A Very Short-Term Forecasting Maximum Lyapunov Exponent Tool for Distributed Photovoltaic Output

Lingwei Zheng, Zhaokun Liu, Junnan Shen, Chenxi Wu
(School of Automation, Hangzhou Dianzi University, Hangzhou, 310018, China)

Abstract: Photovoltaic (PV) power generation varies randomly and intermittently with weather. For a microgrid with PV sources, this fluctuation not only affects the necessary configuration of energy storage capacity chosen in microgrid planning and design but also influences microgrid operation. Consequently, accurately forecasting PV output is crucial. Towards the operation of a PV-dominated microgrid, a new method for very short-term(VST) forecasting based on the maximum Lyapunov exponent (MLE) is proposed. Firstly, historical power generation data are divided into three weather conditions, sunny, cloudy and rainy days. Then, a PV output series in the different weather conditions is constructed, and the chaotic characteristic is verified by reconstructing an attractor graph and calculating the MLE. Finally, based on the MLE method, the PV generation under different historic weather conditions is forecasted. The raw output time series are measured data from a demonstration system installed on the roof of Building 6 at Hangzhou Dianzi University, China. Forecasting performance accuracy is evaluated by several statistical metrics and is also compared to forecasts obtained by the widely used auto-regression (AR) approach. Comparing forecasts indicates the MLE-based method is statistically but not universally more accurate for VST forecasting.
Keywords: photovoltaic power generation; weather type; chaotic sequence; maximum Lyapunov exponent; very short-term forecasting, microgrids

0 Introduction

Renewable energy resources should replace traditional power generation because of their desirable characteristics of sustainability and low pollution. Photovoltaic(PV) electricity generation is an excellent renewable energy source by these criteria, but its output varies significantly depending on the weather, especially in cloudy climates such as southeastern China. PV power is actually both uncertain and variable. On the one hand, large-scale PV being interconnected to the utility grid might affect safety and reliability, raising the highest barrier to its large-scale utilization. On the other hand, if demand is served locally by PV within a microgrid, even at low solar penetration, stability problems could arise for microgrid operation, e.g., it might be hard to maintain the voltage and frequency in an islanded PV-dominated inertialess microgrid. Using battery banks large enough to smooth inevitable demand-supply imbalances is the usual strategy, but it significantly increases capital costs and can lead to pollution from battery disposal. All of these effects have economic implications for future microgrid viability. Accurate prediction of PV power can provide an important input to microgrid power dispatching and operation by facilitating adjustments to the operational plan in time to optimize performance and minimize cost. Because the consequences of PV's randomness and variability can be mitigated by accurate forecasting and its full use effectively encouraged, solar forecasting has received considerable attention in the literatures.

The long history of research on forecasting PV generation has been fruitful. Fundamentally, the technology can be divided into two categories: physical modeling and data-driven methods. The physical modeling approach focuses on studying equivalent circuits of PV cells to forecast power output based on predicted weather input parameters, such as irradiation and temperature[1]. These models can require many circuit parameters, such as series resistance, shunt resistance, diode reverse saturation current, photocurrent, diode impact factor and various temperature coefficients. Acquiring some of these data sets requires complex calculations, while others can only be provided by manufacturers. For example, [2] uses a physical model requiring a large amount of raw sampled data to estimate the PV system parameters used in the simulation. Consequently, physical modeling has limited real-world application, and data-driven based methods are used more in practice. Further, rapid progress in data processing, artificial intelligence, and machine learning are giving data-driven methods a stronger advantage and they are now being used widely in various forecasting applications [3].

Data-driven methods can be further sub-divided mathematically into linear and nonlinear models. The popular linear models applied in PV forecasting are auto-regression (AR), auto-regression extrapolation (ARX), and auto-regressive moving average extrapolation (ARMAX). Reference [4] proposes an online method using AR and ARX respectively to forecast PV production, indicating that AR is effective for forecasting up to 2 h ahead while an ARX model taking numerical weather prediction (NWP) as input performs better over longer horizons. Paper [5] considers a time series ARIMA model unable to include meteorological information and suggests an ARMAX model, which can include exogenous inputs to improve accuracy. Work [6] proposes a vector auto-regression framework be applied to PV forecasting with a spatial-temporal forecasting method supplying 6 h ahead forecasts. These linear models are simple in structure but inflexible [7], and actual PV output is absolutely nonlinear, so many researchers focus on studying nonlinear forecasting. Among the nonlinear models, the artificial neural network (ANN), the support vector machine (SVM), and their enhanced counterparts combined with other composite methods are widely used [8-13]. Reference [8] combines ANN with statistical parameters to improve the short-term solar irradiance forecasting accuracy. Reference [12] applies a gradient-descent optimization technique and a shuffled frog leaping algorithm to ANN to secure original parameters and optimize them, respectively. To decrease the computational cost and execution time

while assuring high predictive performance, [13] integrates principal component analysis and least squares SVM to cope with high dimensionality data. Compared to linear methods, these approaches significantly improve precision, but the forecast results depend on model structure, network parameters, and initial values.

Data-driven forecasting methods can also be sub-divided into three models according to the raw data used, i.e. time series, NWP, and sky imaging. Time series models are based only on historic PV generation data. No other input data is necessary, so there are no experimental requirements. Linear models like AR and ARMA and their extended versions are usually used for time series [4], while nonlinear time series models are not common. NWP is commonly used with control inputs [4,14-17]. For example, reference [16] takes the satellite medium-range weather forecast data from the European Centre for Medium-Range Weather Forecasts (ECMWF) in Reading, U.K., and surface data as ANN inputs, greatly improving accuracy. With the progress of image processing, predicting PV output based on sky imaging is attracting researcher attention. This method takes ground-based or satellite cloud images and forecasts their movement and shading, thereby predicting surface insolation [18-24]. Which model(s) to use depends heavily on experimental conditions and available information. For example, NWP methods need considerable weather data and sky imaging requires satellite cloud pictures or all-sky imaging equipment, which are both expensive.

In terms of time horizon, PV forecasts can be categorized as long-term, short-term, or very short-term (VST), but researchers have very different ideas about the time cutoffs. For example, in [25] and [26], the VST horizon is defined as several minutes to several hours, the short-term horizon ranges from a few hours to 3 days, and the long-term is defined as from a week to a year. While 5~8 minutes is defined as short-term in reference [27], in reference [28], two weeks ahead is considered short-term. No matter how they are defined, the roles of the forecast periods are similar. The VST forecast aims at intra-day real-time control and power market participation; the short-term forecast is used for day-ahead economic dispatch; and the long-term one focuses on equipment maintenance scheduling, market participation, etc. [29]. In these three time domains, researchers have addressed current problems in short-term forecasting [10-12,30-33].

The shorter the horizon, the more conducive a forecast is to emergency response. Thus, the VST PV forecast is particularly crucial. VST PV behavior is mainly affected by cloud movement, so this forecast mainly uses the aforementioned NWP, sky images, sensor arrays plus random-sequence and time series, etc. [20,25,34]. These methods make it possible to achieve a precise forecast, but they are limited by the measurement equipment available. For example, cloud imaging of the entire sky needs a fish-eye lens mounted on a whole sky camera with an unrestricted panoramic view of 180 degrees in every direction. These devices are not only expensive but also limited in space-time resolution, and still there are some shortcomings in the PV forecast.

There are still other ways to classify the methods used for PV forecasting. For example, training methods can be classified as online or offline. Compared with offline training, online training adds current values to training samples to achieve a rolling forecast that can provide higher accuracy [4,18,35]. Aiming at the forecast uncertainty from meteorological variables and the forecasting model itself, [36] refreshes PV output prediction by taking recent history outcomes of the prediction error into account.

