Improved Markov-chain-based ultra-short-term PV forecasting method for enhancing power system resilience

Xiaoyang Bai¹,² | Liang Liang¹,² | Xueqin Zhu¹,²

¹ State Key Laboratory of Operation and Control of Renewable Energy and Storage Systems, China Electric Power Research Institute, Beijing, China
² School of Mechanical Engineering and Automation, Harbin Institute of Technology (Shenzhen), Shenzhen, Guangdong, China

Correspondence
Liang Liang
1 State Key Laboratory of Operation and Control of Renewable Energy and Storage Systems, China Electric Power Research Institute, Beijing, China
2 School of Mechanical Engineering and Automation, Harbin Institute of Technology (Shenzhen), Shenzhen, Guangdong, 518055, China.
Email: liangl@hit.edu.cn

Abstract
The awareness capability of output power for renewable resources is essential for enhancing the resilience of power systems. Photovoltaic (PV) forecasting technology is an essential technology for increasing the operation efficiency and controllable resources for power systems after extreme natural events. Conventional Markov chain (MC) methods often ignore the time characteristics and the actual distribution of the PV output power sequence when making PV forecasts. This article proposes improved MC methods of equal quantity and clustering-based division methods. The methods can consider the interval distributions of the PV output power time series and select an hour as the time interval. As a sequence, the predicted power at the next moment can be closer to the expectation of the output power distributions. Such a method is combined with a similar day algorithm to calculate the forecast result. Case studies were conducted with one-year operation data from a 25-MW PV station. The results indicate that the proposed methods can effectively improve the accuracy of prediction results compared with traditional methods.

1 | INTRODUCTION
In recent years, extreme natural events have brought great technical challenges to power system operations. Enhancing the resilience of power systems can increase the ability of power systems to adapt to these challenges. The awareness and number of controllable resources are two key points for enhancing the resilience of power systems.

Renewable energy resources provide an increasing amount of energy for power systems and play an important role in power system operations. The cost of developing and utilizing renewable energy resources has been greatly reduced, the scale is continuously increasing, and the economic efficiency has been significantly improved. With the development of power electronics and energy storage technologies, renewable energy resources can be regarded as a partly controllable resource. In this way, renewable energy resources, including solar and wind energy, can be dispatched to facilitate the restoration process after extreme natural events so that the resilience of the system can be enhanced.

However, there are still technical obstacles to the large-scale integration of renewable energy power generation. The output of renewable energy has strong randomness and fluctuations. As far as photovoltaics (PVs) are concerned, time, weather, and clouds cause fluctuations in output power. In addition to the following load changes, conventional power supplies must consider renewable energy fluctuations. With the gradual expansion of the scale of renewable energy use, high requirements are placed on the regulated power supply, and a large number of flexible peak-shaving energy storage and pumped storage power stations are required if the power of the new energy cannot be accurately predicted. Hence, developing more effective PV power generation forecasting technology can increase the awareness capability of the power system and significantly...
improve the flexibility and resilience of power systems. Moreover, the forecasting technology of PV systems can build a stronger power grid that adapts to the widespread integration of renewable energy resources and promote the implementation of national energy transition policies [1].

The proportion of renewable energy in the power grid has increased due to the innovation of development technology and the reduction of cost. For a grid with large-scale access to renewable energy, extreme weather can have a significant impact on the grid’s power generation capacity. The basic four stages of resilient power system coping with extreme events are as follows: preparation in advance, resistance and absorption, response and adaptation, and rapid recovery.

The power network simulation model is built by predicting the power generation output and grid load. In [2], the simulation results are used to analyse the power failure probability and its impact on the energy market, and a risk-based analysis framework is built to quantify the vulnerability of power grid operation. And the load recovery after the risk is accelerated through the topology control mitigation solutions.

The appropriate reserve capacity for grid operation is determined through forecasting to protect against the risk of high-impact and low-probability events. The innovation of power generation and load forecasting methods plays an extremely important role in improving the resilience of power grid under the condition of high renewable energy source (RES) penetration [3].

Currently, a simple way to implement a PV forecast system in practical engineering projects is to build a simple neural network model by collecting data such as numerical weather forecast results, real-time weather station data, real-time output power data, and inverter unit status. The average system accuracy for such a typical forecast system is approximately 85% [4]. With the increase in RES permeability in the power grid, this accuracy will not meet the actual demand. To improve the prediction accuracy, many related studies have been carried out for the prediction of PV power generation.

In a previous study [5], building a reference evaluation of solar transmittance physical model that does not consider the carbon dioxide transmittance was proposed to predict the solar irradiance data of four stations in India. Solar radiation models were proposed, improved, and applied based on empirical formulas in various places [6–10]. They have good engineering and experimental applicability and are of great significance to regions lacking solar radiation observations. However, the above methods of establishing physical models often put high demands on the quantity and quality of actual parameters.

In recent years, machine-learning-related methods have emerged and have been widely used in the field of PV prediction. Research includes artificial neural networks (ANNs), k-nearest neighbours (KNN), deep neural networks, support vector machines, and the combination of these methods [11–19]. These methods avoid the process of building complex physical models, and the prediction has high accuracy in the case of abundant data.

