Journal of Renewable and Sustainable Energy ARTICLE scitation.org/journal/rse

Clustering-based computation of degradation rate for photovoltaic systems

Cite as: J. Renewable Sustainable Energy **11**, 014701 (2019); doi: 10.1063/1.5042688
Submitted: 04 June 2018 · Accepted: 24 November 2018 · Published Online: 08 January 2019

View Online Export Citation CrossMark

Parveen Bhola and Saurabh Bhardwaj

AFFILIATIONS

Electrical & Instrumentation Engineering Department, TIET, Patiala, India

ABSTRACT

For effective utilization of solar energy, performance monitoring of photovoltaic (PV) systems is required. Two important research goals are to maximize the power output from PV systems and further reduce the economic losses. This paper proposes a model using a clustering-based technique to evaluate the degradation of PV panels with different topologies. Here, the performance ratio (PR) of the PV panels is estimated without physical inspection on-site, making the proposed model beneficial for real-time estimation of the PR and in turn for more robust forecasting of the PV power output. The present work utilizes the segmental K-means clustering technique to obtain clusters of input meteorological data sharing similar features. Various forms of meteorological data, including temperature, relative humidity, wind speed, dew point, solar radiation, and sunshine hours, are given as the input, and solar power data are the output of the proposed model. The proposed model calculates the degradation in output solar power in terms of PR for panels with three different topologies, namely, amorphous silicon (a-Si), polycrystalline silicon (p-Si), and heterojunction with an intrinsic thin layer (HIT), over a period of three years. The degradation rate produced by a-Si technology was lowest, and it was highest for HIT technology. The results obtained showed good agreement with the standard method used for performance evaluation in a similar earlier study. The proposed model has the advantage over other methods that real-time estimation is possible, as this method does not require physical inspection and imaging, which is essential in other techniques.

I. INTRODUCTION

The Republic of India has huge potential in solar energy. With the help of technological advances, India has already achieved the milestone of 20 GW solar power[1] and its target is to further increase this to 100 GW by the end of 2022.

At present, photovoltaic (PV) technology is becoming more widespread globally. Therefore, performance monitoring of PV systems is a major area of interest for researchers. A key metric of performance is resistance to degradation, which is a gradual decrease in power output over the years. Furthermore, as the installation cost of PV systems is higher than that of conventional systems, the design of a solar PV system requires an estimate of energy output at the site where the system is to be installed. Knowledge of the degradation rate is helpful for the accurate forecasting of energy output.

Moreover, the continuous monitoring of the degradation of PV systems in terms of the performance ratio (PR) is useful for the correction of underperforming systems and helpful in reducing the economic losses due to operational problems. The majority of techniques calculate the degradation of PV systems by physical inspection and by capturing images of PV panels. This process is time-consuming and costly and cannot be used for the real-time analysis of degradation. Therefore, keeping in view the above requirements, a clustering-based technique for the estimation of the degradation rate is proposed here.

A comprehensive body of work on degradation is available in the literature. The degradation in performance can occur at the cell, module, or array level. At the cell level, the main factors responsible for degradation include the temperature, humidity, precipitation, snow, dust, and solar radiation, while at the array level, module mismatch and shading contribute to the degradation. Corrosion and discoloration are also major sources of degradation.[2]

An experimental study collected data from 57 crystalline silicon modules in five different climatic zones of India. The study concluded that discoloration-induced degradation is more prominent in hot and dry climate zones, while corrosion is more important in hot and humid zones.[3,4] It was also concluded that cold climatic zones show the least degradation.

The capacity utilization factor (CUF), yield, and PR are the most common parameters to evaluate the performance of PV systems in industry. PR is advantageous over other commonly used parameters as it indicates the actual energy delivered by the plant, instead of the theoretical value, under a given insolation and climate condition.[5] PR is an indicator of losses resulting from the cell/array mismatch, shading, inverter problems, module temperature, etc. Various statistical methods are reported in the literature to calculate the PR. The use of linear regression (LR), classical seasonal decomposition (CSD), and auto-regressive integrated moving average (ARIMA) is described in Ref. 6. The most common method for calculating the degradation is LR, which aims to minimize the sum of squared residuals. However, this method is very sensitive to outliers and seasonal variation and thus has a very large uncertainty. To overcome this limitation and extract the trends from PV time series data, the CSD method is used.[7] In this method, the seasonal component of every month is extracted using centered moving average computation, based upon the assumption that seasonal components remain stable year after year. As the LR and CSD methods are fitted to a fixed model, they are unable to capture some of the important features of solar energy time series due to notable autocorrelations in the model residuals. To overcome this limitation as well as to deal with seasonal variation, random errors, and outliers, the ARIMA model is preferred. This is a combination of two statistical operations, namely, autoregression and moving average. Another statistical decomposition method based on locally weighted scatterplot smoothing (LOESS) is proposed in Ref. 8. In this, the performance of nine different PV panels was evaluated at the Solar Energy Institute of Singapore. It was found that mono-crystalline technology performed better than multi-crystalline and amorphous silicon technologies. The copper–indium–gallium–selenide (CIGS) module experienced the highest degradation rate of 6% per year. The advantage of using LOESS is that it provides robust estimates for the seasonal components and trends as compared to CSD or ARIMA.