Although there are many ways to forecast, it is difficult to say which one is better, given different time scales, geographies, PV capacities, and experimental conditions [37-39]. This paper summarizes the advantages and disadvantages of the AR forecast methods outlined above compared to a proposed VST forecast method based on the maximum Lyapunov exponent (MLE) approach that analyzes the chaotic characteristic of an actual evolving PV generation time series.

While traditional AR model has good linear forecast performance, MLE being a nonlinear time series analysis method promises greater accuracy. Given the sensitivity of chaotic systems, MLE does not directly construct a mathematical model linking output to its influencing factors. Instead, it describes the evolution of the dynamic system according to the exponential separation characteristic of the actual chaotic data to yield a nonlinear time series VST forecast. The approach does not require a complex experiment or expensive equipment and no procedures are needed to estimate parameters for initial inputs. Because of these characteristics the MLE method is widely used in economics, medicine, astronomy, and other fields [40-43]. As examples, [40] applies MLE to perform one-week out-of-sample prediction of the stock returns, a MLE-based method for tourist arrivals forecasting is used in [41], and [43] develops an MLE-based algorithm for epileptic seizure prediction; nonetheless, this approach has hardly ever been used in PV forecasting.

Based on the advantages of MLE, it is adopted here for VST PV forecasting, aiming to seek a precise, low-cost, and stable forecasting method. Firstly, statistics are calculated and a PV output sequence is built for various weather conditions, and then, the chaotic characteristic of the power sequence is tested and verified. Secondly, historical PV generation sequences matched to their corresponding weather forecasts are chosen for learning data sets. Then, the MLE can be estimated by reconstructing the phase space and calculating two of its key factors, delay time and embedding dimension. Forecast values can then be obtained by MLE and the recent forecasted values appended to the base sequence to calculate new exponents, establishing an online multi-step forecast.

1 Data set

This paper uses measured historic data from a China demonstration PV microgrid located on the campus of

Hangzhou Dianzi University (HDU) at latitude 30°19′N and longitude 120°20′E. The microgrid consists of a total 120 kW of PV arrays, a 120 kW diesel generator, a 50 kW × 1 h battery bank, and a set of approximately 100 kW × 2 s supercapacitors, serving the loads in HDU's Buildings 6 and 8. The PV arrays are installed on the root-top of Building 6, facing solar south at an inclination of 20°, and with maximum power point tracking inverters. The Building 6 PV installation is shown in Fig. 1 and the roof layout appears in Fig. 2. PV output power is measured by a Hokio3196 power quality analyzer at a one second sampling rate, and data are logged and stored in an online database.

Fig. 1 Building 6 PV Array

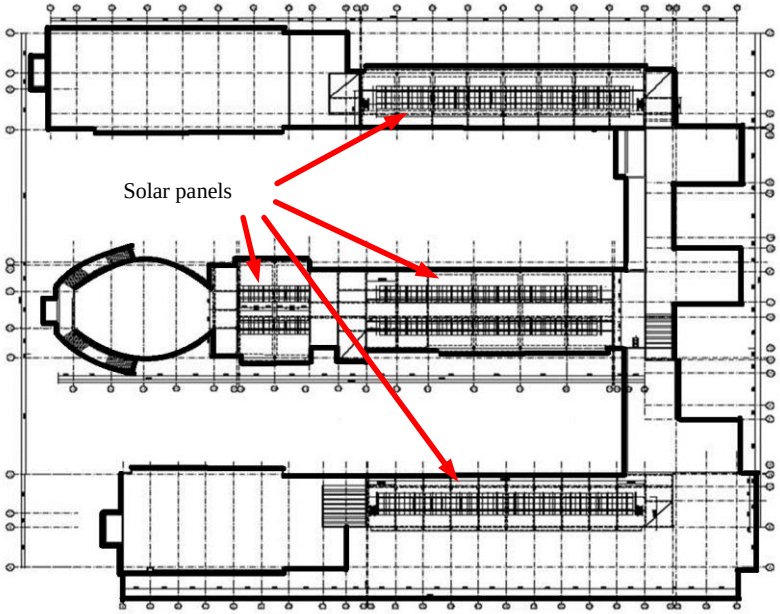
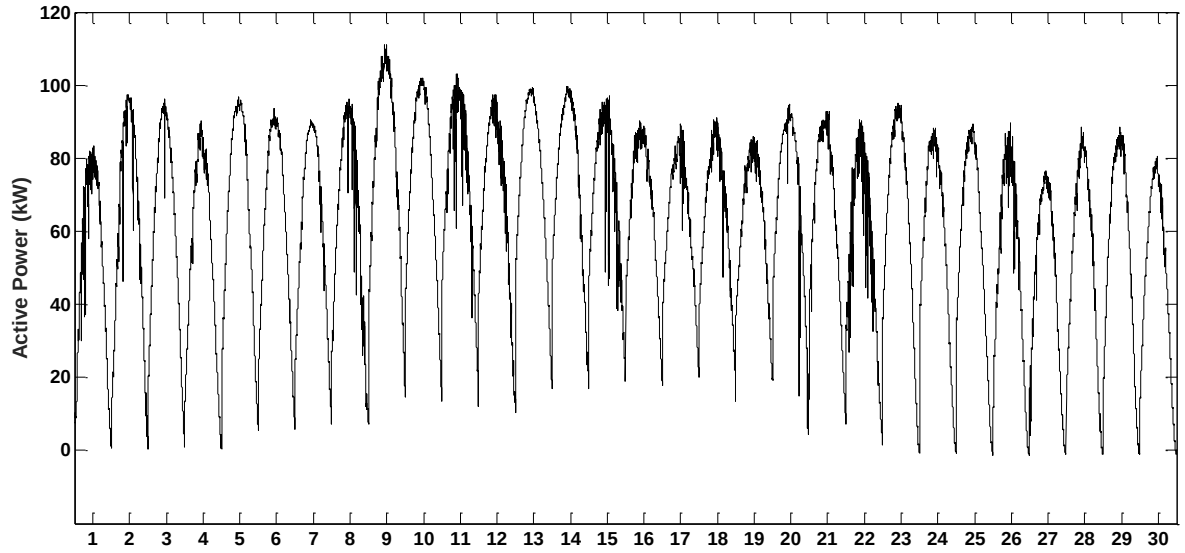
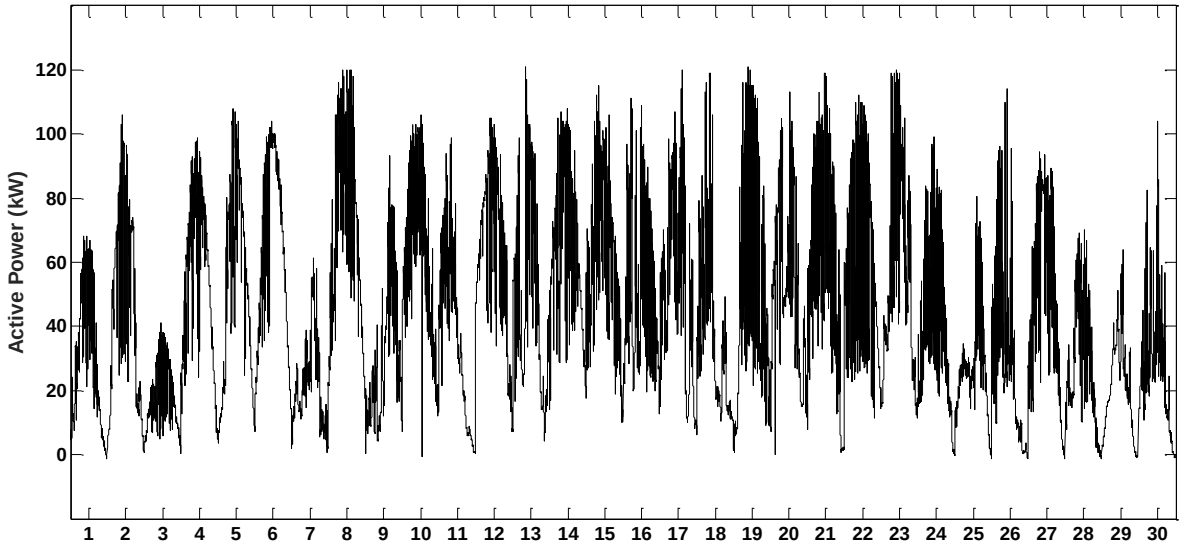