However, in reality, many areas do not have a large amount of accurate weather information to support the training of the above methods. In the case used in Section 4, the time resolution of the meteorological observation data in this area is only 1 h, and the time resolution of the PV output data can reach 5 min. The prediction effect of neural network construction with meteorological characteristics is not ideal. In this case, the time-series method has relative advantages. Because its model is simple and easy to construct, it is necessary to study this type of method.

In previous work [20], multiple state transition matrices were generated according to seasonal, weather, and daily characteristics, and the fluctuation characteristics and the amount of fluctuation were considered to establish a Markov chain model. In other research [21], the Grey–Markov (GM) chain was adopted, using two GM (1,1) models to predict monotonic data, and a Markov chain model was established for the relative error of the prediction. In other work [22], KNN and AdaBoost were used to improve the clustering algorithm to divide samples, and each weather submodel established a Markov chain model with the attenuation coefficient as the input and solar irradiance as the output; then, a PV engineering formula was used to calculate the output power value.

The difference between the above several Markov methods lies in the definition of the model input and output, which can be applied to different problems or combined with other algorithms. To improve the prediction effect, the Markov chain method itself can be improved in terms of interval division.

However, the power of PV power generation has a strong correlation with weather conditions. In other words, the power curves of PV power generation are quite similar under similar weather conditions. Therefore, the sample clustering method is important for PV forecasting. The method of self-organising mapping has been used [23, 24] to cluster the weather conditions into different types (cloudy, sunny, rainy etc.) according to the solar radiation, relative humidity, temperature, and other meteorological conditions. In other work [25–27], the k-means clustering algorithm was used to divide the entire sample based on the meteorological parameters calculated by the correlation. In addition to the commonly used parameters of temperature and irradiance, in another study [28], solar time and geomagnetic declination were considered, and the output of similar dates was mixed with simple mathematical processing and model mixing as the prediction results, which also yielded good results. Other researchers simply distinguished the weather according to the average value of irradiance for all data, and then they established an ANN model [29].

Based on the combination of the similar day algorithm and the traditional Markov chain, two major categories for the division of intervals are proposed—overall sample division and hourly sub-model division—each of which can adopt equal value, equal data volume, and clustering methods. An actual case of a PV power station in Shandong Province illustrates that the improved method reasonably and effectively improves the accuracy of the forecast.

The major contributions of this paper are listed as follows.

1. A data processing model of similar daily algorithm is established. The traditional classification method of weather is to divide the historical weather into three or four types (sunny,
cloudy, rainy etc.) by clustering method. The method proposed in this work can reflect the similarity of weather characteristics and changes in the time period of the day.

2. Several improved methods based on the traditional Markov method are proposed. Several clustering methods are used in the step of state interval division to improve the effective utilisation of data and display the data distribution characteristics.

3. On the basis of the above method, several small sub-models are built in order to accurately reflect the power distribution and variation law in different time periods. The prediction results of the case show that: a) the advantages of the improvement are proved by comparing with the conventional Markov method; and b) compared with long short-term prediction (LSTM), one of the most commonly methods used in time-series prediction issue, the improved method is effective under the condition of limited characteristic data and insufficient historical samples.

The rest of this article is arranged as follows. Section 2 introduces the basic principles of the similar day algorithms and improved Markov chain. On this basis, the complete forecasting process is introduced in Section 3. In Section 4, actual cases and results are used to verify the effectiveness of the forecast. Finally, conclusions are presented in Section 5.

## 2 | BASIC PRINCIPLES

### 2.1 | Similar day algorithm

The similar day algorithm clusters the samples, which can effectively improve the prediction accuracy for specific weather conditions. The correlation analysis between tags and features should be completed first. Some features that are unrelated to output power or low correlation should be screened out to reduce the dimensionality of system features and system complexity. Clustering of samples can help reduce the occurrence of overfitting and improve the calculation speed of the prediction system. The Pearson correlation coefficient is a common method used to measure the linear correlation degree of the change trend of two variable series. The principle of this method is simple and easy to calculate. The main steps are as follows.

By analysing the data sequence, the target variable $X_0$ to be predicted is obtained, and all the feature value sequences $X_i$ that may affect the prediction result are determined. The covariance and the standard deviations of the two sequences are found, and the Pearson correlation coefficient is calculated as

$$
\rho_{X_0, X_i} = \frac{\text{cov}(X_0, X_i)}{\sigma_{X_0} \sigma_{X_i}} = \frac{E(X_0 X_i) - E(X_0)E(X_i)}{\sqrt{E(X_0^2) - E^2(X_0)}} \sqrt{E(X_i^2) - E^2(X_i)} \tag{1}
$$

According to the calculated absolute value of the similarity of the feature, the corresponding relationship between the

| TABLE 1 Pearson coefficient and correlation strength |
|-----------------|-----------------|
| Pearson coefficient (absolute value) | 0.8–1.0 | 0.6–0.8 | 0.4–0.6 | 0.2–0.4 | 0.0–0.2 |
| Correlation strength | High | Strong | Medium | Weak | Very weak or uncorrelated |

Pearson coefficient and the correlation is combined, as shown in Table 1. The retained feature is selected, and the feature is used to determine the similarity date. After all sample data have been normalised, the standardised Euclidean distance between the forecast day and the sample day at the corresponding time point of each day is calculated. The sample set is given as $X = \{X_1, X_2, ..., X_n\}$, and the forecast day feature is $X_0$. Then,

$$
d_i = \sqrt{(X_i - X_0)^2 / S \cdot \omega}, \quad i = 1, 2, ..., n \tag{2}
$$

where $S$ is the standard deviation of the sample set $X$, and $\omega$ is the weight vector of each feature. The final similarity is ranked from small to large, which is the most similar weather condition to the day to be tested.

