiveness of Random Forest feature selection but also demonstrates the improvement of data smoothing on Random Forest performance.

B. Comparison analysis between prediction results and actual data

This part is about the prediction results comparison using smoothed data, unsmoothed data and actual data. We learn from Part A that the feature selection results are distinct using the smoothed and unsmoothed data. After the smoothed data with the filtered features are combined to form a high quality data set, this data set is substituted into GRU model for prediction. The same steps are followed for the unsmoothed prediction. We note from Table 5 that the prediction error of LOWESS smoothing is the smallest. Then, we select the prediction results using LOWESS smoothed data for analysis and comparison. Fig.6 represents the comparison results between actual data and GRU prediction model. From the interval of data distribution, the interval of actual data is 0~27.5 kw · h, the interval of unsmoothed prediction results is 5~17.5 kw · h, while the interval of smoothed prediction results ranges from 2.5~27.5 kw · h. Obviously, the smoothed prediction results are more consistent with the interval of actual data. In terms of the goodness of fit, the trend of smoothed prediction results and actual data are basically the same, but the trend of unsmoothed prediction results has almost no correlation with actual data, then the goodness of fit and prediction effect of smoothed prediction results are much better than unsmoothed prediction results. In conclusion, LOWESS smoothing can effectively reduce the fluctuation range of data, Random is applicable to feature selection, which shows that the combination of LOWESS smoothing and Random Forest can improve quality of input data and prediction accuracy of GRU.

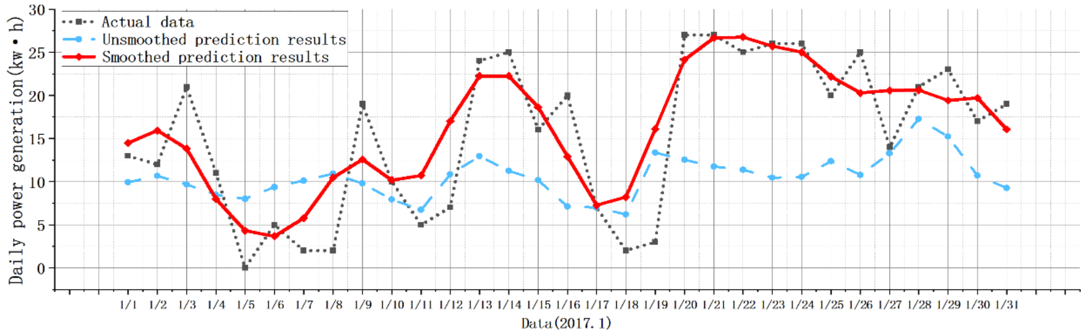

Fig.6. Comparison of actual data with prediction results.

C. Comparison analysis in GRU prediction errors

Different prediction results produce diverse prediction errors. In Table 6, smoothed prediction error is significantly smaller than unsmoothed prediction error, and the R^2 value of smoothed prediction is higher. A higher value of R^2 means a better goodness of fit for prediction model. The R^2 value of smoothed prediction is 0.955, which is close to 1, representing the high fitting ability of prediction method proposed in this paper. The R^2 value of unsmoothed prediction is 0.874, which is lower than that of smoothed prediction. The MAPE of unsmoothed prediction even reached infinity. Moreover, the RMSE of unsmoothed prediction is five times higher than smoothed prediction, and the MSE reaches 20 times. Then, smoothed prediction can result in a lower prediction error and a higher prediction validity.

Table 6. Comparison of unsmoothed and LOWESS smoothed prediction errors.

Evaluation metrics	Unsmoothed	**LOWESS smoothing**
RMSE	10.005	**2.352**
MSE	100.105	**5.530**
MAE	8.435	**1.851**
MAPE	Inf	**19.715**
NRMSE	0.370	**0.102**
R^2	0.874	**0.955**

After comparing the results of above three contexts, we find that data smoothing can effectively reduce the prediction error and improve the model prediction accuracy. Especially, Table 5 indicates that all five smoothing methods used in our experiment can effectively improve the prediction accuracy of the GRU prediction model. However, the unsmoothed prediction performance is poor because of the uncertainty and fluctuation of photovoltaic power generation from the influence of solar radiation and temperature.

5.4.2 Horizontal comparison

(1) Internal comparison of GRU model

A. Comparison analysis of smoothed and un-smoothed prediction errors

This section compares the prediction errors of GRU model using smoothed and unsmoothed data under the same LOWESS smoothed feature selection results. Table 7 shows that the unsmoothed prediction error is higher than that under smoothed prediction, the value of R^2 is significantly lower. The MSE of unsmoothed prediction is much larger than that of smoothed prediction, and its MAPE is also infinite. Although we use the same feature selection results, the unsmoothed prediction result is worse than that under smoothed prediction, which indicates that data smoothing techniques are meaningful for improving prediction accuracy.

Table 7. Comparison of smoothed and unsmoothed GRU internal prediction errors.

Evaluation metrics	**LOWESS smoothing**	Unsmoothed
RMSE	**2.352**	9.112
MSE	**5.530**	83.022
MAE	**1.851**	7.837
MAPE	**19.715**	Inf
NRMSE	**0.102**	0.337
R^2	**0.955**	0.824

B. Comparison analysis of optimized and un-optimized GRU prediction errors

Based on the consistent data processing results, the un-optimized GRU with only one layer is compared with the optimized GRU added RepeatVector layer and TimeDistributed layers. It can be seen from Table 8 that the prediction error of optimized GRU model is evidently less than that of un-optimized GRU model, the value of R^2 is higher than un-optimized GRU, the NRMSE is much lower than un-optimized, the RMSE and MAPE of optimized GRU is several times less than that of un-optimized GRU, so as to show that the optimization of prediction model is also a momentous mean to improve prediction accuracy.

Table 8. Comparison of optimized and un-optimized GRU prediction errors.

Evaluation metrics	**Optimized GRU**	Un-optimized GRU
RMSE	**2.352**	34.901
MSE	**5.530**	5.908
MAE	**1.851**	4.634
MAPE	**19.715**	82.858
NRMSE	**0.102**	0.872
R^2	**0.955**	0.831

(2) External comparison of GRU model

A. Comparison analysis of prediction results

Based on same data processing results, GRU, RNN and LSTM models are used for comparison. We first analyze the prediction results of three prediction models. As can be seen in Fig.7, although the trends of three models are roughly the same as the actual data, the prediction results of GRU are closer to the actual data than those of LSTM and RNN, and have the best goodness of fit with the actual data. Especially, both RNN and LSTM are significantly far from the actual data since January 21, 2017. Therefore, under the consistent data processing results, it is also particularly momentous to choose one appropriate prediction model.

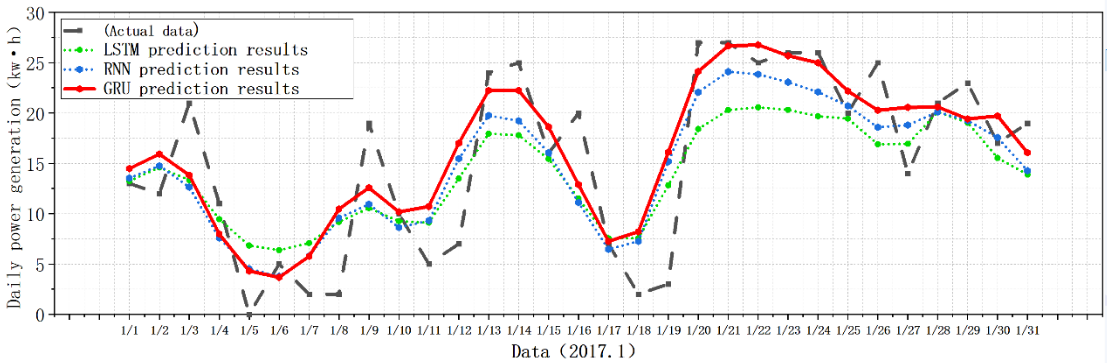

Fig.7. Comparison among prediction results of LSTM, RNN, GRU and actual data.

B. Comparison analysis of prediction errors

Finally, we compare the prediction errors of three models. Table 9 shows that GRU has the smallest prediction error and the highest accuracy. The RMSE, MSE and MAPE of these three models have little difference, but the gap in MAPE is slightly larger. Based on R^2 value, GRU is 0.950, RNN is 0.931, LSTM is 0.860, which prove that GRU has the best goodness of fit, followed by RNN, and LSTM has the worst fit.

Table 9. Comparison of prediction errors of LSTM, RNN, and GRU.

Evaluation metrics	**GRU**	LSTM	RNN
RMSE	**2.352**	3.131	2.602
MSE	**5.530**	9.801	6.771
MAE	**1.851**	2.794	2.241
MAPE	**19.715**	29.056	22.207
NRMSE	**0.102**	0.221	0.128
R^2	**0.950**	0.860	0.931

5.4.3 Impact of uncertainty

To fully analyze the impact of uncertainty on model performance, we consider both data processing and prediction models. In general, the uncertainty in data processing mainly refers to the unavoidable errors in the data due to the limitations of measurement means and machine anomalies. The uncertainty in prediction model mainly includes the model's own structure as well as the optimization algorithm [36]. In addition, the factors affecting the uncertainty of data processing also include natural climatic conditions.

A. Data processing

Fig.5 shows the uncertainty of actual data leads to the lack of ladder nature in the feature selection results. With the data smoothing technology, the results of feature selection are clearly enhanced. Therefore, data smoothing can effectively reduce the negative impact of uncertainty. As seen in Fig.6, the uncertainty of data causes serious deviations between actual data and prediction results, and the prediction accuracy is too low. However, the prediction accuracy is significantly improved after data smoothing and feature selection. Then data processing is meaningful to reduce the adverse effects of uncertainty on the prediction accuracy.

B. Prediction model

For prediction model, Table 9 demonstrates the importance of choosing a suitable prediction model. RNN and LSTM have slightly lower prediction accuracy than GRU due to their own gradient problem and complex parameters. Table 8 shows the model performance of optimized GRU added RepeatVector and TimeDistributed layers is clearly better than that of un-optimized GRU. Thus, choosing the appropriate prediction model and optimization method is also beneficial to reduce the influence of uncertainty.