Fig. 2 Array Configuration of Building 6 at HDU

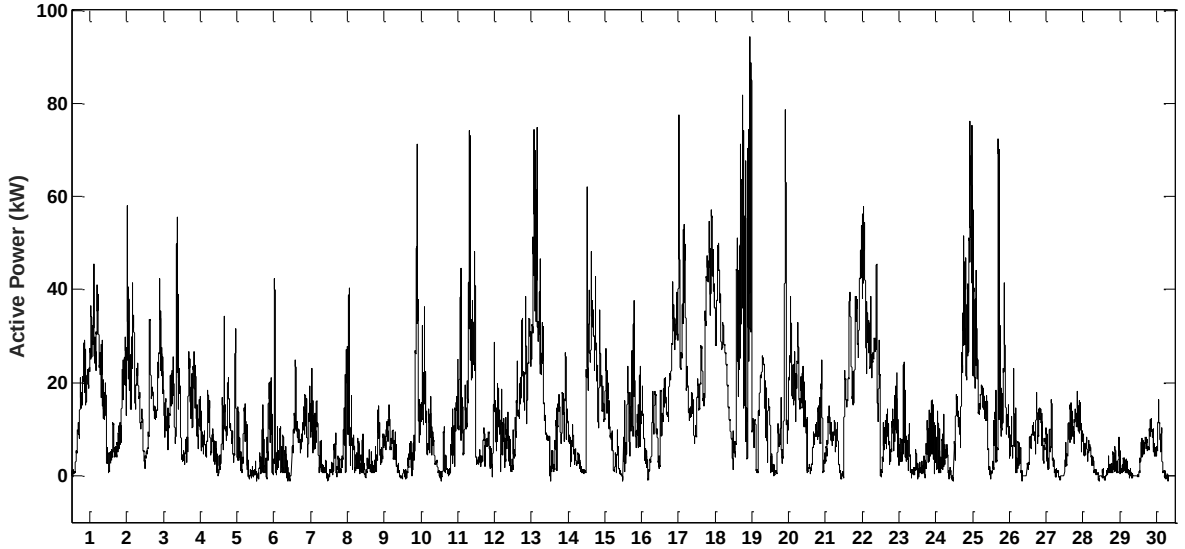
In this work, PV output power was collected from 8:00 to 17:00 for almost every day during 2017, and 30 non-sequential example days were selected for each of three weather types, clear, cloudy, and rainy. The 1-min sample interval selected data curves for the three weather regimes are shown as Fig. 3, and in the same order, clear, cloudy, and rainy. Besides cloudiness, air temperature is another factor affecting PV output, so several days in each season are included for each weather type.



(a) Sunny Days



(b) Cloudy Days



(c) Rainy Days

Fig. 3 PV Power Generation Curves During the Three Weather Types

2 Chaos Analysis of PV Generation Sequences

2.1 Reconstruct the phase space

According to Packard et al., the phase space of a dynamic system can be reconstructed from its time series since the time series itself contains information about all the variables involved in this system [44]. Takens proved this and proposed this lower bound of embedding dimension in phase space, that is [45]

$$m \geq 2d + 1 \tag{1}$$

where, m is the embedding dimension, and d is the correlation dimension of the dynamic system in which the strange attractor lies. Takens shows that as long as the embedding dimension is large enough, a regular trajectory can be recovered from the reconstructed embedding phase space. Therefore, the forecasting problem can be transformed into a short evolutionary process in the phase space, which provides the theoretical basis for the time series forecast.

There are two methods for reconstructing the phase space, derivative reconstruction and coordinated delay reconstruction, with the latter adopted commonly in practical applications. For time series data $x_t, t = 1, 2, \cdots n$, first N data points ($N<n$) can form the first vector $X_1 = (x_1, x_2, \cdots\cdots x_N)$, then each item in the first vector is shifted backward in an orderly way by τ to obtain the second vector $X_2 = (x_{1+\tau}, x_{2+\tau}, \cdots x_{N+\tau})$, and m ($m<N$) moves in this way are executed, extending the time series from one dimension to m dimensions. Now the set $\{X\}$ can be obtained, as shown in Eq. (2):

$$X = \begin{bmatrix} X_1^T \\ X_2^T \\ \cdots \\ X_N^T \end{bmatrix} = \begin{bmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(m-1)\tau} \\ x_2 & x_{2+\tau} & \cdots & x_{2+(m-1)\tau} \\ \cdots & \cdots & \cdots & \cdots \\ x_N & x_{N+\tau} & \cdots & x_{N+(m-1)\tau} \end{bmatrix} \tag{2}$$

$$N = n - (m-1)\tau \tag{3}$$

where, τ is the delay time, which together with embedding dimension are the two key parameters in the reconstruction of phase space. Their proper selection can improve the quality of phase space reconstruction and thereby improve forecast accuracy. In engineering, there are Grassberger-Procaccia (G-P), neighbor point dimension and Cao methods for counting the embedding dimension [46-48]. For the delay time, there are autocorrelation, improved autocorrelation, and mutual information methods [49,50]. Additionally, there are some integrated methods, such as the C-C (named from a set of equations as Eq.(11)-Eq.(14) as explained in [51]), time window, and automatic algorithm methods, to estimate both the delay time and the embedding dimension. The C-C method, using the correlation integral to count both the delay time and the correlation dimension with low calculation cost and strong anti-interference ability [51], is adopted for this study.

2.2 The principle of chaos determination and the MLE

Chaotic behavior, which looks random and disorderly, exists in many natural and man-made systems, such as weather, climate and road traffic[52,53]. There is no universally accepted definition of chaos so far, but the most common one found in expository literature is put forward by Robert Devaney[54]. He asserts a dynamic system can be said to be chaotic if it is sensitive to initial conditions, topologically transitive, and its periodic points are dense[55]. The sensitivity to initial conditions means an arbitrarily small perturbation of the current trajectory may lead to significantly different future behavior, popularly known as the “butterfly effect”. Topological transitivity requires that a dynamic system cannot be broken down or decomposed into two disjoint open subsystems. And, having dense periodic orbits means that in the midst of this random behavior, there is nevertheless an element of regularity, namely dense periodic points.

Chaotic behavior can also be identified by calculating the characteristic parameters of the chaotic signal’s strange attractor. Three representative methods are MLE, Kolmogorov entropy, and the correlation dimension method [56,57].

In this paper the MLE method is used to determine the chaotic characteristic of the analyzed PV output sequence. The Lyapunov exponent is a parameter that describes the adjacent orbit divergence rate of the initial distance in the phase space and plays an important role in characterizing the chaotic properties of a dynamic system. The Lyapunov exponent determines whether adjacent tracks can be brought together to form a stable track or fixed point. When the value of the Lyapunov exponent is positive, the system displays chaotic behavior. Otherwise, the motion of the system will converge to a fixed point. The system appears periodic when the Lyapunov exponent equals zero [58].

In real world chaos system identification, only the MLE is needed, and the common ways to obtain it are Wolf, Jacobian and small data set algorithms. The Wolf method is based on the definition of the Lyapunov exponent, which is estimated from the evolution of the phase trajectory, phase plane, phase volume, etc., and is widely used in chaotic system research[59].