The robust principal component analysis (RPCA) method has also been used for calculating the PR.[9] That study was performed on three different technologies, including monocrystalline silicon, multi-crystalline, and hetero-junction with an intrinsic thin layer (HIT). The results were reasonably accurate for thin-film technology when monthly PR was calculated for eight years of plant data, situated at different locations in Cyprus.

Researchers have generally used the following methods to evaluate the degradation rate of PV modules:[10] module current–voltage (I-V) measurement, metered raw kWh, PR, and performance index. Among these four methods, the I-V method was found to be the best for degradation rate computation in Ref. 10. In Ref. 11, the degradation rate was computed with the help of the regression and year-to-year (YOY) methods and a comparative analysis of the two was also presented. From the analysis, it was clear that the regression method requires filtering, as it is sensitive to outliers, while the YOY method does not require filtering and is also insensitive to outliers but requires multiple years of data. The hybrid combination of YOY with the clear sky model was also used in Ref. 12. This model provides the liberty to use the clear sky irradiance data instead of site sensor data. This method provides reliable degradation rate calculation even when sensor drift, data drift, and soiling are present. When compared to other methods, the results produced by this method showed the lowest uncertainty in the value of the degradation rate. Rooftop PV systems placed at the National Renewable Energy Laboratory (NREL), USA, were investigated for degradation after 20 years of operation.[13] Two arrays of mono-Si technology showed a degradation rate of 0.8% per year calculated through historical mean values of PR. In a similar study, the degradation in PR of 90 mono-crystalline PV modules was calculated, which were 22 years old and installed at the rooftop of the guesthouse of the National Institute of Solar Energy (NISE), Gurgaon, India.[14] Here, the average degradation rate was found to be 1.9% per year with a maximum value of 4.1% per year and a minimum reported value of 0.3% per year. Elsewhere, another study of thermal degradation was conducted using meteorological data and the effective temperature.[15] The degradation effect in opaque and semitransparent PV modules was presented in Ref. 16. The degradation rate in opaque modules was higher than that in semitransparent modules.

The present work proposes a machine learning-based technique to estimate the degradation of PR, which does not require physical inspection. This work proposes a technique by which real-time estimation of the PR is possible. A clustering-based approach for calculating the PR is proposed. The study is performed on three different topologies of PVs, namely, amorphous silicon (a-Si), polycrystalline silicon (p-Si), and hetero-junction with an intrinsic thin layer (HIT). Monthly PRs for all three topologies are calculated for the years 2010–2012, and finally, the degradation performance for the year 2012 with respect to 2010 is obtained using clustering-based computation of degradation rate (CCDR). The results of the present work show a good agreement with the study in Ref. 17, which assessed the performance of different PV technologies under similar outdoor conditions using the same types of data. According to the present study, the degradation rate of HIT technology is highest, followed by p-Si and then a-Si technology.

This paper is organized as follows: Sec. II covers the details of the site location and the data used in the present work. Section III presents the performance analysis parameters used for performance evaluation of the PV systems, along with the clustering technique used to obtain clusters of meteorological and solar power data. Section IV explains the methodology used for computation of the CCDR, and Sec. V is dedicated to the results and conclusions.