In summary, data smoothing techniques reduce the fluctuation of data, Random Forest filters out reasonable feature selection results, which solve the uncertainty in data processing to some extent. Furthermore, the GRU model added RepeatVector and TimeDistributed layers improves the prediction accuracy of GRU model. Therefore, the developed forecasting method can effectively reduce the uncertainty impact of photovoltaic power generation on the prediction performance.

6 Conclusion

Aiming to enhance the accuracy and stability of forecasting, this paper considers the impact of data fluctuations on the prediction of photovoltaic power generation, and develops an optimized GRU forecasting method which includes data smoothing technology, feature selection, and optimization of GRU. Firstly, we uses different data smoothing technologies to reduce the fluctuation of actual data and choose the best one with the least prediction error. Secondly, to obtain higher-quality input data, Random Forest method is used to select natural factors affecting photovoltaic power generation. Finally, the RepeatVector layer and TimeDistributed layer are used to optimize the GRU model. Through case studies and experimental results, the conclusions are obtained as follows:

(1) Five different smoothing methods can all improve the prediction performance of GRU model. Among them, the LOWESS smoothing can generate the smallest prediction error.

(2) The Random Forest feature selection can simplify the number of features and optimize the prediction performance of GRU model.

(3) Under the consistent results of data processing, GRU model is more suitable for the prediction of photovoltaic power generation than LSTM and RNN.

(4) The GRU model with RepeatVector layer and Timedistributed layer has a better prediction performance than the un-optimized single-layer GRU.

In summary, compared with original GRU model, the proposed forecasting method in this paper improves data processing and optimizes GRU model. The applications of different data smoothing techniques reduce the fluctuation of daily power generation and improve the quality of input data. Random Forest selects the characteristics of natural factors affecting photovoltaic power generation. The optimized GRU with RepeatVector and TimeDistributed layers not only enriches GRU layer structure, but also effectively improves the prediction accuracy. In terms of practical application value, high-precision generation forecasting method is an effective approach for integrating solar energy resources into power grid. Our results show that the proposed method could be a useful tool to forecast the short-term photovoltaic power generation with an acceptable degree of accuracy. Besides, according to the prediction results, photovoltaic power plants can arrange future power generation, adjust electrovalence, and provide technical support to make timely and reasonable scheduling decision for power grid.

The integration of photovoltaics into the power grid present both opportunities and challenges. The instability of photovoltaic power generation has caused challenges and impacts on the power grid though it can reduce the pollution and loss caused by traditional power generation. Accurate prediction is conducive to the safe and stable and economic operation of photovoltaic plants after grid connection. Compared with the original GRU, the developed prediction method in this paper improves the prediction accuracy through smoothing technology, feature selection and optimized GRU, but there are still some limitations. For example, the loss of prediction accuracy, the reduction of error between smoothed data and actual data, the unity of the feature selection method, the feasibility of interval prediction, and better improvement of GRU prediction performance. These are all problems we intend to solve. In addition, the profound development of photovoltaic power generation needs the support of medium and long-term forecasting. In the future, we collect data suitable for medium and long-term forecasting, study relevant forecasting technologies, and use medium and long-term forecasting to make scientific decisions and plans for the development of photovoltaic power.

Acknowledgements

This work was supported by the National Natural Science Foundation of China [No. 72171126], Ministry of Education Project of Humanities and Social Science [No.20YJA630009], Social Science Planning Project of Shandong Province [No. 20CSDJ15]. The third author (Mingming Leng) was supported by the Faculty Research Grant of Lingnan University under the grant number DB21B1.

References

[1] Lin B, Zhu J. Policy effect of the Clean Air Action on green development in Chinese cities. J Environ Manage 2020;258:110036. https://doi.org/10.1016/j.jenvman.2019.110036.

[2] Jin S, Greaves D. Wave energy in the UK: Status review and future perspectives. Renew Sustain Energy Rev 2021;143:110932. https://doi.org/10.1016/j.rser.2021.110932.

[3] Gandoman FH, Abdel Aleem SHE, Omar N, Ahmadi A, Alenezi FQ. Short-term solar power forecasting considering cloud coverage and ambient temperature variation effects. Renew Energy 2018;123:793–805. https://doi.org/10.1016/j.renene.2018.02.102.

[4] Cheng L, Zang H, Wei Z, Ding T, Xu R, Sun G. Short-term Solar Power Prediction Learning Directly from Satellite Images with Regions of Interest. IEEE Trans Sustain Energy 2022;13:629-639. https://doi.org/10.1109/TSTE.2021.3123476.

[5] Rodríguez F, Genn M, Fontán L, Galarza A. Very short-term temperature forecaster using MLP and N-nearest stations for calculating key control parameters in solar photovoltaic generation. Sustain Energy Technol Assessments 2021;45. https://doi.org/10.1016/j.seta.2021.101085.

[6] Almonacid F, Pérez-Higueras PJ, Fernández EF, Hontoria L. A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator. Energy Convers Manag 2014;85:389–98. https://doi.org/10.1016/j.enconman.2014.05.090.

[7] Khan AT, Khan AR, Li S, Bakhsh S, Mehmood A, Zaib J. Optimally configured Gated Recurrent Unit using Hyperband for the long-term forecasting of photovoltaic plant. Renew Energy Focus 2021;39:49–58. https://doi.org/10.1016/j.ref.2021.07.002.

[8] Yang D. A guideline to solar forecasting research practice: Reproducible, operational, probabilistic or physically-based, ensemble, and skill (ROPES). J Renew Sustain ENERGY 2019;11. https://doi.org/10.1063/1.5087462.

[9] Guijo-Rubio D, Durán-Rosal AM, Gutiérrez PA, Gómez-Orellana AM, Casanova-Mateo C, Sanz-Justo J, et al. Evolutionary artificial neural networks for accurate solar radiation prediction. Energy 2020;210:118374. https://doi.org/10.1016/J.ENERGY.2020.118374.

[10] Dogan A, Birant D. Machine learning and data mining in manufacturing. Expert Syst Appl 2021;166:114060. https://doi.org/10.1016/J.ESWA.2020.114060.

[11] Jeffrey Kuo CF, Su TL, Jhang PR, Huang CY, Chiu CH. Using the Taguchi method and grey relational analysis to optimize the flat-plate collector process with multiple quality characteristics in solar energy collector manufacturing. Energy 2011;36:3554–62. https://doi.org/10.1016/J.ENERGY.2011.03.065.

[12] Adar M, Najih Y, Gouskir M, Chebak A, Mabrouki M, Bennouna A. Three PV plants performance analysis using the principal component analysis method. Energy 2020;207:118315. https://doi.org/10.1016/J.ENERGY.2020.118315.

[13] Jiang Z, Cai J, Moses PS. Smoothing control of solar photovoltaic generation using building thermal loads. Appl Energy 2020;277:115523. https://doi.org/10.1016/J.APENERGY.2020.115523.

[14] Mariani MC, Basu K. Local regression type methods applied to the study of geophysics and high frequency financial data. Phys A Stat Mech Its Appl 2014;410:609–22. https://doi.org/10.1016/J.PHYSA.2014.05.070.

[15] Quan Q, Cai KY. Time-domain analysis of the Savitzky–Golay filters. Digit Signal Process 2012;22:238–45. https://doi.org/10.1016/J.DSP.2011.11.004.

[16] Lin L, Ma S, Ma M. A group neighborhood average clock synchronization protocol for wireless sensor networks. Sensors (Switzerland) 2014;14:14744–64. https://doi.org/10.3390/S140814744.

[17] Duin RPW, Haringa H, Zeelen R. Fast percentile filtering. Pattern Recognit Lett 1986;4:269–72. https://doi.org/10.1016/0167-8655(86)90007-3.

[18] Tronarp F, Särkkä S. Iterative statistical linear regression for Gaussian smoothing in continuous-time non-linear stochastic dynamic systems. Signal Processing 2019;159:1–12. https://doi.org/10.1016/J.SIGPRO.2019.01.013.

[19] Rostami M, Berahmand K, Nasiri E, Forouzande S. Review of swarm intelligence-based feature selection methods. Eng Appl Artif Intell 2021;100:104210. https://doi.org/10.1016/J.ENGAPPAI.2021.104210.

[20] López G, Batlles FJ, Tovar-Pescador J. Selection of input parameters to model direct solar irradiance by using artificial neural networks. Energy 2005;30:1675–84. https://doi.org/10.1016/J.ENERGY.2004.04.035.

[21] Bas E, Yolcu U, Egrioglu E. Picture fuzzy regression functions approach for financial time series based on ridge regression and genetic algorithm. J Comput Appl Math 2020;370:112656. https://doi.org/10.1016/J.CAM.2019.112656.

[22] Rafati A, Joorabian M, Mashhour E, Shaker HR. High dimensional very short-term solar power forecasting based on a data-driven heuristic method. Energy 2021;219:119647. https://doi.org/10.1016/J.ENERGY.2020.119647.

[23] Dutta S, Li Y, Venkataraman A, Costa LM, Jiang T, Plana R, et al. Load and Renewable Energy Forecasting for a Microgrid using Persistence Technique. Energy Procedia 2017;143:617–22. https://doi.org/10.1016/J.EGYPRO.2017.12.736.

[24] Wang M, Peng J, Luo Y, Shen Z, Yang H. Comparison of different simplistic prediction models for forecasting PV power output: Assessment with experimental measurements. Energy 2021;224:120162. https://doi.org/10.1016/J.ENERGY.2021.120162.

[25] Bin Shams M, Haji S, Salman A, Abdali H, Alsaffar A. Time series analysis of Bahrain's first hybrid renewable energy system. Energy 2016;103:1–15. https://doi.org/10.1016/J.ENERGY.2016.02.136.

[26] Jung Y, Jung J, Kim B, Han SU. Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: Case study of South Korea. J Clean Prod 2020;250:119476. https://doi.org/10.1016/J.JCLEPRO.2019.119476.

[27] Ahmed R, Sreeram V, Mishra Y, Arif MD. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. Renew Sustain Energy Rev 2020;124:109792. https://doi.org/10.1016/J.RSER.2020.109792.