### 2.2 | Markov chain

The Markov model was proposed by the Russian mathematician Markov in 1906. It is an important method for dealing with discrete random processes. When the state of the random process of research and analysis in the future moment is only related to the current value but not correlated to the previous moment, this property is called "no aftereffect". After long-term research and experimentation, the Markov model has a good effect on short-term prediction. It has been widely used in communication, meteorology, geography, economics, and many other fields.

Assuming a random process,

$$
X = \{X_i, \ldots, X_n\}, \quad t = 0, 1, 2, \ldots \tag{3}
$$

The set of possible values of the $i$ random variables is called the "state space", $S = \{S_1, S_2, \ldots, S_n\}$, where $n$ is the total number of states that divide the state space, and the Markov process satisfies the following formula:

$$
P(X_t = S_j | X_{t-1} = S_i) = P(X_{t-1} = S_i | X_{t-2} = S_j) = p_{ij} (t - 1). \tag{4}
$$

This indicates the probability of the system transitioning from state $i$ to state $j$ at time $t - 1$. Moreover,

$$
0 \leq p_{ij} \leq 1, \sum_{j=1}^{n} p_{ij} = 1, \quad i = 1, 2, \ldots, n. \tag{5}
$$
When \( P_{ij} \) irrespective of the time \( t \), the sequence is called a “homogeneous Markov chain”. From the state transition probability, the state transition matrix is obtained as

\[
P = \begin{bmatrix}
p_{11} & p_{12} & \cdots & p_{1n} 
p_{21} & p_{22} & \cdots & p_{2n} 
\vdots & \vdots & \ddots & \vdots 
p_{n1} & p_{n2} & \cdots & p_{nn}
\end{bmatrix},
\]

(6)

It is supposed that the probability that the system is in state \( S_i \) at time \( t \) is

\[
\pi_i(t) = p[X_t = S_i], \quad i = 1, 2, \ldots, n
\]

(7)

and the probability should meet

\[
\sum_{i=1}^{n} \pi_i(t) = 1, \quad 0 \leq \pi_i(t) \leq 1.
\]

(8)

The probability when \( t=0 \) is called the initial probability, and the corresponding initial probability vector is

\[
\pi_i(0) = p[X_0 = S_i], \quad i = 1, 2, \ldots, n.
\]

(9)

The probability vector at time \( t \) is

\[
\Pi(t) = \Pi(0) \ast P^t, \quad t = 1, 2, \ldots
\]

(10)

### 2.3 Improved Markov chain

In Markov chain modelling methods, the state interval division process mainly adopts the traditional equivalent method [30–32]. This state interval division method depends on experience and experimentation.

This classification of samples may result in a small number of data in some intervals or a lack of statistics for the process of power rise rapidly, which may not reflect the characteristics of time-series data.

Therefore, the clustering method is proposed for the division of Markov state space, and it is proved effective through case studies in Section 4.

In the following, three methods of dividing state intervals are introduced.

#### 2.3.1 Equivalent division

Equivalent division is the most common and basic division method. According to the maximum and minimum values of historical data \( x_{\text{max}} \) and \( x_{\text{min}} \), the interval width is

\[
d = \frac{1}{n} (x_{\text{max}} - x_{\text{min}} + \delta)
\]

(11)

where \( \delta \) is a small reserved amount for the convenience of statistics and calculations. For a PV power plant with a rated capacity of \( p_N \), the above formula is written as

\[
d = \frac{p_N}{n}.
\]

(12)

Then, the \( n \) state intervals are divided into

\[
[0, d], [d, 2d], \ldots, [(n-1)d, nd].
\]

(13)

#### 2.3.2 Equal quantity division

In order to accurately reflect the power distribution and variation law, it is necessary to ensure that there is sufficient data in each divided interval. Otherwise, accidental factors and measurement errors may increase the impact on the results.

At the same time, the results cannot be calculated if the matrix is sparse. Therefore, all historical data are sorted by power value. The samples after sorting are as follows:

\[
X' = \{X'_1, \ldots, X'_N\}, \quad N = 0, 1, 2, \ldots
\]

(14)

Then, the \( n \) state intervals are divided into

\[
[0, \frac{X'_0}{N/d}], [\frac{X'_0}{N/d}, \frac{X'_2}{N/d}], \ldots, [\frac{X'_{(n-1)d}}{N/d}, \frac{X'_{n}}{N/d}].
\]

(15)

#### 2.3.3 One-dimensional k-means clustering division

Another division method is based on the k-means one-dimensional clustering method, which finds the optimal cluster centre for the entire sample set.