II. SITE AND DATA DESCRIPTION

The National Institute of Solar Energy (NISE), a pioneer institute in the field of solar energy situated in Gurgaon city, was the site considered in this study. It is located in close proximity to the country's capital with the geographical location of 28° 37′ N and 77° 04′ E. As per the Bureau of Indian Standards, it falls in the "composite" climate category and experiences very dry summers, a wet rainy season, and cold winters. The daily average temperature in summer, from April to June, varies from 46 °C to 33 °C with a mean temperature of 41 °C. During the winter, the average daily temperature is below 26 °C with the minimum registered in the months of December and January (9 °C during the day). The relative humidity is high with respect to the annual average for India, varying from 42% to 70% during the rainy season from July to August. The wind speed is light and moderate with an annual average of 1.5 m/s. The monthly global horizontal solar radiation varies from 2.6 KW_P/m^2 per day in January to 6.1 KW_P/m^2 per day in May. The experimental setup installed at NISE is shown in Fig. 1. PV panels based on the three different technologies, namely, a-Si, p-Si, and HIT, are shown in Fig. 1(a). The specification of the modules for all three topologies is provided in Ref. 17. The nominal rating of the p-Si array is 1.6 KW_p, that of the HIT array is 1.6 8 KW_p, and that of the a-Si array is 1.2 KW_p. The a-Si PV array consists of 16 modules of 75 W_p each, the HIT array comprises 8 modules of 210 W_p each, and the p-Si array comprises 10 modules having a rated value of 160 W_p. The I-V performance data of each PV array are taken every 10 min. The analyzer identifies the maximum power P_{Max}, maximum voltage V_{Max}, and maximum current I_{Max}, which are stored in the data logger. Therefore, the solar power data of the three technologies are stored every 10 min. These data are further processed to obtain the hourly and daily database for solar power. The meteorological data are obtained from the weather monitoring station at NISE, Gurgaon, as shown in Fig. 1(b). The parameters recorded at the weather station are ambient temperature, relative humidity, atmospheric pressure, wind speed, dew point, wind direction, and solar radiation. The database is available for the three years from 2010 to 2012 for every minute. The sunshine data are not available in the database; instead, they were calculated from the solar radiation data. The database records for every minute were processed to obtain the hourly and daily database.

III. PERFORMANCE ANALYSIS OF THE PV SYSTEM USING K-MEANS CLUSTERING

A. Performance of the PV system

The performance monitoring parameters used for performance evaluation of a PV system are described in the literature.[18,19] The performance is measured in terms of energy produced, system losses, PR, and various yields. The DC energy produced by the PV system on a daily basis is given by Eq. (1), whereas the monthly DC energy produced by the PV system is given by Eq. (2)

$$E_{dc}, d = \sum_{t=1}^{t=T_{rp}} V_{dc}*I_{dc}*T_r, \tag{1}$$

$$E_{dc}, m = \sum_{d=1}^{N} E_{dc}, d, \tag{2}$$

where T_{rp} is the reporting time during which sunlight is available and T_r is the recording time; N is the number of days in a given month. V_{dc} and I_{dc} are the open circuit voltage and short circuit current produced by the panel. The DC energy produced by the PV system is converted to AC with the help of an inverter. The power recorded at the output terminal of the inverter represents the energy generated or delivered to the grid, given as follows:

$$E_{ac}, d = \sum_{t=1}^{t=T_{rp}} V_{ac}*I_{ac}*T_r, \tag{3}$$

where T_{rp} and T_r are the reporting and recording time; V_{ac} and I_{ac} are the AC voltage and AC current at the output terminals of the inverter. The array yield (Y_a) is equal to the DC energy produced by the PV array when it is operating at the

Figure 1(a) Experimental setup

Figure 1(b) Weather station

FIG. 1. Experimental setup at NISE, Gurgaon.

rated power. The mathematical equation for the array yield is given as Eq. (4). Here, E_{dc} is the DC energy produced by the PV panel. $P_{mp(rated)}$ is the rated power of the PV panel

$$Y_a, d = \frac{E_{dc}, d}{P_{mp(rated)}}. \tag{4}$$

When the energy produced by the array is expressed in terms of AC energy, operating at the rated power, the final yield (Y_f) is calculated as follows:

$$Y_f, d = \frac{E_{ac}, d}{P_{mp(rated)}}. \tag{5}$$

Although the final yield or specific yield (Y_f) is an important parameter used for performance monitoring,[20] it cannot be used for comparing two PV plants located in two different regions due to significant variance in the values of insolation. CUF is another important parameter used for performance monitoring, defined as the ratio of the maximum output of the PV plant to the maximum output under ideal conditions, given as follows:

$$CUF = \frac{Y_f, a}{24*365} = \frac{E_{dc}, d}{P_{mp(rated)}*8760}. \tag{6}$$

CUF does not reflect the actual performance of the PV plant as it does not account for factors like environmental effects. It is expressed as the ratio of the actual annual energy output (E_{dc}, d) of the PV system to the amount of energy that would be generated by the PV system if it is operated at full rated power for 24 h per day in a given year. Therefore, PR is commonly used to indicate the actual energy delivered by the plant, instead of the theoretical value, under a given insolation and climate condition.[21] PR is defined as the ratio of the final yield (Y_a) to the reference yield (Y_r), given as follows:

$$PR = Y_f/Y_r = \frac{E_{dc}, d}{P_{mp(rated)}} \Big/ \frac{H}{G_{(STC)}}. \tag{7}$$

PR is a unitless quantity, and its value lies between zero and one. According to the European PV standard, the values of PR between 0.80 and 0.85 are considered as good, while the values below 0.75 reflect poor performance over time. The present work uses the PR, which is calculated by the clustering technique, to estimate the degradation shown by the three technologies.