[28] Han S, Qiao Y hui, Yan J, Liu Y qian, Li L, Wang Z. Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network. Appl Energy 2019;239:181–91. https://doi.org/10.1016/J.APENERGY.2019.01.193.

[29] Pan C, Tan J, Feng D. Prediction intervals estimation of solar generation based on gated recurrent unit and kernel density estimation. Neurocomputing 2021;453:552–62. https://doi.org/ 10.1016/ j.neucom.2020.10.027.

[30] Abdel-Basset M, Hawash H, Chakrabortty RK, Ryan M. PV-Net: An innovative deep learning approach for efficient forecasting of short-term photovoltaic energy production. J Clean Prod 2021;303:127037. https://doi.org/10.1016/J.JCLEPRO.2021.127037.

[31] Nie Y, Zamzam AS, Brandt A. Resampling and data augmentation for short-term PV output prediction based on an imbalanced sky images dataset using convolutional neural networks. Sol Energy 2021;224:341–54. https://doi.org/10.1016/j.solener.2021.05.095.

[32] Nguyen NQ, Bui LD, Doan B Van, Sanseverino ER, Cara D Di, Nguyen QD. A new method for forecasting energy output of a large-scale solar power plant based on long short-term memory networks a case study in Vietnam. Electr Power Syst Res 2021;199:107427. https://doi.org/10.1016/j.epsr.2021.107427.

[33] Idris A, Rizwan M, Khan A. Churn prediction in telecom using Random Forest and PSO based data balancing in combination with various feature selection strategies. Comput Electr Eng 2012;38:1808–19. https://doi.org/10.1016/J.COMPELECENG.2012.09.001.

[34] Conlin R, Erickson K, Abbate J, Kolemen E. Keras2c: A library for converting Keras neural networks to real-time compatible C. Eng Appl Artif Intell 2021;100:104182. https://doi.org/10.1016/J.ENGAPPAI.2021.104182.

[35] https://www.lvsedianli.com/perHome.html.

[36] Barukčić M, Hederić, Hadžiselimović M, Seme S. A simple stochastic method for modelling the uncertainty of photovoltaic power production based on measured data. Energy 2018;165:246–56. https://doi.org/10.1016/J.ENERGY.2018.09.134.

LOWESS Smoothing and Random Forest Based GRU Model: A Short-term Photovoltaic Power Generation Forecasting Method

Yeming Dai[1*], Yanxin Wang[1], Mingming Leng[2], Xinyu Yang[1], Qiong Zhou[1]

1. *School of Business, Qingdao University, Qingdao 200071, China*

2. *Faculty of Business, Lingnan University, Hong Kong.*

Abstract: Accurate prediction of photovoltaic power generation is vital to guarantee smooth operation of photovoltaic power stations and ensure the electricity consumption of end users. As a good forecasting tool, Gated Recurrent Unit method has been widely used in different forecasting areas. However, the existing studies ignore the impact of data fluctuations on prediction accuracy since photovoltaic power generation is intermittent and uncertain, then the prediction results of Gated Recurrent Unit are facing challenges. To fill the gaps and enhance prediction accuracy, this paper develops an improved Gated Recurrent Unit photovoltaic generation prediction method. Several different data smoothing techniques are introduced and compared to reduce fluctuations, Random Forest method is used for feature selection, and RepeatVector layer extended by attribute dimensions and TimeDistributed layer with full connectivity are utilized to optimize the Gated Recurrent Unit model. A real-world case from the photovoltaic power plant in Xuhui District, Shanghai, China, is adopted to evaluate the performance of proposed method. After comparing with the prediction results of Recurrent Neural Networks and Long Short-Term Memory, and the actual data as well, it is found that the proposed prediction method can effectively improve the prediction accuracy of photovoltaic power generation.

Keywords: Photovoltaic Power Generation; Prediction; Locally Weighted Scatterplot Smoothing; Random Forest; Gated Recurrent Unit

1 Introduction

Energy saving and carbon emission reduction is a topic of common concern all over the world. In China, the government has drawn the historic lessons from the failure of energy saving and carbon emission and sought a climate governance road with Chinese characteristics. Moreover, the clean air action plan, carbon peaking and carbon neutral strategies to achieve energy conservation and environment protection goals were claimed in 2021 [1]. Other countries are also making contributions to environment protection now. For instance, UK announced a policy for its Net Zero Strategy to support the transition of British businesses and consumers to clean energy and green technologies [2].

In particular, as a complementary and multi-system coordinated energy supply and consumption mode, renewable energy (wind energy, water energy, solar energy, geothermal energy, etc.) in the electricity market has become an important means to improve the efficiency of energy utilization and make great contributions to energy saving and carbon emission reduction in recent

[*] *Corresponding author: E-mail: yemingdai@163.com.*

Nomenclature

$(x, \hat{y})$	Center value of regression line	kNN	k-Nearest Neighbor algorithm
i	Smoothing window	MAE	Mean Absolute Error
h_i	Smoothing coefficient	MAPE	Mean Absolute Percentage Error
$w(i, j)$	Weight value	ML	Machine Learning
$f(i, j)$	Original data	MLP	Multilayer Perceptron
$g(x, y)$	Smoothed data	MSE	Mean Square Error
x^t	Current input	LOESS	Locally Estimated Scatterplot Smoothing
h^{t-1}, h^t	Hidden state	LOWESS	Locally Weighted Scatterplot Smoothing
y^t	Output	LR	Linear Regression
$y_i, y_i^{'}$	real value, predicted value	LSTM	Long-Short Term Memory
$y_{\max} - y_{\min}$	Full distance of real value	NN	Neural Network
AAv	Adjacent Average method	NRMSE	Normalized Root Mean Squared Error
AI	Artificial Intelligence	OOB	Out Of Bag
ARD	Automatic Relevance Determination method	PCA	Principal Component Analysis
CNN	Convolutional Neural Network	PF	Percentile Filtering
DEM	Digital Elevation Model	R^2	R-Squared
EANN	Evolutionary Artificial Netural Network	RMSE	Root Mean Square Error
ES	Exponential Smoothing	RMSPE	Root-Mean-Square Percentage Error
GRA	Grey Relational Analysis	RNN	Recurrent Neural Networks
GRU	Gated Recurrent Unit	SG	Savitzky-Golay Smoothing
		SVM	Support Vector Machine
		Var	Variance value

years. Compared with other renewable energy sources, solar energy has the advantages of high flexibility, superior adaptability, and low development cost, which has broaden the social development potentials and prospects [3]. Photovoltaic power generation then has been a priority since it can convert solar energy into electricity. With the popularity of photovoltaic power generation, more and more countries and regions have been implementing their initiatives of integrating photovoltaic power generation into power grids, which has resulted in an increase in daily power supply and a reduction in carbon emissions. Nevertheless, the development of renewable energy is a combination of opportunities and challenges. One of the major challenges is that renewable energy is affected by natural factors and can be unstable when connected to the power grid [4]. The photovoltaic power generation is greatly affected by weather factors. This results in its intermittent defects, which, in turn, is not conducive to the stable operation of power grid. Therefore, accurate prediction results of photovoltaic power generation can make appropriate operations and scheduling efforts and alleviate the instability issues.

Existing studies have shown that the length of prediction period has an important impact on the prediction accuracy and application scenarios. According to the length of prediction period, forecasting methods can be divided into three types: very short-term [5], short-term [6], and medium and long-term [7]. For very short-term forecasting, it is accurate in seconds to minutes and is suitable for real-time dispatching different sizes grids, so as to reserve spare capacity for power grid in a timely manner [6]. In terms of short-term forecasting with the range from hour to day, it concerns economic dispatch and decision making of power grid to balance power market transactions whose meaning is adjusting the unit combination scheme and optimizing the generation plan [5]. In addition, medium and long-term prediction focuses on day to week/month/year based on the very short-term and short-term forecasting, which provides a long range plan for power grid and shows the ability for the equipment maintenance and the siting of new energy base stations. Many facts have shown that medium and long-term forecasting is filled with the big picture concept, which is also the development direction of forecasting in many economic and industrial areas [7]. In general, the three different prediction methods are adopted according to the characteristics of power grid in terms of time scale, application scope, and purpose. Since short-term prediction concerns the economic dispatch and decision making of power grid to balance power market transactions, our paper focuses more on short-term prediction of photovoltaic power generation in order to make better generation planning and more timely power plant offers to the dispatch center, as well as to improve the security and economy of power grid [8].

The short-term prediction process of photovoltaic power generation mainly includes data processing and model prediction [9]. Data processing is an important prerequisite for prediction, which is shown as data cleaning, data integration, data transformation, and data protocol [10-12]. However, existing literatures ignore the importance of data fluctuations, there are still fewer analysis of data fluctuation in the prediction of photovoltaic power generation [13]. Therefore, it is necessary to pay more attention to the research on these smoothing methods. To the best of our knowledge, Locally Weighted Scatterplot Smoothing (LOWESS), Locally Estimated Scatterplot Smoothing, LOESS) [14], Savitzky-Golay Smoothing (SG) [15], Adjacent Average method (AAv) [16], and Percentile Filtering (PF) [17] have been representative smoothing methods with convenient operation and rapid arithmetic in recent years and have not been applied to the prediction of photovoltaic power generation [10,18]. Accordingly, our paper applies LOWESS, LOESS, SG, AAv, and PF smoothing methods to the data processing in the prediction of photovoltaic power generation to improve the data quality. We also compare the above five smoothing methods based on different evaluation metrics, and then filter out the best method with the lowest prediction error, so as to achieve the goal of improving prediction accuracy.

In addition, we also use feature selection in the data processing. Feature selection refers to the process of selecting some effective features from existing features to reduce the data dimension, mainly including filter, embedded, and wrapper [19]. For example, in the filter method, Ref. [20] used the Automatic Relevance Determination method (ARD) to point out the most relevant input for the accurate monthly average daily solar radiation prediction, Ref. [21] used ridge regression algorithm in the embedded method. These feature selection methods improve the prediction accuracy of the model, but the performance and calculation speed are not as good as the wrapper method. Wrapper method mainly includes Random Forest, SVM(Support Vector Machine), and k-Nearest Neighbor algorithm (kNN) [19]. Compared with SVM and kNN algorithm, Random Forest can process high-dimensional data, deal with many problems such as classification, feature selection, and regression [22]. Existing research and experimental results on Random Forest have exposed that Random Forest feature selection can effectively improve the prediction accuracy [22]. Then, this paper also applies Random Forest for data processing to analyze the factors that affect power generation and obtain higher input data quality.