Step 1: The number of cluster centres \( n \), which is the number of divided state intervals, is set.

Step 2: Initialisation: the initial cluster centres \( C_1, C_2, \ldots, C_n \) are assigned randomly.

Step 3: The Euclidean distance between each point in the sample and each cluster centre is calculated. The smallest distance among them is taken, and it is divided into the corresponding cluster centre group. Thus, the cluster centres corresponding to all samples are grouped into \( G_1, \ldots, G_n \). The amount of sample data in each group is \( N_i, i = 1, 2, \ldots, n \). The sum of all minimum distances is defined as the loss function.

Step 4: The cluster centres are recalculated as

\[
G_i = \frac{1}{N_i} \left( \sum_{X_j \in G_i} X_j \right), \quad i = 1, 2, \ldots, n.
\]

(16)

Step 5: The maximum number of iterations and the convergence threshold of the objective loss function are set, and the above process is repeated until the maximum number of iterations is reached or the difference between the two iterations of the objective loss function is less than the given value—that is, the process
is stopped when the loss function tends to converge. The final cluster centre is recorded, and the grouping of the state space is completed.

By redefining the state space of the Markov chain every hour, the predicted process can further reflect the time characteristics of PV power. In this way, at different time points, the PV power changes are clearer and more specific, reducing the existence of state intervals with sparse data.

The above three methods can be divided according to the overall dataset and divided according to the hourly dataset. “Divided according to the hourly sample” means that several sub-models are used to predict; each is responsible for outputting the power prediction of particular moment. For example, the $M_8$ means the status transition matrix used for predicting the PV power transition between 8:00 and 9:00 a.m. The matrix is obtained by statistics on all the datasets which time label is between 8:00 and 9:00 a.m.

In this article, they are referred to as follows: overall equivalent division (OEV), overall equal quantity division (OEQ), overall clustering division (OC), hourly equivalent division (HEV), hourly equal quantity division (HEQ), and hourly clustering division (HC). Among them, the OEV method is the conventional Markov prediction method. This paper mainly reflects the improved advantages by comparing the remaining methods with the conventional methods.

### 3 MAIN PROCESS

The conventional method uses historical samples to build a single Markov model for prediction. To reflect the characteristics of PV power under different time and weather characteristics, for improvement on this basis, each hour is modelled separately. For all samples of the same hour in the historical data of similar days, the state space is divided according to the proposed methods, and the state transition matrix is calculated at each time. According to the weather conditions on the day to be predicted and the actual output at that moment, the state probability vector at the next moment is predicted, and the predicted power value is calculated.

The specific process, shown in Figure 1, is as follows.

**Step 1:** According to the weather forecast data of the day to be tested and the principle of the similar day algorithm, the date serial number of the historical sample closest to the day to be tested is found. According to all the data of these dates, a set of statistical power sequences is established by hour.

**Step 2:** According to the power history data collection, several methods proposed in the improved Markov chain are used to divide the state intervals and determine the expected value of each interval.

**Step 3:** According to the divided interval, the state transition frequency matrix and the state transition probability matrix are calculated.

**Step 4:** This step is the Markov property test.

Passing the Markov property test is the premise of using the Markov chain to solve the prediction problem, which can be completed by using the $\chi^2$ test in the following equations:

\[
\hat{P}_j = \left( \sum_{i=1}^{n} m_{ij} \right) \left/ \left( \sum_{i=1}^{n} \sum_{j=1}^{n} m_{ij} \right) \right.
\]

\[
\hat{P}_{ij} = \left( m_{ij} \right) \left/ \left( \sum_{i=1}^{n} m_{ij} \right) \right.
\]

\[
\chi^2 = 2 \sum_{j=1}^{n} \sum_{i=1}^{n} m_{ij} \ln \left( \hat{P}_{ij} / \hat{P}_j \right)
\]
where $m_{ij}$ is the frequency corresponding to the transition from state $i$ to state $j$ in the time-series history sample, and $P_{ij}$ is the corresponding frequency. If the statistic obeys the $\chi^2$ distribution with $(q - 1)^2$ degrees of freedom, the confidence level is $\alpha$, which satisfies

$$\chi^2 \geq \chi^2_{\alpha}(q - 1)^2. \quad (20)$$

Then, the sequence satisfies the Markov property test and can be processed by the Markov chain.

**Step 5:** The actual PV output at each moment is divided into the corresponding state space $i$, and then, the state probability vector at that moment meets

$$\begin{cases} P_i(t) = P[X_t = s_i] = 1 \\ P_i(t) = P[X_t = s_j] = 0, j \neq i. \end{cases} \quad (21)$$

It is multiplied by the state transition matrix to find the state probability vector at the next moment. According to the expected value of each interval, the final prediction value of the next moment is obtained.