According to the reconstructed phase space X_i in Eq. (2), if we take the initial position as $X(t_0)$, its adjacent point is $X'(t_0')$. Then the Euclidian Distance between $X(t_0)$ and $X'(t_0')$ is L_0, and

$$L_0 = \left\| X(t_0) - X'(t_0') \right\|, \quad |t_0 - t_0'| > p \tag{4}$$

where, p is the average period calculated by Fast Fourier Transform (FFT). The initial position and the adjacent

point evolves to $X(t_0+\Delta t)$ and $X'(t_0'+\Delta t)$, respectively, after n iterations. Evolution of the two points with time evolving are tracked until the distance between the two points exceeds a predetermined threshold, ε (ε>0), as shown in Eq. (5):

$$L_0' = \left\| X(t_0+\Delta t) - X'(t_0'+\Delta t) \right\| > \varepsilon \tag{5}$$

According to the exponential separation rule of chaotic phase space:

$$L_0' = L_0 e^{\lambda_0 \Delta t} \tag{6}$$

yielding the exponential growth rate, λ_0:

$$\lambda_0 = \frac{1}{\Delta t} \ln \frac{L_0'}{L_0} \tag{7}$$

Similarly, it is assumed $X(t_0+\Delta t)$ is taken as a new initial point, $X(t_1)$, and its adjacent point $X'(t_1')$ is found such that the Euclidian distance between the two points, L_1, complies with Eq. (8):

$$L_1 = \left\| X(t_1) - X'(t_1') \right\| < \varepsilon \tag{8}$$

In the evolution process, the angle should be kept between L_0 and L_1, and as small as possible. The iterative evolution continues, and the exponential growth rate λ_2 is estimated until $X(t)$ reaches the end of the time series. Taking the total steps as M, the estimated value of λ_m can be derived from the average of exponential growth as Eq. (9):

$$\lambda_m = \frac{1}{M} \sum_{i=1}^{M} \frac{1}{\Delta t} \ln \frac{L_i'}{L_i} \tag{9}$$

2.3 Chaos determination of PV output sequences

As mentioned above, PV power strongly depends on weather, and is more uncertain on cloudy days than sunny days. Historic data is classified according to prevailing weather. In the era of big data, historical weather can generally be easily obtained from the Internet, but if historical weather is not available, it can be calculated by the Euclidian distance between the object day and a typical day.

Here the Euclidian distance represents the distance between two days in the time series, as shown in Eq. (10)

$$d = \left\| x_i - y_i \right\| \tag{10}$$

where, x_i is the average power generation at each moment in the typical weather type; and y_i is the power generation power at each moment in the objective day.

In this paper, the HDU data set described in Section 1 is sampled and typical PV output under three weather types is chosen to create a time series based on historical weather forecasts from the Internet. The raw sampling time interval is 1 s making the data set unmanageably large, with too much time needed for processing. To adequately and feasibly identify the chaotic characteristic of the PV output sequence, the data set is reduced by counting the average power at three time intervals, 30 s, 60 s, and 90 s. The MLEs calculated from the data are shown in Table 1. The three weather types have different delay times and embedding dimensions at the three sampling time intervals, as well as different sequence lengths, but the MLEs are all positive. Accordingly, it is concluded that PV power generation exhibits chaotic behavior no matter which weather type prevails.

Table 1. Characteristic Parameters of the Time Series Under Three Weather Types

Weather Type	Simpling Interval (s)	Length of Time Series (observations)	Embedding Dimension	Delay Time (min)	MLE
Sunny	90	3600	2	15	0.0474
	60	5400	3	12	0.0294
	30	10800	3	7.5	0.0273
Cloudy	90	3600	2	13.5	0.6145
	60	5400	2	11	0.4233
	30	10800	3	8	0.3251
Rainy	90	3600	2	15	0.3806
	60	5400	2	11	0.3369
	30	10800	3	6	0.101

Taking a sunny day as an example, the attractor of reconstructed phase space, the progression of numerical values towards which the system tends to evolve, is shown in Fig.4. As expected, the attractor in the three-dimensional embedding space exhibits a remarkable trajectory. It is topologically transitive and has dense periodic orbits, which are consistent with two necessary properties of chaotic systems.

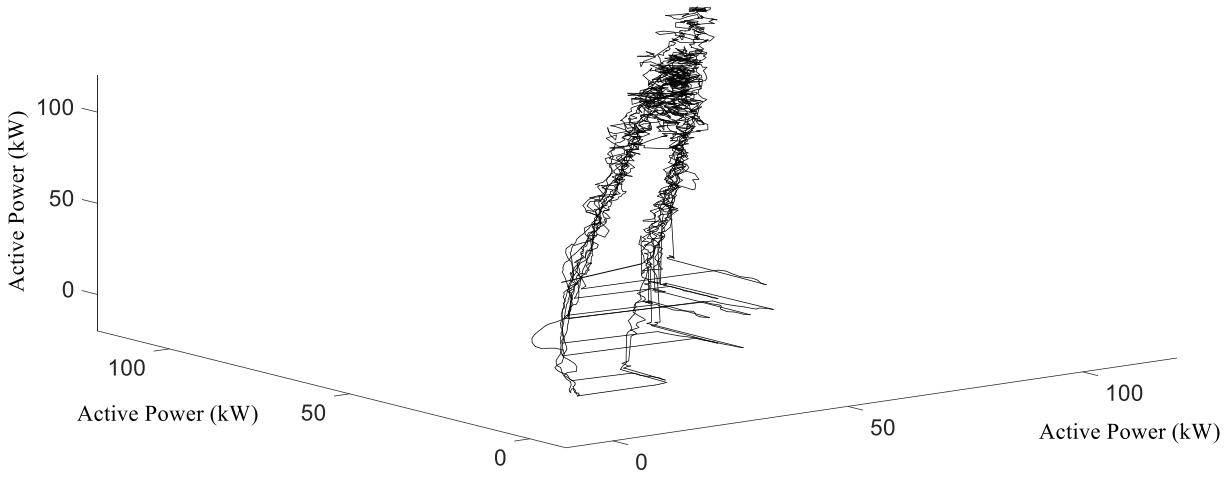


Fig. 4 Reconstructed Attractor Graphs of Sunny Days

3 VST Forecast Based on the MLE for PV power generation

3.1 Forecasting Steps

MLE-based forecasting tracks the evolution of nearest points in the reconstructed phase space and calculates MLEs that realize the forecast. The detailed steps are as follows:

1) Establish PV power generation time series

In this step, historical data is sorted by prevailing weather to form historical time series under each weather regime $\{x(t), t=1,2,\cdots,n\}$, and then, combine the current weather type and generation to construct the time series.

2) Apply a FFT to calculate average period [60]

The time series $\{x(t), t=1,2,\cdots,n\}$ is converted to frequency components $F(k)$ by FFT.

$$F(k)=\sum_{n=1}^{N} x(n)e^{-j2\pi(k-1)\frac{n-1}{N}} \quad k=1,2\ldots N \tag{11}$$

If it is assumed

$$F(K)=MAX(F(2),F(3),\cdots F(N)) \tag{12}$$

the average period is:

$$P=\frac{1}{K} \tag{13}$$

where, K is the frequency of the maximum frequency component; P is the average period.

3) Calculate the parameters of reconstructed phase space

The delay time, τ, and correlation dimension, d, are counted via the C-C method and the embedding dimension m is obtained according to Eq. (1).

4) Reconstruct the phase space

The phase space is reconstructed as $\{X(i),\ i=1,2,\cdots,N\}$ based on τ and m.

5) Calculate the MLE, λ_m, and forecast horizon T

The MLE, λ_m, is calculated using Eq. (10) and the forecast horizon T is determined by Eq. (14)

$$T=\frac{1}{\lambda_m} \tag{14}$$

The forecast horizon T represents the maximum time required to double the system state error and is considered one of the reliability indices of the VST forecast.