B. K-means clustering

K-means clustering is a widely used method for clustering data. We have used this method to extract clusters of meteorological data sharing similar features. This is an unsupervised learning technique to form groups or patterns of given data points in such a way that patterns in the same cluster are similar in nature, while patterns belonging to other clusters are different.[22] The formation of clusters is carried out using the centroid technique. In this technique, a centroid is defined for each cluster, and the objective function is minimized as follows:

$$E = \sum_{j=1}^{K} \sum_{p \in C_i} dist(p, C_i)^2, \tag{8}$$

where E is the sum of squared error for the entire set of objects in the dataset, p represents the positions of all the objects in space, and C_i is the centroid of the cluster. The selection of the number of clusters to be taken is an important parameter. In our case, the number of clusters, K, is based on the seasonal variation at the plant location. Once the number of clusters is known, the clustering technique is applied to the meteorological dataset, using the daily average values, in order to use these clusters as input variables, as required in the proposed model.

IV. METHODOLOGY

A. Clustering-based computation of degradation rate (CCDR)

The present work proposes a new methodology using the unsupervised clustering technique to compute the PR and degradation rate of PV modules. The method requires knowledge of previous meteorological data but has no need of any physical inspection of the modules on-site. The degradation rate is the decline in power output for the same input conditions over a time period and is a crucial parameter that reflects the performance of the plant. The present methodology applies the powerful pattern-recognition capability of clustering to obtain clusters of similar input conditions. The technique is applied to the meteorological data of three years (2010–2012). The complete procedure for the CCDR technique is shown in Fig. 2. In the proposed technique, the pre-processed data are used for the degradation rate calculation. The data consist of meteorological parameters as the input vector and solar power as the output. After pre-processing, the next step is to find patterns/clusters of similar weather conditions in the input data. The meteorological data in a specific cluster provide approximately uniform input conditions for all three years, i.e., 2010–2012, for all the PV topologies under investigation. In the present work, the classical K-means clustering method is used to obtain these clusters in the data. The value of K is precisely chosen in such a way that it covers the seasonal variation during a year. The principal seasons at the plant location are summer, rainy season, autumn, winter, and finally spring.

Experiments were conducted in which a range of values of K was considered. The optimized value of K was found to be 12, which covered all the seasonal variability at the plant. It is important to mention that the clustering algorithm was applied to the whole dataset for the years 2010–2012, but to calculate the degradation rate, similar input patterns of each month are required. Therefore, considering the seasonal variability, we have further subdivided the clusters according to their month for all three years. This step arranges the whole dataset on a monthly basis, with each monthly group (also termed common cluster) showing the same weather conditions. Once the monthly groups of data are known, the corresponding solar power is determined by taking the averages of the common clusters for each year separately. This process provides the average solar power of each year for similar input conditions on a monthly basis. This makes it straightforward to calculate the

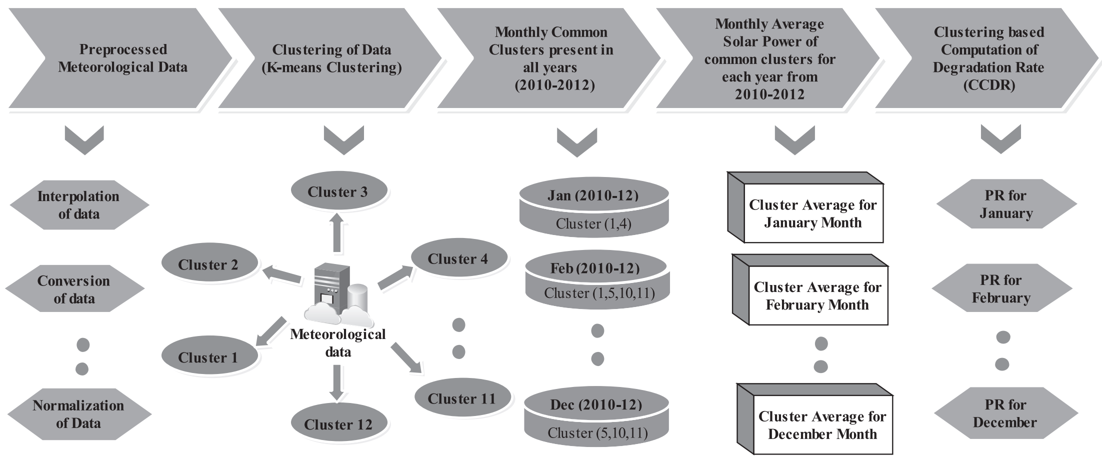

FIG. 2. Clustering-based computation of degradation rate.

change in power and hence the degradation rate. For ease of understanding, the proposed methodology is demonstrated for a subset of data. The complete methodology is summarized in the following steps:

Steps

Step 1: Pre-processing of meteorological data: The meteorological data are processed through the following substeps before the clustering algorithm is applied:

(i) Interpolation of the missing values in the meteorological data:
As the obtained input dataset is corrupted, a few of its values are missing. Therefore, the interpolation technique is applied to obtain the missing values. The new data point for each of the missing values is calculated, and the corrupted data points are removed from the dataset.