The next step after data processing is the prediction with a suitable model. At present, the prediction models of photovoltaic power generation are mainly divided into four categories: persistence forecast of "today equals tomorrow" [23], physical model based on terrain research [24], statistical techniques related to time series [25], and Artificial Intelligence (AI) prediction represented by Machine Learning (ML) [26]. The first three categories possess some prediction flaws because of an increase in time span and abnormal sudden changes in weather [27]. ML mainly includes Linear Regression (LR), SVM, and Neural Network (NN). Especially, NN can efficiently process a large amount of data, improve the prediction accuracy and solve the defects in persistence forecast, physical model, and statistical techniques [26]. Today, NN has become the primary choice of prediction methods in many fields.

As one of NN, Gated Recurrent Unit (GRU) can solve gradient disappearance and explosion of Recurrent Neural Networks (RNN), simplify parameters of Long-Short Term Memory (LSTM) [7,29], which shows excellent performance in prediction and obtains some further improvement [30]. However, the impact of data smoothing on prediction accuracy is not considered, the input data in multiple dimensions is not analyzed, and the information is shared diversely. There is still much room for the improvement of GRU model. Therefore, we develop an improved GRU model by introducing RepeatVector layer and TimeDistributed layer to optimize the GRU model, which is different from other optimized GRU models without diverse and multidimensional improvements in term of model hierarchy. To provide an improved reflection made to the GRU model in this paper and indicate the differences with other NN-based literatures, we have made a comprehensive comparison in Table 1. Moreover, our paper introduces data smoothing techniques while none of the remaining references introduces them. For feature selection, we use Random Forest for feature analysis which is similar to [11,30,32] while the other literatures do not perform feature selection. In summary, compared to other NN-based references, the innovation of this paper mainly includes the application of data smoothing techniques and the optimization of GRU model.

Based on the above analysis, we first introduce LOWESS, LOESS, SG, PF, and AAv data smoothing methods and compare them to filter the best method with the least error. Secondly, considering the variety of natural factors that affect photovoltaic power generation, we use Random Forest for feature selection. Finally, we optimize the GRU model for prediction by using RepeatVector layer and TimeDistributed layer. The main contributions of this paper are as follows:

(1) We consider different data smoothing technologies to reduce the data fluctuation of daily photovoltaic power generation. We also compare these data smoothing techniques to find the best smoothing method that has the least prediction error.

(2) We use Random Forest to extract the characteristics of natural factors affecting daily photovoltaic power generation.

(3) We add RepeatVector layer and TimeDistributed layer into the GRU model to improve its prediction accuracy.

(4) We utilize the dataset from Shanghai, China and three prediction models to verify the accuracy and feasibility of our proposed method.

The remainder of this paper is arranged as follows: Section 2 introduces the basic concepts of the five smoothing methods mentioned in this paper, the generation process of Random Forest, and original GRU model. Section 3 describes the structure of optimized GRU with the addition of RepeatVector layer and TimeDistributed layer. Section 4 gives the specific forecasting steps. In Section 5, we provide details for our case study, which include a description of experimental data, the selection of evaluation metrics, the setting of model parameters, and our comprehensive analysis of experimental results. Section 6 presents a summary of our works in this paper.

Table 1. Comparison between this paper and other NN-based references.

Study	Year	Prediction period	Data source	Data processing	Prediction model	Optimization method	Prediction error
This paper	2021	Short-term	China	Data smoothing and Random Forest feature selection	GRU	Use RepeatVector layer and TimeDistributed layer to optimize GRU	RMSE: 2.352 MAE: 1.851 MAPE: 19.715 MSE: 5.530 R^2: 0.955 NRMSE: 0.102
[5]	2021	Very short-term	Basque	N-nearest-station model	MLP (Multilayer Perceptron)	Optimize the length of the input window	RMSE: 0.2515 R^2: 0.9985
[9]	2020	Short-term	Spain	Pearson correlation	EANN (Evolutionary Artificial Netural Network)	Evolutionary algorithm	MBE: 0.30 MAE: 33.46 RMSE: 0.9709
[26]	2020	Medium and long-term	Korea	DEM (Digital Elevation Model)	LSTM-RNN	Use LSTM layer to optimize stacked RNN	R^2: 0.724 RMSE: 14.003 NRMSE: 7.416 MAPE: 10.805
[28]	2019	Medium and long-term	China American	Copula function	LSTM	Joint prediction (wind and photovoltaic power generation)	MAPE: 6.65 RMSPE: 8.43
[30]	2021	Short-term	Australia	Remove outliers and feature normalization	Conv-GRU	Use convolutional layers to optimize GRU	R^2: 0.8938 RMSE: 2.630
[31]	2021	Very short-term	American	Data augmentation techniques	CNN (Convolutional Neural Network)	Adam algorithm	RMSE: 3.259
[32]	2021	Short-term	Vietnam	Pearson correlation and remove outliers	LSTM	Replace the historical weather data entered into the model with forecast weather data	MSE: 56.348 RMSE: 7.507 MAE: 4.743 MAPE: 9.881

2 Theoretical basis

This section presents the base methods used in data processing and prediction. The data processing mainly includes LOWESS, LOESS, SG, AAv and PF in data smoothing and Random Forest in feature selection. We specify the original GRU model as a base prediction model.

2.1 Data smoothing

2.1.1 LOWESS smoothing

① Definition: take point x as the center, intercept a section of proportional data forward and backward respectively, make weighted linear regression with weight function W for this section of data, $(x,\hat{y})$ is the center value of the regression line, $\hat{y}$ represents the corresponding value after fitting the curve, all n data can make n weighted regression lines. The connection of the central value of each regression line is the LOWESS smooth curve of this data [14].

② Weight function W

The commonly used weight function is the cubic function $W(x)$.

$$W_{(x)} = \begin{cases} (1-|x|)^3, & if \ |x|<1; \\ 0, & if \ |x|\ge 1. \end{cases} \quad \text{(1)}$$

2.1.2 LOESS smoothing

LOESS smoothing divides the samples into multiple cells, performs polynomial fitting on the interval samples, repeats the fitting process continuously, and obtains multiple weighted regression curves, finally connects the center of the curve to obtain the smooth curve [14].

2.1.3 Savitzky-Golay smoothing

Savitzky-Golay smoothing is based on the least square principle and performs k -order polynomial fitting for data points in a certain length window [15]. In formula (2), i represents the i th smoothing window, h_i represents the smoothing coefficient, h_i/H is solved by the least square method.

$$x_{k,smooth} = \bar{x}_k = \frac{1}{H}\sum_{i=-w}^{+w} x_{k+i} \cdot h_i \quad \text{(2)}$$

2.1.4 AAv smoothing

AAv smoothing is a smoothing method for calculating the arithmetic mean of several adjacent data [16], use neighborhood average (formula (3)) or weighted average (formula (4)) for smoothing. $w(i,j)$ represents weight value, $f(i,j)$ represents the original data, M denotes the number of adjacent data, $g(x,y)$ is the smoothed data.

$$g(x,y) = \frac{1}{M}\sum_{i,j\in s} f(i,j) \quad \text{(3)}$$

$$g(x,y) = \frac{1}{M}\sum_{i,j\in s} w(i,j) f(i,j) \quad \text{(4)}$$

2.1.5 PF smoothing

PF smoothing is a non-linear smoothing method that calculates a specified quantile value for local data and replaces the original data with this quantile value, which is suitable for signal smoothing with pulse characteristics [17].

2.2 Random Forest

The generation process of Random Forest is to put back samples from the original training samples to obtain numerous subsets. These subsets train different base classifiers, and the optimal classification results are determined by the voting of the base classifier [33]. The evaluation of Random Forest performance mainly uses Out Of Bag error (OOB error). When the total number of samples=N, the importance of features is calculated by formula (5), $errOOB_1$ represents the Out Of Bag data error of each decision tree, $errOOB_2$ represents the Out Of Bag data error after adding noise interference.

$$\sum \frac{errOOB_2 - errOOB_1}{N} \quad \text{(5)}$$

2.3 GRU prediction model

The realization process of GRU is as follows: combine the current input x^t and the hidden state h^{t-1} passed down from the previous node to obtain the output y^t of the current hidden node and the hidden state h^t passed to the next node [7]. The basic GRU model has only one layer, and there is room for optimization. Fig.1 shows the network structure of GRU model, and Fig.2 shows the module internal structure of GRU model.

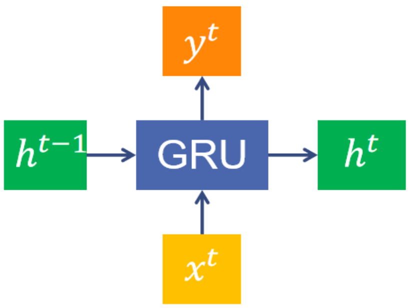

Fig.1. GRU model network structure.

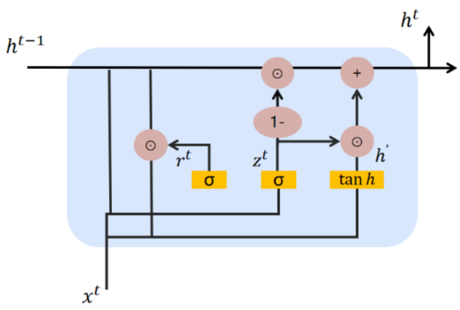

Fig.2. GRU model internal structure.

The procedure of GRU is as follows [29]:

Step 1: Calculate the gate: obtain the weight and parameter status of the update gate and reset gate through h^{t-1} and x^t, σ represents the ***sigmoid*** activation function.