Several indicators are used to evaluate the forecast results:

1. Coefficient of determination $R^2$
   $$R^2 = \frac{\sum (P - \bar{P})^2}{\sum (P - \bar{P})^2} \quad (22)$$

2. Root-mean-square error (RMSE)
   $$\text{RMSE} = \sqrt{\frac{1}{N} \sum (P - \bar{P})^2} \quad (23)$$

3. Mean absolute percentage error (MAPE)
   $$\text{MAPE} = \frac{1}{N} \sum \left| \frac{P - \bar{P}}{C_i} \right| \quad (24)$$

where $\bar{P}$ is the average predicted power of all samples, $P$ is the actual power, $\bar{P}$ is the predicted power, and $C_i$ is the total capacity.

## 4 CASE STUDIES

The case studies are based on a dataset of a 25-MW PV station along the coast of Shandong collected in 2019. The time resolution of the selected dataset was 5 min, and the data were combined with weather data [33] downloaded from the weather website for the entire year with a resolution of 1 h. In this work, the time resolution is 5 min and time ranges from 8 a.m. to 5 p.m., 9 hours per day. The data of 11 months were selected as historical samples to calculate the transition matrix for Markov chain methods. The 28 days in December when the power data could be obtained were used for verification. Time period of prediction: $9 \times 28 = 252$ h. Various weather conditions of sunny, cloudy, and rainy were also imported and tested by the proposed methods to verify the effectiveness of the methods.

### 4.1 Data processing

The first step of data processing is to delete unnecessary data in the historical dataset. In practical projects, the PV station has no output after sunset and before the sun rises. Such zero-output power can be deleted from the historical dataset to save space and reduce processing time for the proposed forecasting methods.

### 4.2 Similar day algorithm

The similar day algorithm clusters the samples, which can effectively improve the prediction accuracy for specific weather conditions. Distinguishing weather by characteristics generally divides the weather into sunny, cloudy, and rainy days, so the characteristics of cloud cover, solar radiation, weather, and rainfall conditions have a great impact on the similarity of weather. The correlation index between the 24 features of the dataset and the output power value Pearson coefficient is obtained, as shown in Table 2.

The calculated results from Table 2 and the scatter diagram drawn by the output power and short-wave radiation are shown in Figure 2. Consistent with the calculated results, they are important factors to forecast the PV power that must be considered when analysing meteorological conditions. The same method was used to determine several other features with higher correlation. Finally, relative humidity, total cloud coverage, sunshine duration, and solar shortwave radiance were selected as the characteristics of similar day division.
### 4.3 Markov chain calculation

The following describes the overall forecasting calculation process, taking December 1 as an example, with the OEV method.

1. **Obtain the similar days in the data set**: In the case where the number of similar days is selected as 15, the calculated similar days number in order of similarity are 64, 74, 82, 34, 30, 50, 139, 42, 26, 149, 208, 168, 167, 1, and 61.

2. **State interval division**: The division of five state intervals is taken as an example. According to the historical PV output, the highest PV output is less than 25 MW, and the PV power is divided into 0–5, 5–10, 10–15, 15–20, and 20–25 MW, which are recorded as State 1 to State 5. The expected value of power prediction in each interval is 2.5, 7.5, 12.5, 17.5, and 22.5 MW, respectively.

3. **Calculate the state transition matrix**: Taking 9:00 a.m. on December 1 as an example, the statistics of the transition frequency matrix and transition probability matrix of the sample day are as follows:

   \[
   P_{s9(5)} = \begin{bmatrix}
   48 & 8 & 1 & 0 & 0 \\
   7 & 49 & 9 & 0 & 0 \\
   0 & 8 & 28 & 3 & 0 \\
   0 & 0 & 1 & 3 & 0 \\
   0 & 0 & 0 & 0 & 0
   \end{bmatrix}
   \]

   \[
   P_{c9(5)} = \begin{bmatrix}
   0.842 & 0.140 & 0.017 & 0 & 0 \\
   0.108 & 0.754 & 0.138 & 0 & 0 \\
   0 & 0.718 & 0.077 & 0 & 0 \\
   0 & 0.25 & 0.75 & 0 & 0 \\
   0 & 0 & 0 & 0 & 0
   \end{bmatrix}.
   \]

   After normalising all sample data, the distance between the forecast day and the sample day at the corresponding time point of each day is calculated. Arrange the dates in order of weather similarity from smallest to largest. The calculated Pearson coefficient value is taken as the weight of similarity calculation \( \omega = [0.502, 0.482, 0.553, 0.772] \).

### Table 2: Pearson coefficients of all features

| Feature                          | \( \rho \) | Feature                          | \( \rho \) | Feature                          | \( \rho \) |
|----------------------------------|------------|----------------------------------|------------|----------------------------------|------------|
| Temperature (°C)                 | 0.288      | Cloud cover high (%)             | -0.31      | Evapotranspiration (mm)          | 0.363      |
| Relative humidity (%)            | -0.502     | Cloud cover medium (%)           | -0.438     | CAPE (J/kg)                      | -0.037     |
| Mean sea level pressure (hPa)    | -0.074     | Cloud cover low (%)              | -0.281     | PBL height (m)                   | 0.414      |
| Precipitation total (mm)         | -0.132     | Sunshine duration (min)          | 0.553      | Wind gust (km/h)                 | 0.176      |
| Precipitation runoff (mm)        | -0.08      | Direct shortwave radiation (W/m) | 0.772      | Wind speed (km/h)                | 0.08       |
| Cloud cover total (%)            | -0.482     | Diffuse shortwave radiation (W/m)| 0.25       | Wind direction                   | 0.089      |
| Geopotential height (Gpm)        | -0.063     | Date distance                    | -0.082     | Soil temperature (°C)            | 0.243      |
| Soil moisture (fraction)         | -0.182     | Vapour pressure deficit (hPa)    | 0.427      | Total column cloud water (mm)    | -0.288     |