6) Forecast PV generation via the MLE λ_m

In this step, the last value of the time series is taken to be the center point to obtain forecast values. If X_N is in the time series and X_{N+1} is its adjacent point

$$\left\|X_{N+1}-X'_{N+1}\right\|=\left\|X_N-X'_N\right\|e^{\lambda_m} \tag{15}$$

Eq. (15) can be expanded as

$$(x_{n+1}-x'_{n+1})^2+\sum_{i=1}^{m-1}(x_{N+i\times\tau+1}-x'_{N+i\times\tau+1})^2=\left\|X_N-X'_N\right\|^2 e^{2\lambda_m} \tag{16}$$

From Eq. (16), the forecast value is:

$$x_{n+1} = x'_{n+1} \pm \sqrt{\left\| X_N - X'_N \right\|^2 e^{2\lambda_m} - \sum_{i=1}^{m-1} (x_{N+i\times\tau+1} - x'_{N+i\times\tau+1})^2} \tag{17}$$

7) Prepare for the next forecast

The forecast value x_{n+1} is appended to the base series $\{x(t), t = 1, 2, \cdots, n\}$ to construct a new series and the new MLE is calculated from the extended series. Steps 6 and 7 are repeated till the forecast horizon is reached.

A forecasting flow chart based on the steps described above is shown in Fig.5.

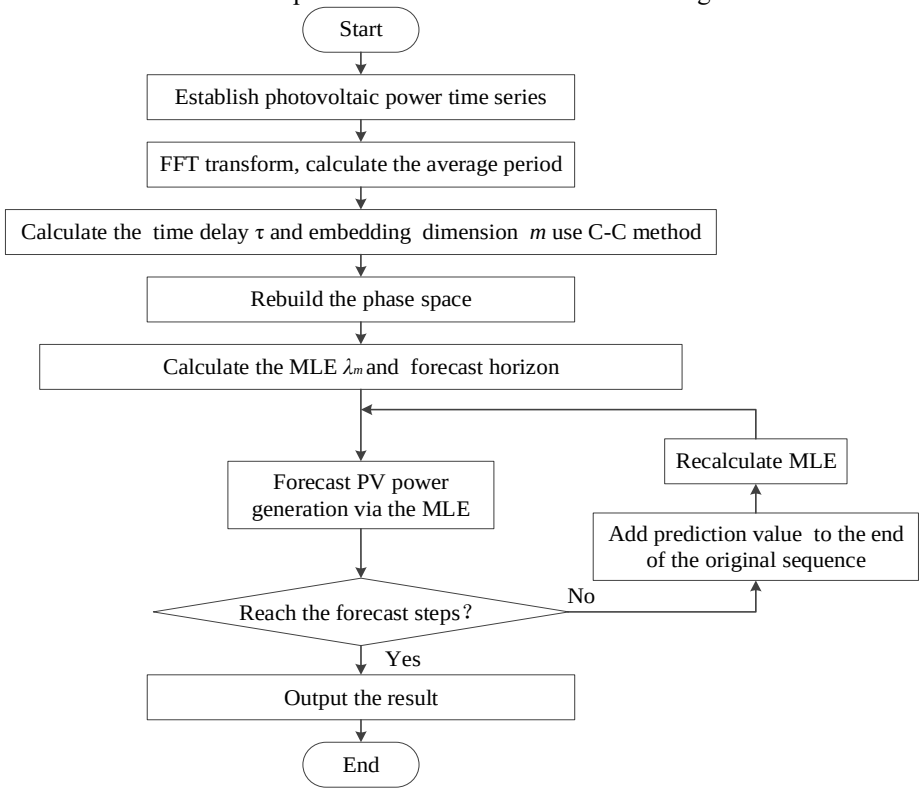


Fig. 5 VST PV Forecast Flow Using MLE

3.2 Assessment metrics

In order to assess forecasting method accuracy, three of the most common statistical error indices are considered [1]:

mean absolute percentage error (MAPE)

$$E_{MAPE} = \frac{1}{N} \sum_{i=1}^{N} \frac{\left| W_{forecast,i} - W_{true,i} \right|}{W_{true,i}} \times 100\% \tag{19}$$

relative root mean squared error (rRMSE)

$$E_{rRMSE} = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^{N} (W_{forecast,i} - W_{true,i})^2}}{\frac{1}{N} \sum_{i=1}^{N} W_{true,i}} \times 100\% \tag{20}$$

mean bias error (MBE)

$$E_{MBE} = \frac{1}{N} \sum_{i=1}^{N} (W_{forecast,i} - W_{true,i}) \tag{21}$$

In the above definitions, $W_{forecast,i}$ and $W_{true,i}$ are the forecasted and measured PV output at time point *i*, respectively.

4 Results and Discussion

As a comparison to MLE, the performance of an AR model, the most common forecast method for time series, is evaluated. All the raw data are from the HDU data set.

4.1 Parameters in the MLE method

Based on the PV output sequence shown in Fig.3, the MLE parameters for three weather types are calculated and shown in Table 2.

Table 2 Comparison of the MLE Parameters

Weather	Embedding Dimension	Delay Time (min)	MLE	Forecast Horizon (min)
Sunny	3	16	0.0234	42
Cloudy	3	16	0.3311	3
Rainy	3	16	0.0867	11

The embedding dimension and delay time are all the same, 3-dimensions and 16 min. The MLE results confirm expectations: the sunny day MLE is the smallest, so its forecastable horizon is the longest, 42 min; output varies most on cloudy days, and consequently, the MLE is by far the largest, yielding a forecastable horizon size of only 3 min; while rainy days lie between, and can be used for VST forecasts of up to 10 min. Results are consistent with intuition, i.e. the most drastic variation occurs on cloudy days meaning VST forecasts are more difficult.

4.2 Forecasting result comparison to an AR model

As mentioned before, the linear time series AR model is a common forecasting method; therefore, this work uses it for comparison. The two methods are applied to HDU microgrid PV generation and 30 comparison forecasts have been conducted. The forecast horizon is determined by the MLE method. To achieve results in different hours of the day and evaluate the performance statistically, the forecasts are made in 5 time-steps. Meanwhile, 4 representative sunny 2017 days are selected to evaluate temperature's influence on forecast performance: 14 Apr., 28 Aug., 9 Nov., and 25 Dec.. Because of low PV output on cloudy and rainy days, it can be assumed their PV output is insensitive to temperature; therefore, only one day for each of these two weather types are forecast: 27 Dec. for cloudy and 28 Dec. for rainy. Note that the time periods, namely the forecast horizons, differ dramatically for three weather regimes, e.g., the time period for sunny days is about 40 min, but only 3 min for cloudy days. The error indices comparison for sunny days in all four seasons, and cloudy and rainy days for winter only are shown in Tables 3 and 4. As is clear in Table 3, the MLE method performs well on sunny days for this 120 kW PV array, with most forecasting errors under 10 kW, and the MAPEs are mostly below 10%. The MLE method's forecast accuracy is usually better than the AR, but not universally. From Table 4, the MLE method shows distinct advantage in the cloudy day and rainy day. Short forecast horizons are the MLE's weakness.