① Update gate:

$$z_t = \sigma(W_z \cdot [h^{t-1}, x^t]) \quad (6)$$

The parameter representation of the update gate:

$$[(x_{dim} + h_{dim}) * h_{dim} + h_{dim}] \quad (7)$$

② Reset gate:

$$r_t = \sigma(W_r \cdot [h^{t-1}, x^t]) \quad (8)$$

The parameter representation of the reset gate:

$$2 * [(x_{dim} + h_{dim}) * h_{dim} + h_{dim}] \quad (9)$$

③ Total parameter formula expression: it is obtained by adding formula (7) and formula (9):

$$3 * [(x_{dim} + h_{dim}) * h_{dim} + h_{dim}] \quad (10)$$

Step 2: Capture information: directly extract the local information $h^{t-1'}$ from the long-term information h^t through r_t.

$$h^{t-1'} = h^{t-1} \otimes r_t \quad (11)$$

Step 3: Obtain the current information $h^{'}$: splice $h^{t-1'}$ and x^t, and use $\tan h$ to form compression.

$$h^{'} = \tan h(W \cdot [h^{t-1'}, x^t]) \quad (12)$$

Step 4: Generate a new output h^t: fuse h^{t-1} and $h^{'}$ by taking a part of each weight. Among them, $(1-z_t) \otimes h^{t-1}$ is the selective "forgetting" of h^{t-1} unimportant information, $z_t \otimes h^{'}$ is the selective memory of $h^{'}$.

$$h^t = (1-z_t) \otimes h^{t-1} + z_t \otimes h^{'} \quad (13)$$

3 Multi-layer optimized GRU model with RepeatVector layer and TimeDistributed layer

Original GRU model has a GRU layer only, which possesses randomness and uncertainty. We add RepeatVector layer and TimeDistributed layer to make the GRU layer more diversified.

First, we add the RepeatVector layer to the GRU hierarchy to ensure the same vector in each time step, which specifically refers to increase the dimension of input data and add attribute dimension, then the model can be analyzed in all aspects from various dimensions [34]. The parameter is represented by *n*. For example, when *n*=3, it means that the dimension of input data increases to 2 dimensions;

Secondly, we add the TimeDistributed layer using time series for tensor operations to obtain a better weight information sharing, and the same fully connected layer can be applied to each time step, thus achieving full connectivity in the time dimension [34]. The structure layer and setting basis are shown in Table 2.

Table 2. GRU model structure layer and setting basis.

Model layer (in order)	Input and output	Number of parameters (201216+394752+1028=596996)	Calculation method (obtained according to formula (10))	Setting basis
GRU_1	Input:(None, 256) Output: (None, 256)	201216	3*[256*(256+5)+256]	Layer 1 GRU prediction model
RepeatVector	Input:(None, 256) Output:(None, 2, 256)	0	0	Repeated input of potential vectors can increase attribute dimensions, which is beneficial to multi-dimensional analysis of the model.
GRU_2	Input:(None, 256) Output:(None,2, 256)	394752	3*[256*(256+257)+256]	The second layer GRU prediction model, double-layer GRU to some extent improves the model prediction performance.
TimeDistributed	Input:(None, 2, 256) Output: (None, 2, 4)	1028	256*4+4 (Input*Output+Output)	The distributed temporal feature representation is mapped to the sample marker space for full connection in the temporal dimension.

4 Multi-layer optimized GRU prediction method based on LOWESS smoothing and Random Forest

The prediction method developed in this paper mainly includes four steps: data smoothing, feature selection, prediction, and outcome analysis.

Step1: **Data smoothing**. Five different smoothing methods of origin software, LOWESS, LOESS , PF, SG, and AAv smoothing, are respectively used to process the daily power generation, so as to compare the prediction results. Sort the root mean square error between the smoothed data and the actual data in descending order.

Step2: **Random Forest feature selection**. Random Forest is introduced for feature selection, Python is used to sort and screen the importance of factors affecting photovoltaic power generation.

Step3: **Model prediction**. The data set is divided into train set and test set, the results of Random Forest feature selection are combined with smoothed data to form a new set of high-quality data set, i.e., the results of Random Forest feature selection are used as the input features of the model. On the basis of GRU model, the RepeatVector and TimeDistributed layers are added for optimization. The optimized GRU model is then employed for prediction.

Step4: **Model evaluation and comparison**. The experimental results are split into vertical comparison and horizontal comparison for analysis.

(1) Vertical comparison: The smoothing method with the least error is first selected. We compare this smoothing method with the prediction results with no smoothing.

(2) Horizontal comparison: Under the GRU model, we conduct the comparison between smoothed and unsmoothed, the comparison between optimized and un-optimized, and the comparison between the accuracies of GRU, RNN, and LSTM prediction models.

The flow chart of prediction steps is shown in Fig.3.

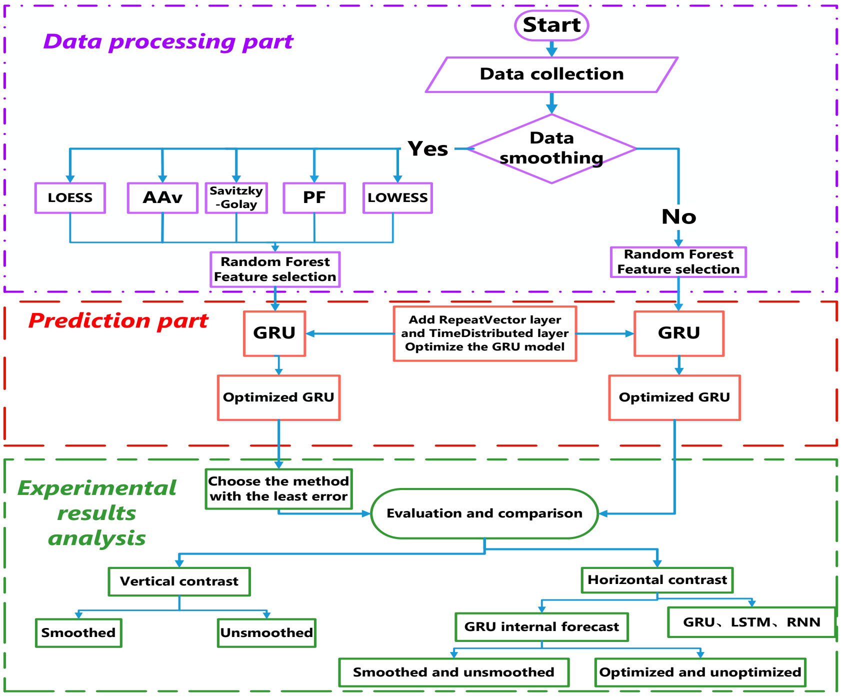

Fig.3. Flow chart of photovoltaic power generation prediction.

5 Case study

5.1 The experimental data

This experiment uses the daily power generation data of the photovoltaic power station of Xuhui District Government in Shanghai, China from January 1, 2015, to December 31, 2016 (train set) and January 2017 (test set) in smart PV website (https://www.lvsedianli.com/perHome.html) [35], with an interval of 24 hours.

The natural factors affecting the daily photovoltaic power generation are the data of Shanghai meteorological station from the national greenhouse system website, and the interval is also 24 hours. 11 natural factors affecting photovoltaic power generation are cumulative precipitation from 20 to 20 o'clock (mm) (hereafter, referred to as cumulative precipitation), average wind speed (m/s), maximum wind speed (m/s), average temperature (℃), daily maximum temperature (℃), daily minimum temperature (℃), sunshine hours (h), daily cumulative radiation (MJ/m^2), average relative humidity (%), minimum relative humidity (%), evaporation (mm).

5.2 Evaluation metrics

Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), NRMSE (Normalized Root Mean Squared Error) and R^2 (R-Squared) are selected as the metrics for evaluating the GRU model. Among these six evaluation metrics, the value of R^2 ranges from 0 to 1; and, the closer it is to 1, the better the fit of the model is. A smaller value of the remaining five evaluation criteria implies a higher prediction accuracy. Where y_i of formula (14-17) represents the real value, $y_i^{'}$ of formula (14-17) refers to the predicted value. In formula (18), $y_{\max} - y_{\min}$ means the full distance of the true value. Var denotes the variance value of formula (19), $n = 31$.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_i^{'}\right| \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(14)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - y_i^{'}\right)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(15)$$

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-y_i^{'}\right)^2} \quad (16)$$

$$MAPE=\frac{1}{n}\sum_{i=1}^{n}\frac{\left|y_i-y_i^{'}\right|}{y_i}\times 100\% \quad (17)$$

$$NRMSE=\frac{RMSE}{y_{\max}-y_{\min}} \quad (18)$$

$$R^2=1-\frac{RMSE}{Var} \quad (19)$$

5.3 Experimental setup

(1) Parameter settings of Random Forest

The accuracy of the Random Forest is mainly determined by the number of decision trees, the maximum depth of decision trees, the setting of random numbers, and the minimum sample number of leaf nodes. The parameter settings of Random Forest are shown in Table 3.

Table 3. Parameter settings of Random Forest.

Algorithm	Parameter meaning	Parameter	The parameter value	Setting basis
Random Forest	the number of decision trees	*n*_estimators	200	Specify the number of classifiers. If the number is too small, it is not fitted, and the training rate is too much, it needs to be compromised.
	the maximum depth of decision trees	max_depth	3	The common value range is 10~100, which can be modified appropriately when there are many sample sizes and characteristic quantities.
	the setting of random numbers	random_state	42	It is used to ensure that the experiment is divided into the same training set and test set every time.
	the minimum sample number of leaf nodes	min_samples_leaf	2	It is related to decision tree pruning, which is generally set to 1, it can be increased under the condition of a large sample size.
	the number of unit layers	num_layers	2	The default is 1 layer, if there are 2 layers, two GRUs are stacked together to form a unit.

(2) Parameter settings of GRU

The accuracy of GRU prediction model mainly depends on the number of neurons, the number of unit layers, time step, hidden layer width, and iteration times. The parameter settings of GRU are shown in Table 4.

Table 4. Parameter setting of GRU.

Model	Parameter meaning	Parameter	The parameter value	Setting basis
GRU	the number of neurons	unit	256	It is a key parameter affecting the accuracy and cannot be increased indefinitely.
	the number of unit layers	num_layers	2	The default is 1 layer, if there are 2 layers, two GRUs are stacked together to form a unit.
	time step	time_step	2	The difference between the two-time points before and after, and this experiment is the prediction of daily power generation.
	hidden layer width	batch_size	256	The number of statements entered into the GRU at one time, there is no fixed value.
	iteration times	epochs	50	It is related to the computing capacity of the computer, and too many iterations are time-consuming and labor-intensive.