### Table 3: Markov property test results

| Methods  | States | Similar days |
|----------|--------|--------------|
|          |        | 2 | 5 | 10 | 15 | 20 | 30 |
| Times of \( \chi^2 \) test | |  | | | | | |
| not past | |  | | | | | |
| OEV      | 5      | 152 31 | 10 | 6  | 3  | 0  |     |
|          | 12     | 252 207| 19 | 11 | 3  | 0  |     |
| OEQ      | 5      | 127 17 | 7  | 3  | 1  | 0  |     |
|          | 12     | 252 197| 12 | 1  | 0  | 0  |     |
| OC       | 5      | 125 17 | 7  | 3  | 1  | 0  |     |
|          | 12     | 252 204| 19 | 11 | 3  | 0  |     |
| HEV      | 5      | 10 2   | 0  | 0  | 0  | 0  |     |
|          | 12     | 252 76 | 0  | 0  | 0  | 0  |     |
| HEQ      | 5      | 10 0   | 0  | 0  | 0  | 0  |     |
|          | 12     | 252 23 | 0  | 0  | 0  | 0  |     |
| HC       | 5      | 11 0   | 0  | 0  | 0  | 0  |     |
|          | 12     | 252 33 | 0  | 0  | 0  | 0  |     |

4. **Markov property test**: By calculating the \( \chi^2 \) distribution statistics, the results of the Markov property test of all state sequences constructed for a total of 29 days are shown in Table 3.

For five state intervals, the chi-square distribution is obeyed with degrees of freedom \((5 - 1)^2 = 16\). At a confidence level \( \alpha = 0.05 \), according to the chi-square distribution table, we have \( \chi^2_{0.05(16)} = 26.296 \).

For 12 state intervals, the chi-square distribution is obeyed with degrees of freedom \((12 - 1)^2 = 121\). At a confidence level \( \alpha = 0.05 \), according to the chi-square distribution table, we have \( \chi^2_{0.05(121)} = 147.674 \).

5. **Result analysis of Markov property test**: According to Table 3, when there are a smaller number of similar days, the time series lacks Markov property, and it is not suitable to use the Markov chain method. In the overall division method (OEV, OEQ, and OC), the power remains at a low level during the earliest and latest time periods of the day, which leads to almost all samples belonging to the state 51, so the calculated value of \( \chi^2 \) is lower than the standard. The forecast for this period cannot be processed in the Markov chain and can
be processed using other time-series methods. For example, the average value of the sample data is used as the predicted value here.

The hourly division method can effectively improve the Markov property of the time series and reduce the number of times that the $\chi^2$ distribution is not satisfied. This also shows that this improved division method is more reasonable.

After the Markov property test is passed, the PV power at the target time can be predicted. The specific steps are shown in Figure 1.

4.4 Forecast result analysis

In the case of selecting similar days as 2, 5, 10, 15, 20, and 30 days, the predicted power value is compared with the actual value. Figure 3 shows that, for clustering weather conditions, the number of days with similar weather conditions cannot be too small. If it is, the sample lacks universality and is greatly affected by special circumstances.

When the number of days is greater than 15, the accuracy of the forecast is not significantly improved, but the time consumption increases proportionally. Therefore, after comprehensive consideration, 15 similar days are selected as the forecast condition.

Figure 4 shows the influence of the number of state intervals on the prediction results. Under the condition of ensuring the Markov properties, the more detailed the division of the historical data samples, the higher the prediction accuracy often is. For the clustering method shown in the figure, the number of states changes very little at more than 10, and the increase in the number of state spaces also means that the predicted time consumption increases proportionally. Therefore, in this study, the number of states of 5 and 12 were taken as examples for prediction.

Some of the predicted results are shown in Figures 4 and 5. Figure 5 shows a magnified view of the time period in Figure 4 for a clearer comparison. When the results of the 5-state OEV are compared with the 12-state OEV, and the 12-state OEV with the 12-state HC, it is obvious that increasing the number of divided states and using the clustering method divided by hours...
can effectively make the predicted curve closer to the true value. The comparison can be made more accurately by calculating the overall evaluation index of the forecast result.