Table 3 Sunny Day Forecasting Result Comparison for all Four Seasons

		Forecast Horizon (min)	MLE			AR		
			MAPE	rRMSE	MBE (kW)	MAPE	rRMSE	MBE (kW)
Spring	1	45	4.76%	5.21%	3.6079	6.75%	8.11%	5.2864
	2	48	6.77%	7.09%	4.8854	6.76%	8.00%	5.0323
	3	45	4.17%	5.16%	3.3295	6.99%	8.08%	5.5994
	4	39	7.74%	8.27%	7.6206	12.93%	13.78%	12.681
	5	39	9.34%	9.68%	9.2691	15.40%	17.08%	15.284
Summer	1	33	2.40%	2.96%	2.1958	7.69%	9.11%	6.107
	2	35	2.39%	2.79%	2.2041	8.62%	9.83%	6.8561
	3	41	12.78%	13.11%	11.793	7.55%	8.90%	6.0663
	4	38	10.73%	11.48%	9.9358	7.22%	8.62%	5.8411
	5	34	2.00%	2.68%	1.8216	7.73%	9.58%	6.3146
Autumn	1	32	10.54%	10.99%	9.0898	12.40%	13.34%	12.206
	2	31	8.31%	8.62%	7.2728	16.02%	17.78%	15.971
	3	31	6.97%	7.11%	6.0932	17.34%	18.64%	17.285
	4	29	6.70%	6.84%	5.8029	18.97%	20.10%	18.833
	5	30	5.57%	5.77%	4.8124	18.97%	18.21%	16.321
Winter	1	42	10.42%	10.96%	6.8873	11.87%	14.39%	7.5163
	2	30	8.02%	8.68%	6.1441	10.38%	11.86%	7.9696
	3	29	5.69%	5.95%	4.454	5.85%	6.48%	4.5886
	4	37	5.36%	5.67%	4.181	6.63%	7.38%	5.1544
	5	27	10.36%	10.93%	7.828	14.71%	15.89%	11.134

Table 4 Cloudy and Rainy Day Forecasting Result Comparison

		Forecasts Horizon (min)	MLE			AR		
			MAPE	rRMSE	MBE (kW)	MAPE	rRMSE	MBE (kW)
Cloudy day	1	2	17.35%	17.82%	4.8047	22.78%	24.50%	6.1844
	2	2	2.55%	2.65%	0.743	68.83%	77.36%	20.069
	3	3	10.36%	11.56%	2.942	19.42%	20.17%	5.7526
	4	2	0.52%	0.64%	0.134	61.85%	76.32%	16.324
	5	2	23.71%	23.72%	7.851	62.83%	63.61%	21.571
Rainy day	1	11	8.63%	12.39%	0.7331	73.61%	74.90%	6.4974
	2	11	97.07%	98.93%	6.273	81.68%	92.36%	5.4886
	3	11	52.26%	60.07%	3.306	51.22%	55.62%	3.6944
	4	11	26.41%	30.38%	2.766	70.75%	75.90%	7.3554
	5	11	32.42%	43.20%	5.041	56.99%	57.89%	8.1202

Box and whisker graphs of the Table 3 and 4 results appear as Figs. 6 and 7, with the MAPE index selected to compare for sunny and cloudy days. It should be noted that on the rainy days, the MAPE and rRMSE are very large even though the absolute forecasting error is as small as the other days' because of the low total output. Accordingly, it is more illuminating to evaluate prediction errors using MBE on rainy days.

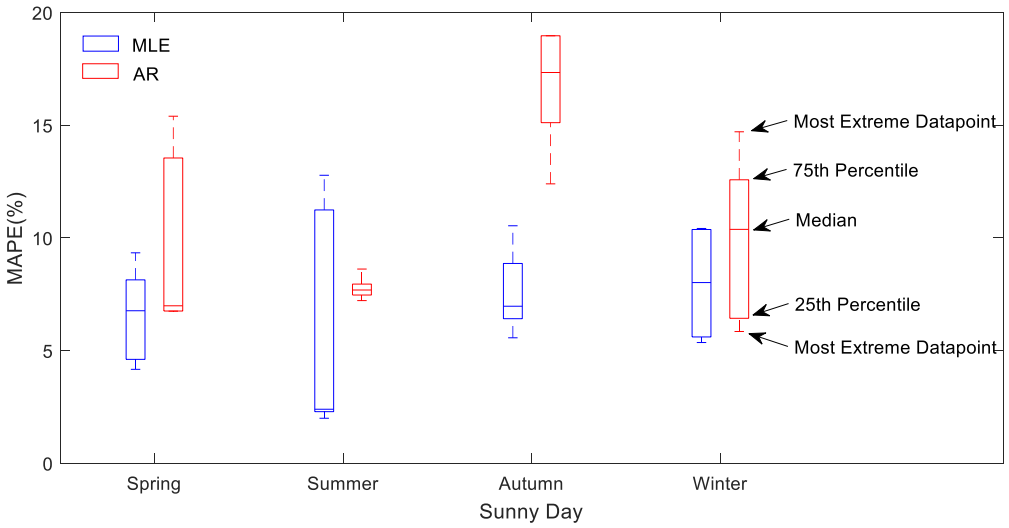


Fig. 6 Sunny Days Comparison for all Four Seasons

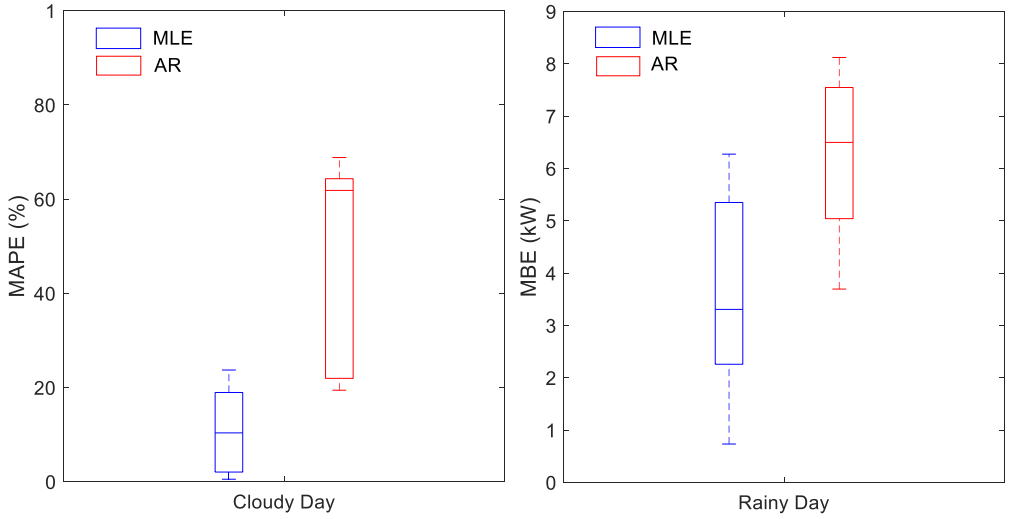


Fig.7 Cloudy and Rainy Day Comparison

From Fig.6, although the two MAPE medians in spring are close, the MLE has more low-MAPEs than AR. In

summer, the AR delivers a more stable forecasting performance, even if the MLE has the lower median. MLE is clearly better than AR in autumn, and winter is similar to spring. Further, from Fig.7, the MLE performs clearly better on the challenging cloudy day. In summary, the MLE forecast results are statistically better than the AR model's, but in some individual cases AR also performs quite well.

Three groups of typical forecasted curves are shown in Fig. 8 as examples. These example results show three magnifications of short time periods during example days of each weather type, 25 December, 11:40~12:17 (sunny day), 27 December, 13:51~13:53 (cloudy day), and 28 December, 13:40~13:51 (rainy day), respectively.