5.4 Experiment Results

In our experiments, the results are divided into vertical comparison and horizontal comparison. The vertical comparison refers to the comparison between the introduction of data smoothing and the non-introduction of data smoothing. For the prediction resulting from the involvement of data smoothing, the specific steps including LOWESS, LOESS, PF, SG, and AAv can smooth the actual data, then utilize Random Forest to select the features of smoothed data, and finally combine the smoothed data with the results of feature selection and substitute them into the GRU model for prediction. The non-introduction of data smoothing indicates that the actual data is directly featured through Random Forest and then is substituted in the GRU prediction model.

The horizontal comparison is divided into the internal comparison of GRU and the external comparison between GRU, RNN and LSTM. The internal comparison based on GRU model is the comparison between prediction results from smoothed and unsmoothed data under the same feature selection results. In addition, the internal comparison also includes the compared results of optimized GRU and un-optimized GRU models. Moreover, three prediction models with the same data processing results are compared in terms of prediction accuracy.

5.4.1 Vertical comparison

(1) The comparison with different data smoothing methods

Considering that the daily photovoltaic power generation is subject to solar radiation, temperature, and other factors, which have a large fluctuation ranges, the data smoothing technologies are introduced to reduce the noise and fluctuation range. We use LOWESS, LOESS, PF, SG, and AAv to process the daily power generation data. Then, we use Random Forest for feature selection. Since different data smoothing methods have different results of feature selection and prediction, it is necessary to compare and analyze the final prediction results, which are shown in Table 5. (Aim to find the smoothing method with the best prediction results by adopting four evaluation metrics)

Table 5. Comparison of prediction errors of five data smoothing methods.

Smoothing methods	RMSE	Random Forest feature selection results (in order of importance)	GRU model prediction results			
			RMSE	MSE	MAE	MAPE
Savitzky-Golay	6.271	① daily cumulative radiation ② daily maximum temperature ③ average relative humidity	7.263	52.750	6.589	120.148
LOWESS	**6.472**	**① daily maximum temperature** **② daily cumulative radiation** **③ daily minimum temperature**	**2.352**	**5.530**	**1.851**	**19.715**
LOESS	6.993	① daily maximum temperature ② daily cumulative radiation ③ daily minimum temperature ④ evaporation	5.163	26.658	4.182	Inf
AAv	7.339	① daily cumulative radiation ② daily maximum temperature ③ evaporation	4.486	20.120	3.630	29.321
PF	8.998	① daily cumulative radiation ② daily maximum temperature ③ evaporation	4.910	24.109	3.822	27.343

According to Table 5, we find that when LOWESS smoothing method and three features (daily maximum temperature, daily cumulative radiation, and daily minimum temperature) are selected by Random Forest, GRU prediction model can obtain the lowest experimental error. In fact, RMSE, MSE, MAE and MAPE represent the smoothed prediction error values. We find that MSE, MAE and MAPE are all lowest under LOWESS smoothing method. The RMSE values in the second column of Table 5 represent the error between the smoothed data and actual data. It is obvious that SG smoothing with the lowest RMSE has the largest prediction error. At this time, the prediction error of PF smoothing with highest RMSE is relatively low. Therefore, improving the accuracy of

GRU prediction model but not changing the attributes and authenticity of actual data as much as possible is the key to this experiment. Compared with other smoothing methods, the RMSE of LOWESS smoothing ranks second, which has less error with the actual data and the lowest prediction error, LOWESS smoothing method then reasonably become our experiment selection.

(2) Comparison of unsmoothed and smoothed experimental results

In this section, we discuss the comparison between smoothed and unsmoothed experimental results under three different contexts (Random Forest feature selection, prediction results, and prediction errors) to reflect the importance of data processing.

A. Comparison analysis in Random Forest feature selection

We use Random Forest to rank and filter the feature importance of actual data and smoothed data. The results of feature importance ranking are presented in Fig.4 and Fig.5.

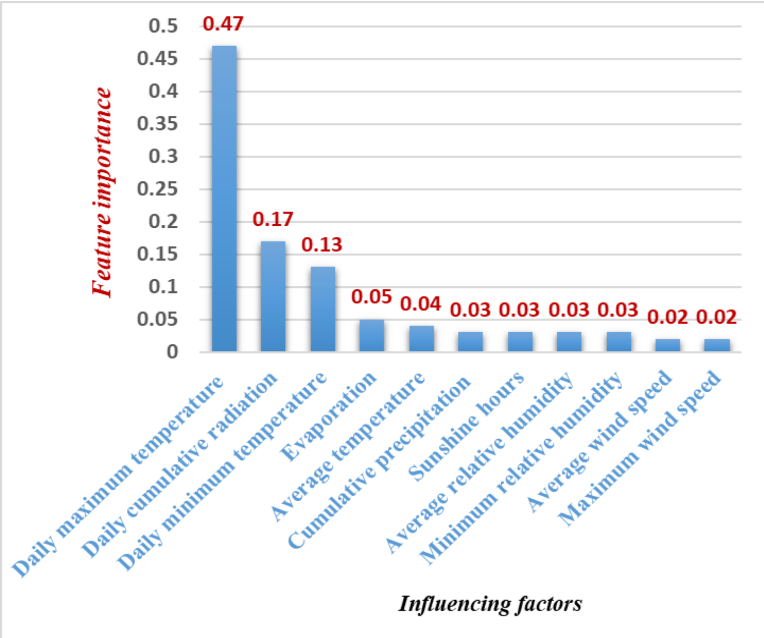

Fig.4. Random Forest feature selection results (Data after LOWESS smoothing).

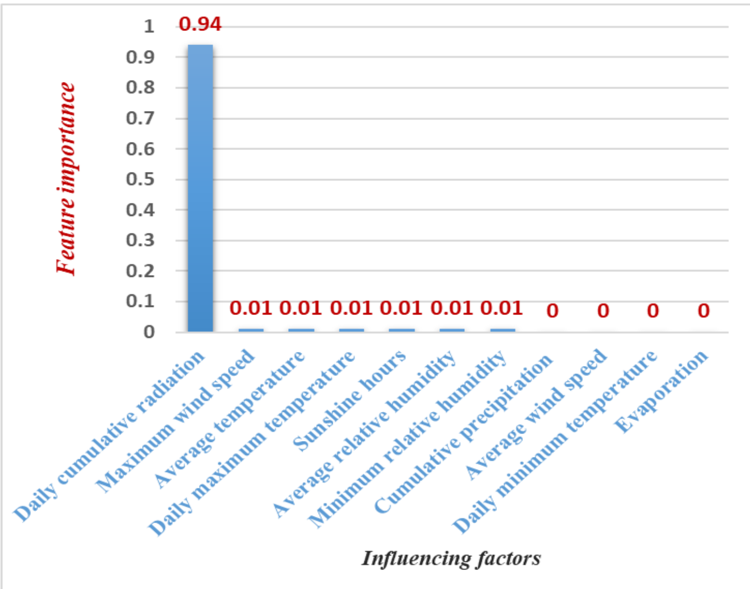

Fig.5. Random Forest feature selection results (Actual data).

As shown in Fig.4, for the LOWESS smoothed data, the results of Random Forest feature selection indicate a hierarchical gradient downward effect with a stepwise nature. According to the order of importance, the filtered features are daily maximum temperature, daily cumulative radiation and daily minimum temperature. As for the feature selection results considering actual data in Fig.5, the seven features filtered are daily cumulative radiation, maximum wind speed, average temperature, daily maximum temperature, sunshine hours, average relative humidity and minimum relative humidity, which filters out more features than smoothed data. The ratio of daily cumulative radiation is very high, and the remaining impact factors are extremely low, even to 0, with no stepwise. Therefore, the feature selection results considering actual data are obviously not as good as that considering smoothed data, which not only proves the effectiveness of Random Forest feature selection but also demonstrates the improvement of data smoothing on Random Forest performance.

B. Comparison analysis between prediction results and actual data

This part is about the prediction results comparison using smoothed data, unsmoothed data and actual data. We learn from Part A that the feature selection results are distinct using the smoothed and unsmoothed data. After the smoothed data with the filtered features are combined to form a high quality data set, this data set is substituted into GRU model for prediction. The same steps are followed for the unsmoothed prediction. We note from Table 5 that the prediction error of LOWESS smoothing is the smallest. Then, we select the prediction results using LOWESS smoothed data for analysis and comparison. Fig.6 represents the comparison results between actual data and GRU prediction model. From the interval of data distribution, the interval of actual data is 0~27.5 kw·h, the interval of unsmoothed prediction results is 5~17.5 kw·h, while the interval of smoothed prediction results ranges from 2.5~27.5 kw·h. Obviously, the smoothed prediction results are more consistent with the interval of actual data. In terms of the goodness of fit, the trend of smoothed prediction results and actual data are basically the same, but the trend of unsmoothed prediction results has almost no correlation with actual data, then the goodness of fit and prediction effect of smoothed prediction results are much better than unsmoothed prediction results. In conclusion, LOWESS smoothing can effectively reduce the fluctuation range of data, Random is applicable to feature selection, which shows that the combination of LOWESS smoothing and Random Forest can improve quality of input data and prediction accuracy of GRU.

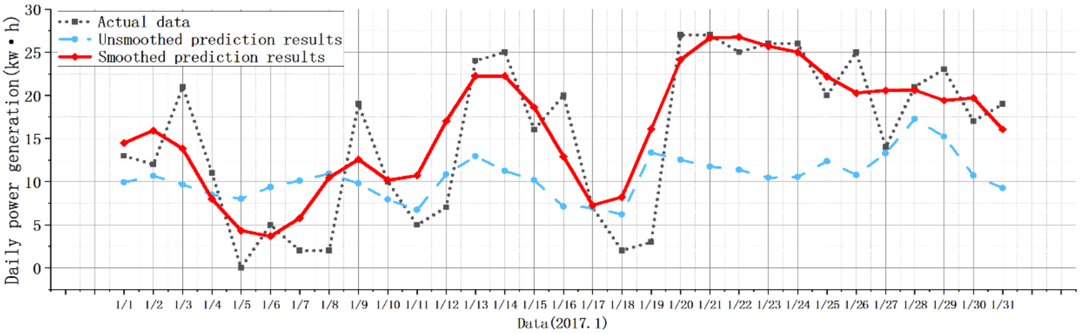

Fig.6. Comparison of actual data with prediction results.