(ii) Conversion of data from minute to daily format:
The per-minute meteorological data are obtained from the plant, but as per the requirement of the proposed model, the data are converted into the hourly and finally into the daily format, by a simple averaging technique.

(iii) Normalization of data:
The input vectors of the meteorological data have different ranges, which are difficult to model. So, to constrain the data into the same range, we next normalize the data by using the max-min normalization (also known as feature scaling) to restrict the input parameters in the range [0 1]. The feature scaling is performed using Eq. (9). Here, X' is the normalized value of the data and X is the real value of the meteorological parameters

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}, \quad (9)$$

where X_{min}, X_{max} = minimum and maximum values of the attribute.

Step 2: Clustering of input meteorological data: The ambient temperature, humidity, wind speed, pressure, dew point, irradiance, and sunshine hours are used as the input vector. To find similar patterns in the input data, the segmental K-means clustering algorithm is used. After the application of the clustering algorithm, the data are allocated to the clusters shown in Table I.

TABLE I. Input meteorological data with the assigned clusters.

Time	Temp	Humidity	Wind speed	Pressure	Dew Pt.	Irradiance	Sunshine hours	Cluster
01/01/2010	12.63	69.81	0.91	986.10	6.44	137.69	7	5
02/01/2010	9.50	91.15	0.94	988.76	8.08	50.25	3	10
06/01/2010	10.20	79.9	0.55	985.10	6.47	126.14	7	6
01/01/2011	10.77	77.33	1.57	984.43	6.73	120.31	6	10
02/01/2011	9.92	70.36	1.78	984.54	4.68	109.68	6	7
05/01/2010	10.20	79.9	0.55	985.10	6.47	126.14	7	6
01/01/2012	14.76	77.08	0.66	984.31	10.7	33.58	2	5
02/01/2012	13.91	81.42	0.46	985.53	10.6	56.88	4	10
03/01/2012	10.20	79.9	0.55	985.10	6.47	126.14	7	4
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

TABLE II. Monthly arrangement of common clusters.

Time	Temp	Humidity	Wind speed	Pressure	Dew Pt.	Irradiance	Sunshine hours	Common cluster
02/01/2010	9.50	91.15	0.94	988.76	8.08	50.25	6	10
07/01/2010	10.43	84.29	0.42	987.09	7.55	88.09	6	10
08/01/2011	10.77	77.33	1.57	984.43	6.73	120.31	6	10
09/01/2011	9.92	70.36	1.78	984.54	4.68	109.68	6	10
05/01/2012	14.76	77.08	0.66	984.31	10.7	73.58	5	10
06/01/2012	13.91	81.42	0.46	985.53	10.6	56.88	6	10
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Step 3: Month-wise arrangement of cluster data: After arranging the input data into clusters, the next step is to organize the clustered data month-wise into common clusters, as shown in Table II.

Step 4: Monthly average solar power: After month-wise allocation, the next step is to process the power outputs corresponding to the input meteorological vectors. Here, the monthly cluster data are grouped according to their year. Then, the monthly averaged output power of each common cluster is calculated for each year. The monthly average solar power of the a-Si technology for the years 2010–2012 is shown in Table III.

Step 5: Performance ratio calculation: The next step is to calculate the PR from the monthly power output of the respective years calculated in step 4 as follows:

$$Performance\ Ratio\ (PR) = \frac{Monthly\ Power\ in\ 2011}{Monthly\ Power\ in\ 2010}. \tag{10}$$

The resulting monthly PRs are shown in Table IV.

Step 6: Clustering-based Computation of Degradation Rate (CCDR): Finally, the PRs computed in step 5 are used to estimate the degradation rate for the three different technology modules. The degradation rates are computed with the help of the standard least-squares regression method. The slope of the regression line for the monthly PR determines the degradation rate, as shown in Fig. 3.

V. EXPERIMENTAL RESULTS

The performance of the proposed model is validated by computing the PR along with the degradation rate for the a-Si, p-Si, and HIT technology panels. The monthly obtained average solar power in the year 2010–2011 for a-Si, p-Si, and HIT is shown in Fig. 4. The model estimates the PR for three years, i.e., from 2010 to 2012; the values of PR for the three technology panels are presented in Table V. The performance of the three technologies is shown in Fig. 5 along with the comparative analysis among them.