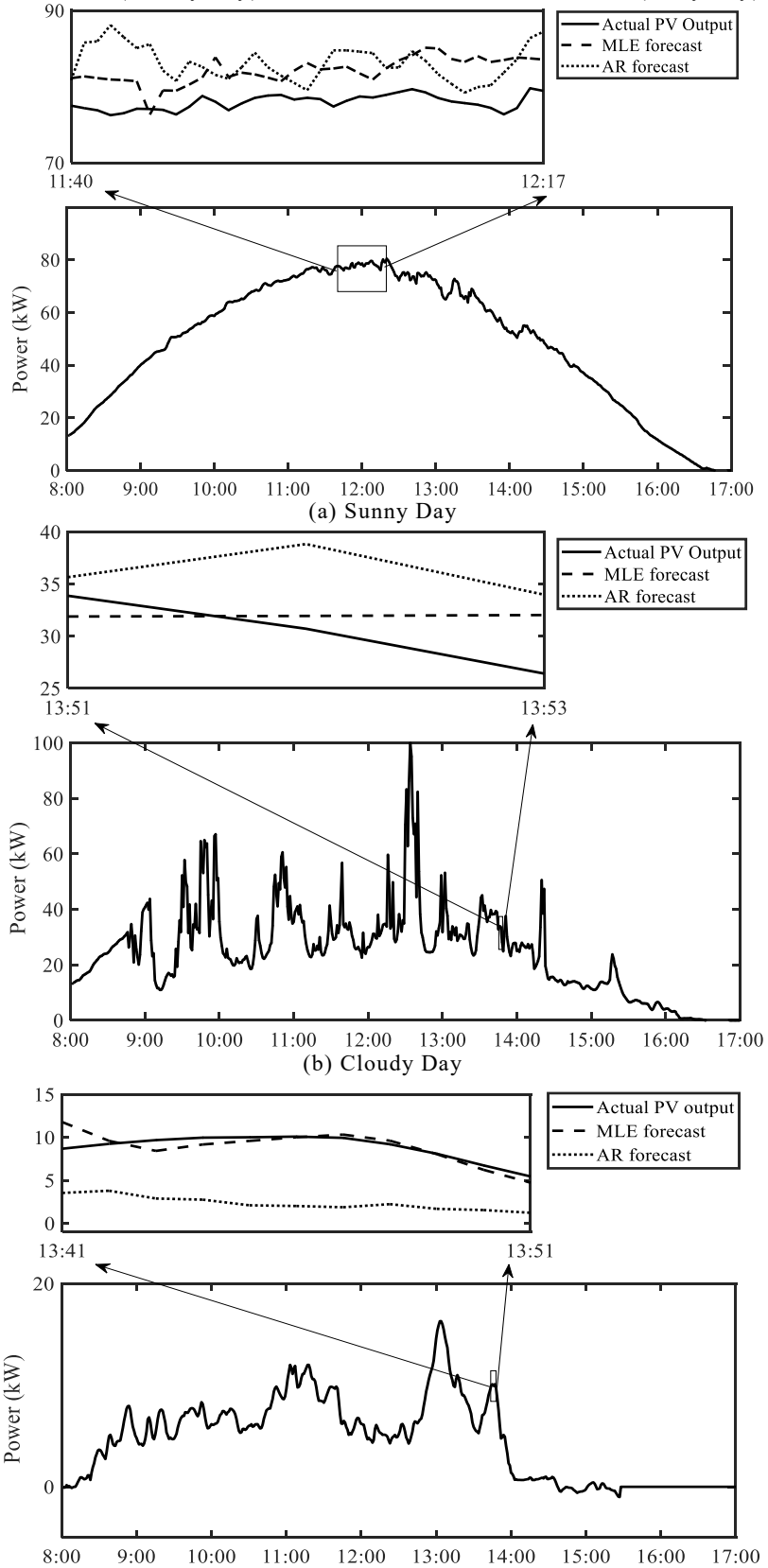


Fig. 8 The PV Power Generation Forecasting

5 Conclusion

The widespread and rapidly growing application of PV makes fully capturing its environmental benefits as effectively as possible a pressing priority, and consequently, output forecasting is justifiably receiving increasing attention. Based on chaos theory and the phase space reconstruction method, this paper proposes a VST forecasting method using MLE. The performance of the MLE method is validated using HDU's demonstration PV system, leading to the following observations:

1) Combined with phase space reconstruction and according to the delay time and the embedding dimension from the C-C method, an MLE is calculated. It confirms that the time series of PV power generation has chaotic characteristic.

2) Based on these, the MLE method is used to forecast PV output under the three weather types. Statistical results show that weather has a great influence on forecasting results from the MLE method. As one might expect, the highest forecast accuracy and longest forecasting horizon are exhibited on the sunny day, followed by the rainy day, and the cloudy day poses the biggest challenge. Accuracy of the MLE method is usually but not universally higher than the AR method for all three weather types.

3) The benefits of the MLE-based approach lie in: its better than linear time series model performance, its use of a nonlinear time series model closer to underlying the physical system, and its reduced experimental costs and greater feasibility compared with NWP and sky imaging methods.

In future research, data collection and archiving data for the HDU demonstration platform will continue, facilitating more in-depth research on VST PV forecasting considering similar days in several different years, and seeking other drivers of better and worse forecasting performance. Also, since results are inconsistent, more comprehensive indicators will also be developed to gauge overall forecasting performance over a wider range of weather and other conditions.

Acknowledgement

The work was supported in part by the National Natural Science Foundation of China (51407052). The authors would like to thank Chris Marnay for his kind and great help in the writing process.

References

[1] Ogliari E, Dolara A, Manzolini G, et al. Physical and hybrid methods comparison for the day ahead PV output power forecast[J]. Renewable Energy, 2017, 113:11-21.
[2] Takeda, Hisashi. Short-term ensemble forecast for purchased photovoltaic generation. Solar Energy 149 (2017): 176-187.
[3] Mellit A, Kalogirou S A. Artificial intelligence techniques for photovoltaic applications: A review[J]. Progress in Energy & Combustion Science, 2008, 34(5):574-632.
[4] Bacher P, Madsen H, Nielsen H A. Online short-term solar power forecasting[J]. Solar Energy, 2009, 83(10):1772-1783.
[5] Li Y, Su Y, Shu L. An ARMAX model for forecasting the power output of a grid connected photovoltaic system[J]. Renewable Energy, 2014, 66(6):78-89.
[6] Bessa R J, Trindade A, Miranda V. Spatial-Temporal Solar Power Forecasting for Smart Grids[J]. IEEE Transactions on Industrial Informatics, 2015, 11(1):232-241.
[7] Li Y, He Y, Su Y, et al. Forecasting the daily power output of a grid-connected photovoltaic system based on multivariate adaptive regression splines[J]. Applied Energy, 2016, 180:392-401.
[8] Wang F, Mi Z, Su S, et al. Short-Term Solar Irradiance Forecasting Model Based on Artificial Neural Network Using Statistical Feature Parameters[J]. Energies, 2012, 5(5):1355-1370.
[9] Pedro H T C, Coimbra C F M. Assessment of forecasting techniques for solar power production with no exogenous inputs[J]. Solar Energy, 2012, 86(7):2017-2028.
[10] Cervone G, Clemente-Harding L, Alessandrini S, et al. Short-term photovoltaic power forecasting using Artificial Neural Networks and an Analog Ensemble[J]. Renewable Energy, 2017, 108:274-286.
[11] Rosato A, Altilio R, Araneo R, et al. Prediction in Photovoltaic Power by Neural Networks †[J]. Energies, 2017, 10(7):1003.
[12] Leva S, Dolara A, Grimaccia F, et al. Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power[J]. Mathematics & Computers in Simulation, 2017, 131(C):88-100.
[13] Das U, Tey K, Seyedmahmoudian M, et al. SVR-Based Model to Forecast PV Power Generation under Different Weather Conditions[J]. Energies, 2017, 10(7):876.
[14] Pierro M, Bucci F, De Felice M, et al. Multi-Model Ensemble for day ahead prediction of photovoltaic power generation[J]. Solar Energy, 2016, 134:132-146.
[15] Andrade J R, Bessa R J. Improving Renewable Energy Forecasting with a Grid of Numerical Weather Predictions[J]. IEEE Transactions on Sustainable Energy, 2017, PP(99):1-1.
[16] Aguiar L M, Pereira B, Lauret P, et al. Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting[J]. Renewable Energy, 2016, 97:599-610.