C. Comparison analysis in GRU prediction errors

Different prediction results produce diverse prediction errors. In Table 6, smoothed prediction error is significantly smaller than unsmoothed prediction error, and the R^2 value of smoothed prediction is higher. A higher value of R^2 means a better goodness of fit for prediction model. The R^2 value of smoothed prediction is 0.955, which is close to 1, representing the high fitting ability of prediction method proposed in this paper. The R^2 value of unsmoothed prediction is 0.874, which is lower than that of smoothed prediction. The MAPE of unsmoothed prediction even reached infinity. Moreover, the RMSE of unsmoothed prediction is five times higher than smoothed prediction, and the MSE reaches 20 times. Then, smoothed prediction can result in a lower prediction error and a higher prediction validity.

Table 6. Comparison of unsmoothed and LOWESS smoothed prediction errors.

Evaluation metrics	Unsmoothed	**LOWESS smoothing**
RMSE	10.005	**2.352**
MSE	100.105	**5.530**
MAE	8.435	**1.851**
MAPE	Inf	**19.715**
NRMSE	0.370	**0.102**
R^2	0.874	**0.955**

After comparing the results of above three contexts, we find that data smoothing can effectively reduce the prediction error and improve the model prediction accuracy. Especially, Table 5 indicates that all five smoothing methods used in our experiment can effectively improve the prediction accuracy of the GRU prediction model. However, the unsmoothed prediction performance is poor because of the uncertainty and fluctuation of photovoltaic power generation from the influence of solar radiation and temperature.

5.4.2 Horizontal comparison

(1) Internal comparison of GRU model

A. Comparison analysis of smoothed and un-smoothed prediction errors

This section compares the prediction errors of GRU model using smoothed and unsmoothed data under the same LOWESS smoothed feature selection results. Table 7 shows that the unsmoothed prediction error is higher than that under smoothed prediction, the value of R^2 is significantly lower. The MSE of unsmoothed prediction is much larger than that of smoothed prediction, and its MAPE is also infinite. Although we use the same feature selection results, the unsmoothed prediction result is worse than that under smoothed prediction, which indicates that data smoothing techniques are meaningful for improving prediction accuracy.

Table 7. Comparison of smoothed and unsmoothed GRU internal prediction errors.

Evaluation metrics	**LOWESS smoothing**	Unsmoothed
RMSE	**2.352**	9.112
MSE	**5.530**	83.022
MAE	**1.851**	7.837
MAPE	**19.715**	Inf
NRMSE	**0.102**	0.337
R^2	**0.955**	0.824

B. Comparison analysis of optimized and un-optimized GRU prediction errors

Based on the consistent data processing results, the un-optimized GRU with only one layer is compared with the optimized GRU added RepeatVector layer and TimeDistributed layers. It can be seen from Table 8 that the prediction error of optimized GRU model is evidently less than that of un-optimized GRU model, the value of R^2 is higher than un-optimized GRU, the NRMSE is much lower than un-optimized, the RMSE and MAPE of optimized GRU is several times less than that of un-optimized GRU, so as to show that the optimization of prediction model is also a momentous mean to improve prediction accuracy.

Table 8. Comparison of optimized and un-optimized GRU prediction errors.

Evaluation metrics	**Optimized GRU**	Un-optimized GRU
RMSE	**2.352**	34.901
MSE	**5.530**	5.908
MAE	**1.851**	4.634
MAPE	**19.715**	82.858
NRMSE	**0.102**	0.872
R^2	**0.955**	0.831

(2) External comparison of GRU model

A. Comparison analysis of prediction results

Based on same data processing results, GRU, RNN and LSTM models are used for comparison. We first analyze the prediction results of three prediction models. As can be seen in Fig.7, although the trends of three models are roughly the same as the actual data, the prediction results of GRU are closer to the actual data than those of LSTM and RNN, and have the best goodness of fit with the actual data. Especially, both RNN and LSTM are significantly far from the actual data since January 21, 2017. Therefore, under the consistent data processing results, it is also particularly momentous to choose one appropriate prediction model.

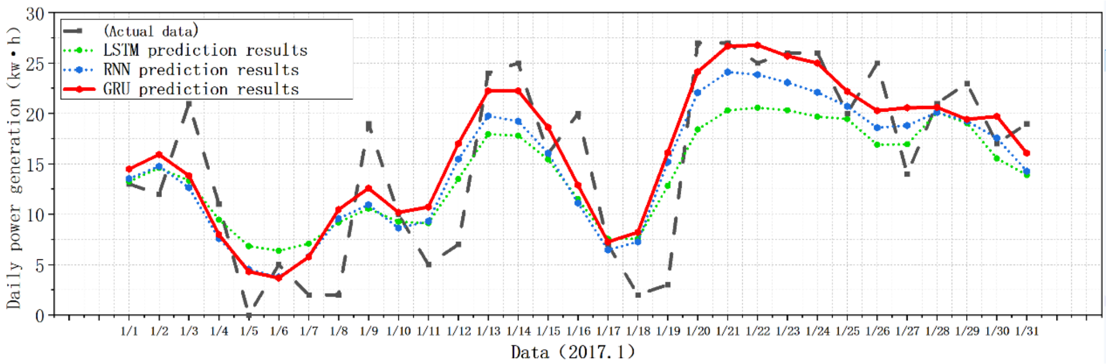

Fig.7. Comparison among prediction results of LSTM, RNN, GRU and actual data.

B. Comparison analysis of prediction errors

Finally, we compare the prediction errors of three models. Table 9 shows that GRU has the smallest prediction error and the highest accuracy. The RMSE, MSE and MAPE of these three models have little difference, but the gap in MAPE is slightly larger. Based on R^2 value, GRU is 0.950, RNN is 0.931, LSTM is 0.860, which prove that GRU has the best goodness of fit, followed by RNN, and LSTM has the worst fit.

Table 9. Comparison of prediction errors of LSTM, RNN, and GRU.

Evaluation metrics	**GRU**	LSTM	RNN
RMSE	**2.352**	3.131	2.602
MSE	**5.530**	9.801	6.771
MAE	**1.851**	2.794	2.241
MAPE	**19.715**	29.056	22.207
NRMSE	**0.102**	0.221	0.128
R^2	**0.950**	0.860	0.931

5.4.3 Impact of uncertainty

To fully analyze the impact of uncertainty on model performance, we consider both data processing and prediction models. In general, the uncertainty in data processing mainly refers to the unavoidable errors in the data due to the limitations of measurement means and machine anomalies. The uncertainty in prediction model mainly includes the model's own structure as well as the optimization algorithm [36]. In addition, the factors affecting the uncertainty of data processing also include natural climatic conditions.

A. Data processing

Fig.5 shows the uncertainty of actual data leads to the lack of ladder nature in the feature selection results. With the data smoothing technology, the results of feature selection are clearly enhanced. Therefore, data smoothing can effectively reduce the negative impact of uncertainty. As seen in Fig.6, the uncertainty of data causes serious deviations between actual data and prediction results, and the prediction accuracy is too low. However, the prediction accuracy is significanty improved after data smoothing and feature selection. Then data processing is meaningful to reduce the adverse effects of uncertainty on the prediction accuracy.

B. Prediction model

For prediction model, Table 9 demonstrates the importance of choosing a suitable prediction model. RNN and LSTM have slightly lower prediction accuracy than GRU due to their own gradient problem and complex parameters. Table 8 shows the model performance of optimized GRU added RepeatVector and TimeDistributed layers is clearly better than that of un-optimized GRU. Thus, choosing the appropriate prediction model and optimization method is also beneficial to reduce the influence of uncertainty.

In summary, data smoothing techniques reduce the fluctuation of data, Random Forest filters out reasonable feature selection results, which solve the uncertainty in data processing to some extent. Furthermore, the GRU model added RepeatVector and TimeDistributed layers improves the prediction accuracy of GRU model. Therefore, the developed forecasting method can effectively reduce the uncertainty impact of photovoltaic power generation on the prediction performance.

6 Conclusion

Aiming to enhance the accuracy and stability of forecasting, this paper considers the impact of data fluctuations on the prediction of photovoltaic power generation, and develops an optimized GRU forecasting method which includes data smoothing technology, feature selection, and optimization of GRU. Firstly, we uses different data smoothing technologies to reduce the fluctuation of actual data and choose the best one with the least prediction error. Secondly, to obtain higher-quality input data, Random Forest method is used to select natural factors affecting photovoltaic power generation. Finally, the RepeatVector layer and TimeDistributed layer are used to optimize the GRU model. Through case studies and experimental results, the conclusions are obtained as follows:

(1) Five different smoothing methods can all improve the prediction performance of GRU model. Among them, the LOWESS smoothing can generate the smallest prediction error.

(2) The Random Forest feature selection can simplify the number of features and optimize the prediction performance of GRU model.

(3) Under the consistent results of data processing, GRU model is more suitable for the prediction of photovoltaic power generation than LSTM and RNN.

(4) The GRU model with RepeatVector layer and Timedistributed layer has a better prediction performance than the un-optimized single-layer GRU.

In summary, compared with original GRU model, the proposed forecasting method in this paper improves data processing and optimizes GRU model. The applications of different data smoothing techniques reduce the fluctuation of daily power generation and improve the quality of input data. Random Forest selects the characteristics of natural factors affecting photovoltaic power generation. The optimized GRU with RepeatVector and TimeDistributed layers not only enriches GRU layer structure, but also effectively improves the prediction accuracy. In terms of practical application value, high-precision generation forecasting method is an effective approach for integrating solar energy resources into power grid. Our results show that the proposed method could be a useful tool to forecast the short-term photovoltaic power generation with an acceptable degree of accuracy. Besides, according to the prediction results, photovoltaic power plants can arrange future power generation, adjust electrovalence, and provide technical support to make timely and reasonable scheduling decision for power grid.