Combining the results in Figures 3–6 and Table 4 reveals that the following:

1. The prediction effect of the improved Markov was obviously better than that of the conventional Markov. A more detailed division of the state interval helps to improve the accuracy of the prediction. In the same method, the error of the prediction results of 12 state intervals is only 60% that of five state intervals. Comparing the three overall interval division methods and their corresponding hourly division methods shows that, for the same interval division, compared with the overall division by hour, the error is only 70–80% of the

| Table 4 Comprehensive results of predictions by several state-space division methods |
|-------------------------------|----------|----------|----------|
| **Method** | **States** | **$R^2$** | **RMSE (MW)** | **MAPE (%)** |
| OEV | 5 | 0.947 | 1.849 | 5.264 |
| | 12 | 0.942 | 1.537 | 3.429 |
| LSTM | 0.931 | 1.719 | 4.966 |
| OEV | 5 | 0.929 | 1.702 | 5.048 |
| | 12 | 0.940 | 1.562 | 3.420 |
| OC | 0.938 | 1.599 | 4.845 |
| | 12 | 0.942 | 1.540 | 3.359 |
| HEV | 5 | 0.945 | 1.504 | 4.224 |
| | 12 | 0.948 | 1.462 | 2.987 |
| HEQ | 0.953 | 1.383 | 3.765 |
| | 12 | 0.969 | 1.128 | 2.751 |
| HC | 0.956 | 1.331 | 3.631 |
| | 12 | 0.969 | 1.129 | 2.776 |
latter, and the MAPE can be reduced by approximately 1%. The improved method can better distinguish the data at this moment, obtain the Markov transition matrix more reasonably and effectively, and greatly improve the accuracy of the prediction results.

2. LSTM that has been developed on the basis of recurrent neural networks is a common method for dealing with time-serial problems. It determines the oblivion, retention, and output of historical information through the forget gate, the input gate, and the output gate and has achieved great success in many areas. Therefore, a four-layer LSTM network model with 20 neurons in each layer is built for training. To avoid overfitting, the dropout parameter is set as 0.2. The time span of data input is 30 min. The time step to forecast is 5 min. The prediction error is shown in Table 4. The LSTM method may be more suitable for long-term prediction than one-step ultra-short prediction. In addition, the temporal resolution of meteorological data is not enough to build multi-feature series model, which limits the improvement of prediction effect. Therefore, the improved Markov proposed in this work may be more suitable in limited data condition.

5 | CONCLUSION

A method based on the traditional Markov-chain-based ultra-short-term PV power prediction method was proposed to increase the prediction accuracy. A non-uniform division method was proposed to determine the state intervals and was combined with a similar day algorithm for the proposed improved Markov chain method. The equal data volume method and one-dimensional k-means clustering division by hour method are applied to determine the state intervals. Compared with the traditional physical modelling method with the data of practical projects, the proposed method is simpler and easier to implement. Moreover, the proposed method is more robust to weather factors.

As verified by the same dataset, processing the data of each hour separately is better than processing all the data. The HEQ and HC methods can improve the accuracy of prediction while ensuring operational efficiency.

Under normal weather conditions, this improvement in accuracy is conducive to the advance planning of the power system and reduces the requirements of the system for climbing ability. In extreme natural conditions, it helps to anticipate risks in advance, prepare backup capacity, and recover quickly after extreme natural events. They may be suitable for application in practical projects to increase the awareness ability and resilience of power systems.

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 52077045, in part by the Natural Scientific Research Innovation Foundation in Harbin Institute of Technology under Grant HIT.NSRIF:2020070, in part by Open Fund of State Key Laboratory of Operation and Control of Renewable Energy and Storage Systems, China Electric Power Research Institute, under Grant DG80-20-001, and in part by the Startup Program of Shenzhen Peacock Plan.

CONFLICT OF INTEREST

We confirm that all authors of this submission have understood the journal’s licensing policy. Further information about the journal’s license agreement(s) and Terms of Use can be found in the journal’s Author Guidelines. We are aware that this journal requires all authors to disclose any potential sources of conflict of interest in the manuscript. We confirm that all the research meets the ethical guidelines, including adherence to the legal requirements of the study country. We confirm that the manuscript has been submitted solely to this journal and is not published, in press, or submitted elsewhere. We confirm that each of the co-authors acknowledges their participation in conducting the research leading to this manuscript and has agreed to its submission to be considered for publication.

PERMISSION TO REPRODUCE

MATERIALS FROM OTHER SOURCES

“None”

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in [“ERA5 download Dongying”] at https://www.meteoblue.com/en/weather/archive/era5/dongying_china_1812101, reference number [33].