[17] Pierro M, Felice M D, Maggioni E, et al. Data-driven upscaling methods for regional photovoltaic power estimation and forecast using satellite and numerical weather prediction data[J]. Solar Energy, 2017, 158:1026-1038.
[18] Lohmann G M, Hammer A, Monahan A H, et al. Simulating clear-sky index increment correlations under mixed sky conditions using a fractal cloud model[J]. Solar Energy, 2017, 150:255-264.
[19] Marquez R, Coimbra C F M. Intra-hour DNI forecasting based on cloud tracking image analysis[J]. Solar Energy, 2013, 91(3):327-336.
[20] Richardson W, Krishnaswami H, Vega R, et al. A Low Cost, Edge Computing, All-Sky Imager for Cloud Tracking and Intra-Hour Irradiance Forecasting[J]. Sustainability, 2017, 9(4):482.
[21] Li M, Chu Y, Pedro H T C, et al. Quantitative evaluation of the impact of cloud transmittance and cloud velocity on the accuracy of short-term DNI forecasts[J]. Renewable Energy, 2016, 86:1362-1371.
[22] Barbieri F, Rifflart C, Ba T V, et al. Intra-Hour Cloud Tracking Based on Probability Hypothesis Density Filtering[J]. IEEE Transactions on Sustainable Energy, 2017, PP(99):1-1.
[23] Chu Y, Pedro H T C, Li M, et al. Real-time forecasting of solar irradiance ramps with smart image processing[J]. Solar Energy, 2015, 114:91-104.
[24] Chu Y, Li M, Pedro H T C, et al. Real-time prediction intervals for intra-hour DNI forecasts[J]. Renewable Energy, 2015, 83:234-244.
[25] Barbieri F, Rajakaruna S, Ghosh A. Very short-term photovoltaic power forecasting with cloud modeling: A review[J]. Renewable & Sustainable Energy Reviews, 2016, 75.
[26]Huang Y, Lu J, Liu C, et al. Comparative study of power forecasting methods for PV stations[C]. International Conference on Power System Technology. IEEE, 2010:1-6.
[27] Elsinga B, Wilfried G.J.H.M. van Sark. Short-term peer-to-peer solar forecasting in a network of photovoltaic systems[J]. Applied Energy, 2017, 206.
[28] Li L L, Cheng P, Lin H C, et al. Short-term output power forecasting of photovoltaic systems based on the deep belief net[J]. Advances in Mechanical Engineering, 2017, 9(9):168781401771598.
[29] Kleissl J. Solar energy forecasting and resource assessment[M]. Academic Press, 2013.
[30] Liu F, Li R, Li Y, et al. Takagi–Sugeno fuzzy model-based approach considering multiple weather factors for the photovoltaic power short-term forecasting[J]. Iet Renewable Power Generation, 2017, 11(10):1281-1287.
[31] Sperati S, Alessandrini S, Monache L D. An application of the ECMWF Ensemble Prediction System for short-term solar power forecasting[J]. Solar Energy, 2016, 133:437-450.
[32] Ni Q, Zhuang S, Sheng H, et al. An ensemble prediction intervals approach for short-term PV power forecasting[J]. Solar Energy, 2017, 155:1072-1083.
[33] Sheng H, Xiao J, Cheng Y, et al. Short-Term Solar Power Forecasting Based on Weighted Gaussian Process Regression[J]. IEEE Transactions on Industrial Electronics, 2017, 65(1):300-308.
[34] Aryaputera A W, Yang D, Zhao L, et al. Very short-term irradiance forecasting at unobserved locations using spatio-temporal kriging[J]. Solar Energy, 2015, 122:1266-1278.
[35] Ferlito S, Adinolfi G, Graditi G. Comparative analysis of data-driven methods online and offline trained to the forecasting of grid-connected photovoltaic plant production[J]. Applied Energy, 2017, 205:116-129.
[36] Gulin M, Pavlović T, Vašak M. A one-day-ahead photovoltaic array power production prediction with combined static and dynamic on-line correction[J]. Solar Energy, 2017, 142:49-60.
[37] Reikard G, Haupt S E, Jensen T. Forecasting ground-level irradiance over short horizons: Time series, meteorological, and time-varying models[J]. Renewable Energy, 2017.
[38] Raza M Q, Nadarajah M, Ekanayake C. On recent advances in PV output power forecast[J]. Solar Energy, 2016, 136:125-144.
[39] Antonanzas J, Osorio N, Escobar R, et al. Review of photovoltaic power forecasting[J]. Solar Energy, 2016, 136:78-111.
[40] Ozkaya A. A model of active trading by using the properties of chaos[J]. Digital Signal Processing, 2015, 39:15-21.
[41] Chen W, Sun J, Gao S, et al. Using a Single Dendritic Neuron to Forecast Tourist Arrivals to Japan[J]. Ieice Transactions on Information & Systems, 2017, E100.D(1):190-202.
[42] Iasemidis L D, Shiau D S, Chaovalitwongse W, et al. Adaptive epileptic seizure prediction system[J]. IEEE transactions on bio-medical engineering, 2003, 50(5):616-27.
[43] Schwarz R, Dvorak R, Süli Á, et al. Survey of the stability region of hypothetical habitable Trojan planets[J]. Astronomy & Astrophysics, 2007, 474(3):1023-1029.
[44] Packard N H, Crutchfield J P, Farmer J D, et al. Geometry from a Time Series[J]. Physical Review Letters, 1980, 45(9):712.
[45] Takens F. Detecting strange attractors in turbulence[M] Dynamical Systems and Turbulence, Warwick 1980. Springer Berlin Heidelberg, 1981:366-381.
[46] Lai Y C, Lerner D. Effective scaling regime for computing the correlation dimension from chaotic time series[J]. Physica D-nonlinear Phenomena, 1998, 115(1-2):1-18.
[47] Sangoyomi T B, Lall U, Abarbanel H D I. Nonlinear Dynamics of the Great Salt Lake: Dimension Estimation[J]. Water Resources Research, 1996, 32(1):149-159.
[48] Grassberger P, Procaccia I. Measuring the strangeness of strange attractors[J]. Physica D Nonlinear Phenomena, 1983, 9(1):189-208.
[49] Aguirre L A. A nonlinear correlation function for selecting the delay time in dynamical reconstructions [J]. Physics Letters

A, 1995, 203(2-3):88-94.
[50] Kraskov A, St Gbauer H, Grassberger P. Estimating mutual information.[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2004, 69(6): 066138.
[51] Kim H S, Eykholt R, Salas J D. Nonlinear dynamics, delay times, and embedding windows[M]. Elsevier Science Publishers B. V. 1999.
[52]Lorenz, Edward N. . Deterministic non-periodic flow. Journal of the Atmospheric Sciences[J].1963, 20 (2): 130–141.
[53] Safonov Leonid A., Tomer Elad; Strygin Vadim V, et al. Multifractal chaotic attractors in a system of delay-differential equations modeling road traffic. Chaos: An Interdisciplinary Journal of Nonlinear Science [J]. 2002, 12 (4): 1006-1014.
[54] Hasselblatt, Boris, Anatole Katok . A First Course in Dynamics: With a Panorama of Recent Developments[M]. Cambridge University Press, 2003.
[55] Devaney, R. L. An introduction to chaotic dynamical system, Addison Wesley Massachusetts New York,1989.
[56] Froeschlé C, Lega E, Gonczi R. Fast Lyapunov indicators. application to asteroidal motion [J]. Celestial Mechanics & Dynamical Astronomy, 1997, 67(1):41-62.
[57] Grassberger P, Procaccia I. Estimation of the Kolmogorov entropy from a chaotic signal[J]. Physical Review A, 1983, 28(4):2591-2593.
[58] Iasemidis L D, Sackellares J C, Zaveri H P, et al. Phase space topography and the Lyapunov exponent of electrocorticograms in partial seizures[J]. Brain Topography, 1990, 2(3):187-201.
[59] Wolf A, Swift J B, Swinney H L, et al. Determining Lyapunov exponents from a time series[J]. Physica D Nonlinear Phenomena, 1985, 16(3):285-317.
[60] Zhu-ping G. The Calculating Method of the Average Period of Chaotic Time Series[J]. Systems engineering, 2010, 28(12): 111-113.