- The a-Si technology showed consistently less values of PR in the year 2012 when compared to the power produced in 2010, as indicated by PR values consistently less than 1. The PR (2012/2010) values varied between 0.80 and 0.93, with the minimum value in the month of July and the maximum value in October, as shown in Fig. 5(a) and Table V. The average value of PR (2011/2010) was 0.89, which dropped to 0.87 for (2012/2010). Further, the mean degradation rate reported was 0.85% per year for the a-Si technology, calculated by the regression method, as shown in Table VI.
- The p-Si technology achieved a better PR as compared to the a-Si technology. The PR (2012/2010) values varied between 0.87 and 0.96, with the minimum in the month of July and the maximum value in December, as shown in Fig. 5(b) and Table V. The average value reported for PR (2011/2010) was 0.95, which dropped to 0.91 for (2012/2010), representing a lower extent of degradation compared to the a-Si technology. The mean degradation rate for p-Si technology was found to be 0.95% per year, which is slightly higher than the a-Si technology (Table VI).

TABLE III. Monthly average solar power of a-Si technology.

Months	2010	2011	2012
January	119.65	113.72	102.18
February	193.82	190.39	176.50
March	307.41	276.92	255.36
April	273.73	251.51	241.36
May	266.86	243.70	236.00
June	265.54	261.97	236.23
July	240.29	229.52	194.68
August	156.98	152.72	145.49
September	192.05	189.51	173.01
October	251.31	242.88	235.76
November	201.74	189.56	182.35
December	166.68	149.46	144.65

TABLE IV. Monthly performance ratios.

Months	PR(11/10)	PR(12/11)	PR(12/10)
January	0.95	0.89	0.85
February	0.98	0.92	0.91
March	0.90	0.92	0.83
April	0.91	0.95	0.88
May	0.91	0.96	0.88
June	0.98	0.90	0.88
July	0.95	0.84	0.81
August	0.97	0.95	0.92
September	0.98	0.91	0.90
October	0.96	0.97	0.93
November	0.93	0.96	0.90
December	0.89	0.96	0.86

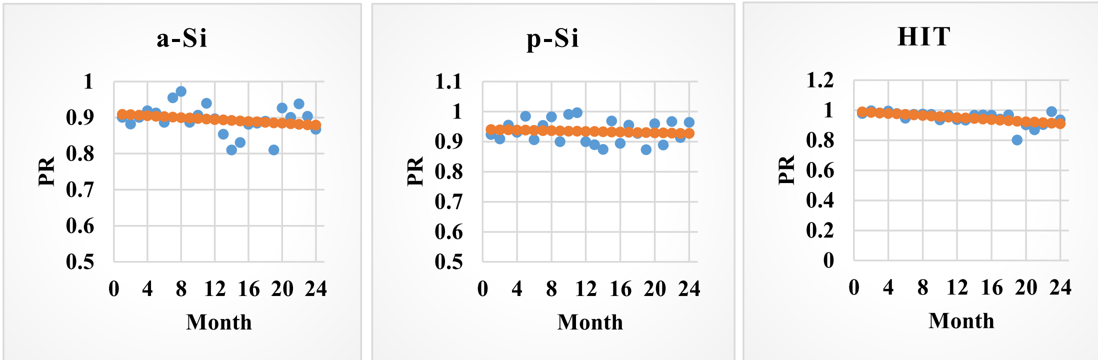

FIG. 3. Degradation rate of a-Si, p-Si, and HIT technology PV modules.