ORCID

Xiaoyang Bai https://orcid.org/0000-0003-3680-461X

REFERENCES

1. M. Liserre, et al.: Stability of photovoltaic and wind turbine grid-connected inverters for a large set of grid impedance values. IEEE Trans. Power Electron. 21(1), 263–272 (2006)
2. P. Dehghanian, et al.: Predictive risk analytics for weather-resilient operation of electric power systems. IEEE Trans. Sustainable Energy 10(1), 3–15 (2019)
3. M. Mahzarnia, et al.: A review of the measures to enhance power systems resilience. IEEE Syst. J. 14(3), 4059–4070 (2020)
4. Photovoltaic Power Prediction System (SPSF-3000). http://sprinx.chinapower.com.cn/sell/itemid-1814.shtml accessed 26 May 2020
5. Alam, S.: Prediction of direct and global solar irradiance based on cloud observations. Sol. Energy 78(2), 157–162 (2005)
6. Chang, J.M., et al.: A proposed modified efficiency for thermosyphon solar heating systems. Sol. Energy 76, 693–701 (2004)
7. C. Ertekin, F. Evrendilek: Spatio-temporal modeling of global solar radiation dynamics as a function of sunshine duration for Turkey. Agric. For. Meteorol. 145(1–2), 36–47 (2007)
8. S.G. Jimmy Ehnberg, H.J. Math Bollen: Simulation of global solar radiation based on cloud observations. Sol. Energy 78(2), 157–162 (2005)
9. L. Gu, et al.: Shugart Cloud modulation of solar irradiance at a pasture site in southern Brazil Agric. For. Meteorol. 106, 117–129 (2001)
10. D.L. Liu, B.J. Scott: Estimation of solar radiation in Australia from rainfall and temperature observations. Agric. For. Meteorol. 106(1), 41–59 (2001)
11. O. El Alani, H. Ghennioui, A. Ghennioui: Short term solar irradiance forecasting using artificial neural network for a semi-arid climate in Morocco. In: International Conference on Wireless Networks and Mobile Communications, Fez, Morocco, pp. 1–7 (2019)
12. M.Q. Raza, et al.: An improved neural ensemble framework for accurate PV output power forecast. In: Australasian Universities Power Engineering Conference, Brisbane, QLD, Australia, pp. 1–6 (2016)
13. X. Yang, F. Jiang, H. Liu: Short-term solar radiation prediction based on SVM with similar data. In: 2nd IET Renewable Power Generation Conference, Beijing, China, pp. 1–4 (2013)
14. H. Zhou, et al.: Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism. IEEE Access 7, 78063–78074 (2019)
15. Z. Zhen, et al.: Deep learning based surface irradiance mapping model for solar PV power forecasting using sky image. IEEE Trans. Ind. Appl. 56(4), 3385–3396 (2020)
16. H. Eom, Y. Son, S. Choi: Feature-selective ensemble learning-based long-term regional PV generation forecasting. IEEE Access 8, 54620–54630 (2020)
17. M. Parida, et al.: Combined EMD-ELM and OS-ELM techniques based on feed-forward networks for PV power forecasting. In: Technologies for Smart-City Energy Security and Power, Bhubaneswar, India, pp. 1–6 (2018)
18. L. Chong, et al.: Short-term PV generation forecasting based on weather type clustering and improved GPR model. In: China International Conference on Electricity Distribution, Xi’an, China, pp. 1–5 (2016)
19. Long, H., et al.: Analysis of daily solar power prediction with data-driven approaches. Appl. Energy 126, 29–37 (2014)
20. Eom, H., Son, Y., Choi, S.: Feature-selective ensemble learning-based long-term regional PV generation forecasting. IEEE Access 8, 54620–54630 (2020)
21. Y. Li, et al.: Short-term forecast of power generation for grid-connected photovoltaic system based on advanced grey-Markov chain. In: International Conference on Energy and Environment Technology, Guilin, China, pp. 275–278 (2009)
22. J. Tan, C. Deng: Ultra-short-term photovoltaic generation forecasting model based on weather clustering and Markov chain. In: IEEE 44th Photovoltaic Specialist Conference, Washington, DC, pp. 1158–1162 (2017)
23. H. Yang, et al.: A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output. IEEE Trans. Sustainable Energy 5(3), 917–926 (2014)
24. C. Pan, J. Tan: Day-ahead hourly forecasting of solar generation based on cluster analysis and ensemble model. IEEE Access 7, 112921–112930 (2019)
25. K.Y. Bae, et al.: Hourly solar irradiance prediction based on support vector machine and its error analysis. IEEE Trans. Power Syst. 32(2), 935–945 (2017)
26. Y. Yu, J. Cao, J. Zhu: An LSTM short-term solar irradiance forecasting under complicated weather conditions. IEEE Access 7, 145651–145666 (2019)
27. Y. Wang, et al.: LASSO and LSTM integrated temporal model for short-term solar intensity forecasting. IEEE Internet Things J. 6(2), 2933–2944 (2019)
28. X. Zhang, et al.: A solar time based analog ensemble method for regional solar power forecasting. IEEE Trans. Sustainable Energy 10(1), 268–279 (2019)
29. A. Mellit, et al.: Short-term forecasting of power production in a large-scale photovoltaic plant. Sol. Energy 105, 401–413 (2014)
30. Munkhammar, J., et al.: Probabilistic forecasting of high-resolution clear-sky index time-series using a Markov-chain mixture distribution model. Sol. Energy 184, 688–695 (2019)
31. M. Ding, N. Xu: A method to forecast short-term output power of photovoltaic generation system based on Markov chain. Power Syst. Technol. 1, 152–157 (2011)
32. S.M. Verma, et al.: Markov models based short term forecasting of wind speed for estimating day-ahead wind power. In: International Conference on Power, Energy, Control and Transmission Systems, Chennai, India, pp. 31–35 (2018)
33. Dongying weather data. Accessed on: May 26, 2020. [Online]. Available: https://www.meteoblue.com

How to cite this article: Bai X, Liang L, Zhu X. Improved Markov-chain-based ultra-short-term PV forecasting method for enhancing power system resilience. J Eng. 2021;2021:114–124. https://doi.org/10.1049/tje2.12015