The integration of photovoltaics into the power grid present both opportunities and challenges. The instability of photovoltaic power generation has caused challenges and impacts on the power grid though it can reduce the pollution and loss caused by traditional power generation. Accurate prediction is conducive to the safe and stable and economic operation of photovoltaic plants after grid connection. Compared with the original GRU, the developed prediction method in this paper improves the prediction accuracy through smoothing technology, feature selection and optimized GRU, but there are still some limitations. For example, the loss of prediction accuracy, the reduction of error between smoothed data and actual data, the unity of the feature selection method, the feasibility of interval prediction, and better improvement of GRU prediction performance. These are all problems we intend to solve. In addition, the profound development of photovoltaic power generation needs the support of medium and long-term forecasting. In the future, we collect data suitable for medium and long-term forecasting, study relevant forecasting technologies, and use medium and long-term forecasting to make scientific decisions and plans for the development of photovoltaic power.

Acknowledgements

This work was supported by the National Natural Science Foundation of China [No. 72171126], Ministry of Education Project of Humanities and Social Science [No.20YJA630009], Social Science Planning Project of Shandong Province [No. 20CSDJ15]. The third author (Mingming Leng) was supported by the Faculty Research Grant of Lingnan University under the grant number DB21B1.

References

[1] Lin B, Zhu J. Policy effect of the Clean Air Action on green development in Chinese cities. J Environ Manage 2020;258:110036. https://doi.org/10.1016/j.jenvman.2019.110036.

[2] Jin S, Greaves D. Wave energy in the UK: Status review and future perspectives. Renew Sustain Energy Rev 2021;143:110932. https://doi.org/10.1016/j.rser.2021.110932.

[3] Gandoman FH, Abdel Aleem SHE, Omar N, Ahmadi A, Alenezi FQ. Short-term solar power forecasting considering cloud coverage and ambient temperature variation effects. Renew Energy 2018;123:793–805. https://doi.org/10.1016/j.renene.2018.02.102.

[4] Cheng L, Zang H, Wei Z, Ding T, Xu R, Sun G. Short-term Solar Power Prediction Learning Directly from Satellite Images with Regions of Interest. IEEE Trans Sustain Energy 2022;13:629-639. https://doi.org/10.1109/TSTE.2021.3123476.

[5] Rodríguez F, Genn M, Fontán L, Galarza A. Very short-term temperature forecaster using MLP and N-nearest stations for calculating key control parameters in solar photovoltaic generation. Sustain Energy Technol Assessments 2021;45. https://doi.org/10.1016/j.seta.2021.101085.

[6] Almonacid F, Pérez-Higueras PJ, Fernández EF, Hontoria L. A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator. Energy Convers Manag 2014;85:389–98. https://doi.org/10.1016/j.enconman.2014.05.090.

[7] Khan AT, Khan AR, Li S, Bakhsh S, Mehmood A, Zaib J. Optimally configured Gated Recurrent Unit using Hyperband for the long-term forecasting of photovoltaic plant. Renew Energy Focus 2021;39:49–58. https://doi.org/10.1016/j.ref.2021.07.002.

[8] Yang D. A guideline to solar forecasting research practice: Reproducible, operational, probabilistic or physically-based, ensemble, and skill (ROPES). J Renew Sustain ENERGY 2019;11. https://doi.org/10.1063/1.5087462.

[9] Guijo-Rubio D, Durán-Rosal AM, Gutiérrez PA, Gómez-Orellana AM, Casanova-Mateo C, Sanz-Justo J, et al. Evolutionary artificial neural networks for accurate solar radiation prediction. Energy 2020;210:118374. https://doi.org/10.1016/J.ENERGY.2020.118374.

[10] Dogan A, Birant D. Machine learning and data mining in manufacturing. Expert Syst Appl 2021;166:114060. https://doi.org/10.1016/J.ESWA.2020.114060.

[11] Jeffrey Kuo CF, Su TL, Jhang PR, Huang CY, Chiu CH. Using the Taguchi method and grey relational analysis to optimize the flat-plate collector process with multiple quality characteristics in solar energy collector manufacturing. Energy 2011;36:3554–62. https://doi.org/10.1016/J.ENERGY.2011.03.065.

[12] Adar M, Najih Y, Gouskir M, Chebak A, Mabrouki M, Bennouna A. Three PV plants performance analysis using the principal component analysis method. Energy 2020;207:118315. https://doi.org/10.1016/J.ENERGY.2020.118315.

[13] Jiang Z, Cai J, Moses PS. Smoothing control of solar photovoltaic generation using building thermal loads. Appl Energy 2020;277:115523. https://doi.org/10.1016/J.APENERGY.2020.115523.

[14] Mariani MC, Basu K. Local regression type methods applied to the study of geophysics and high frequency financial data. Phys A Stat Mech Its Appl 2014;410:609–22. https://doi.org/10.1016/J.PHYSA.2014.05.070.

[15] Quan Q, Cai KY. Time-domain analysis of the Savitzky–Golay filters. Digit Signal Process 2012;22:238–45. https://doi.org/10.1016/J.DSP.2011.11.004.

[16] Lin L, Ma S, Ma M. A group neighborhood average clock synchronization protocol for wireless sensor networks. Sensors (Switzerland) 2014;14:14744–64. https://doi.org/10.3390/S140814744.

[17] Duin RPW, Haringa H, Zeelen R. Fast percentile filtering. Pattern Recognit Lett 1986;4:269–72. https://doi.org/10.1016/0167-8655(86)90007-3.

[18] Tronarp F, Särkkä S. Iterative statistical linear regression for Gaussian smoothing in continuous-time non-linear stochastic dynamic systems. Signal Processing 2019;159:1–12. https://doi.org/10.1016/J.SIGPRO.2019.01.013.

[19] Rostami M, Berahmand K, Nasiri E, Forouzande S. Review of swarm intelligence-based feature selection methods. Eng Appl Artif Intell 2021;100:104210. https://doi.org/10.1016/J.ENGAPPAI.2021.104210.

[20] López G, Batlles FJ, Tovar-Pescador J. Selection of input parameters to model direct solar irradiance by using artificial neural networks. Energy 2005;30:1675–84. https://doi.org/10.1016/J.ENERGY.2004.04.035.

[21] Bas E, Yolcu U, Egrioglu E. Picture fuzzy regression functions approach for financial time series based on ridge regression and genetic algorithm. J Comput Appl Math 2020;370:112656. https://doi.org/10.1016/J.CAM.2019.112656.
[22] Rafati A, Joorabian M, Mashhour E, Shaker HR. High dimensional very short-term solar power forecasting based on a data-driven heuristic method. Energy 2021;219:119647. https://doi.org/10.1016/J.ENERGY.2020.119647.
[23] Dutta S, Li Y, Venkataraman A, Costa LM, Jiang T, Plana R, et al. Load and Renewable Energy Forecasting for a Microgrid using Persistence Technique. Energy Procedia 2017;143:617–22. https://doi.org/10.1016/J.EGYPRO.2017.12.736.
[24] Wang M, Peng J, Luo Y, Shen Z, Yang H. Comparison of different simplistic prediction models for forecasting PV power output: Assessment with experimental measurements. Energy 2021;224:120162. https://doi.org/10.1016/J.ENERGY.2021.120162.
[25] Bin Shams M, Haji S, Salman A, Abdali H, Alsaffar A. Time series analysis of Bahrain's first hybrid renewable energy system. Energy 2016;103:1–15. https://doi.org/10.1016/J.ENERGY.2016.02.136.
[26] Jung Y, Jung J, Kim B, Han SU. Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: Case study of South Korea. J Clean Prod 2020;250:119476. https://doi.org/10.1016/J.JCLEPRO.2019.119476.
[27] Ahmed R, Sreeram V, Mishra Y, Arif MD. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. Renew Sustain Energy Rev 2020;124:109792. https://doi.org/10.1016/J.RSER.2020.109792.
[28] Han S, Qiao Y hui, Yan J, Liu Y qian, Li L, Wang Z. Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network. Appl Energy 2019;239:181–91. https://doi.org/10.1016/J.APENERGY.2019.01.193.
[29] Pan C, Tan J, Feng D. Prediction intervals estimation of solar generation based on gated recurrent unit and kernel density estimation. Neurocomputing 2021;453:552–62. https://doi.org/ 10.1016/ j.neucom.2020.10.027.
[30] Abdel-Basset M, Hawash H, Chakrabortty RK, Ryan M. PV-Net: An innovative deep learning approach for efficient forecasting of short-term photovoltaic energy production. J Clean Prod 2021;303:127037. https://doi.org/10.1016/J.JCLEPRO.2021.127037.
[31] Nie Y, Zamzam AS, Brandt A. Resampling and data augmentation for short-term PV output prediction based on an imbalanced sky images dataset using convolutional neural networks. Sol Energy 2021;224:341–54. https://doi.org/10.1016/j.solener.2021.05.095.
[32] Nguyen NQ, Bui LD, Doan B Van, Sanseverino ER, Cara D Di, Nguyen QD. A new method for forecasting energy output of a large-scale solar power plant based on long short-term memory networks a case study in Vietnam. Electr Power Syst Res 2021;199:107427. https://doi.org/10.1016/j.epsr.2021.107427.
[33] Idris A, Rizwan M, Khan A. Churn prediction in telecom using Random Forest and PSO based data balancing in combination with various feature selection strategies. Comput Electr Eng 2012;38:1808–19. https://doi.org/10.1016/J.COMPELECENG.2012.09.001.
[34] Conlin R, Erickson K, Abbate J, Kolemen E. Keras2c: A library for converting Keras neural networks to real-time compatible C. Eng Appl Artif Intell 2021;100:104182. https://doi.org/10.1016/J.ENGAPPAI.2021.104182.
[35] https://www.lvsedianli.com/perHome.html.

[36] Barukčić M, Hederić, Hadžiselimović M, Seme S. A simple stochastic method for modelling the uncertainty of photovoltaic power production based on measured data. Energy 2018;165:246–56. https://doi.org/10.1016/J.ENERGY.2018.09.134.