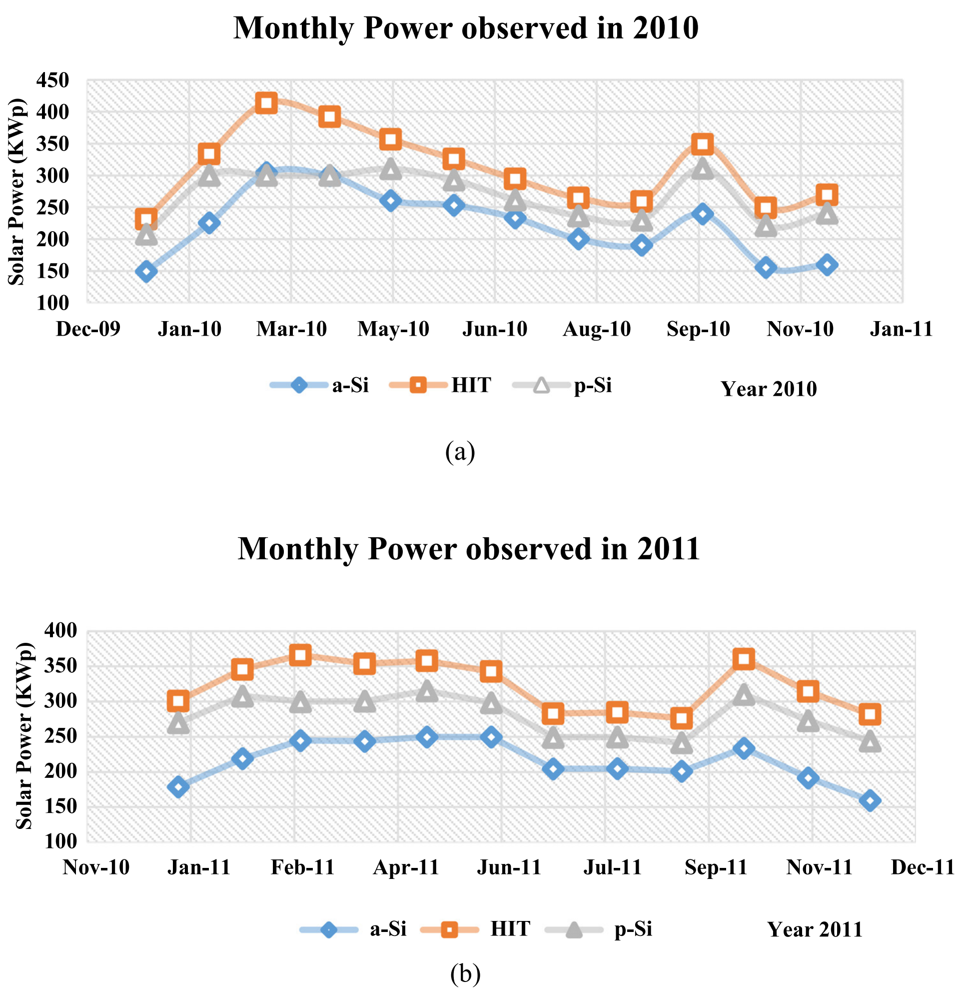

FIG. 4. Monthly average solar power obtained by a-Si, p-Si, and HIT during years 2010–2011. (a) Monthly average power obtained by the three technologies in 2010. (b) Monthly average power obtained by the three technologies in 2011.

- The HIT technology achieved PR (2012/2010) values between 0.80 and 0.97, with the minimum value in July and the maximum value reported in November, as shown in Fig. 4(c) and Table V. The average value for PR (2011/2010) was found to be 0.96, which decreased to 0.92 for (2012/2010). The HIT technology showed the highest degradation rate, with a value of 1.1% per year as shown in Table VI.

TABLE V. Performance of a-Si, p-Si, and HIT technology solar panels during 2010–2012.

	PR for a-Si technology		PR for p-Si technology		PR for HIT technology	
Months	PR (2011/2010)	PR (2012/2010)	PR (2011/2010)	PR (2012/2010)	PR (2011/2010)	PR (2012/2010)
January	0.95	0.85	0.92	0.89	0.97	0.93
February	0.98	0.91	0.90	0.87	0.99	0.96
March	0.90	0.83	0.99	0.96	0.98	0.96
April	0.91	0.88	0.93	0.89	0.99	0.96
May	0.91	0.88	0.98	0.95	0.97	0.93
June	0.98	0.88	0.99	0.92	0.94	0.96
July	0.95	0.81	0.95	0.87	0.97	0.80
August	0.97	0.92	0.98	0.95	0.97	0.90
September	0.98	0.90	0.90	0.88	0.97	0.86
October	0.96	0.93	0.99	0.96	0.93	0.90
November	0.93	0.90	0.99	0.91	0.96	0.98
December	0.89	0.86	0.98	0.96	0.93	0.93

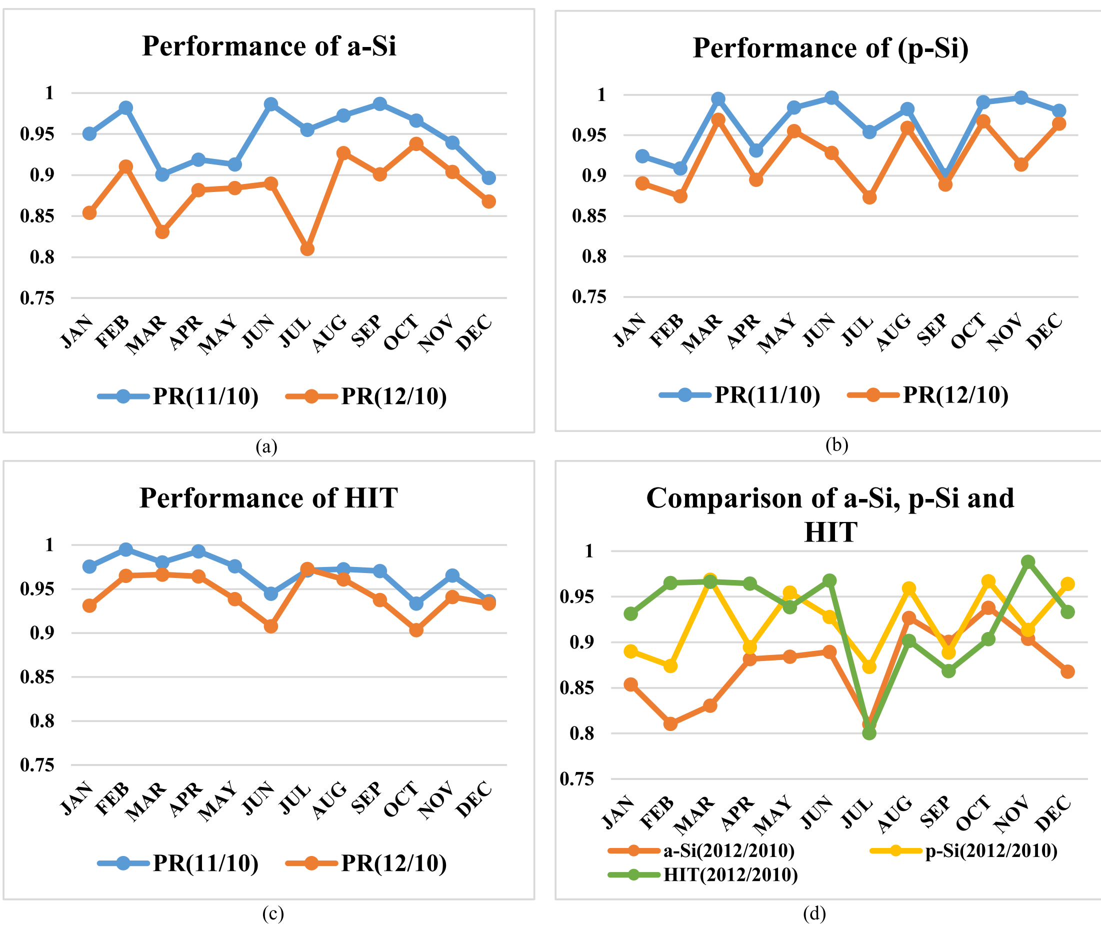

FIG. 5. Performance ratios of a-Si, p-Si, and HIT solar panels and their comparison during 2010–2012. (a) Performance ratio of a-Si. (b) Performance ratio of p-Si. (c) Performance ratio of HIT. (d) Performance comparison of a-Si, p-Si, and HIT technologies.

TABLE VI. Comparison of results with different methods.

Time-frame	Technology	Method	Degradation rate (%/Year)	Location
Monthly	Mono-Silicon	PR-based[10]	0.84	ASU-PRL, USA
	Poly-Silicon		0.95	
	HIT		0.95	
Monthly	c-Si	PR-LLS[23]	0.63	Cyprus
	HIT		2.03	
Monthly	a-Si	CCDR	0.85	NISE Gurgaon
	p-Si		0.95	
	HIT		1.1	

- A comparison of the three technologies is shown in Fig. 5(d) and Table V. The HIT technology performed better than a-Si and p-Si technologies for first half of the year except May month. The p-Si technology performed better for second half of the year except November.
- The highest degradation rate was shown by the HIT technology, with a value of 1.1% per year, followed by p-Si and a-Si technologies. A comparison with other methods for degradation rate estimation is shown in Table VI.

VI. CONCLUSION

The present article has proposed a clustering-based model to estimate the degradation rate of solar panels. The key feature of the proposed model is that it does not require any physical inspection of the panels on-site to calculate the performance ratio of PV panels, and so, it can be used for the real-time estimation of degradation. The degradation in performance for three PV technologies, namely, polycrystalline silicon, amorphous silicon, and hetero-junction with intrinsic thin-layer silicon, was estimated with the help of the model, and the results obtained are in close proximity with the results produced by other methods, as shown in Table VI. It is also summarized that the proposed model has less complexity and is faster than earlier methods. Further, from the experimental results, the following additional conclusions can be drawn.

- HIT technology panels show the highest degradation rate, followed by p-Si and a-Si technologies, in the region of study.
- The a-Si technology has the lowest degradation rate in the region of study. The performance ratio of a-Si was lower than that for HIT technology and p-Si except in the autumn season from August to November.

ACKNOWLEDGMENTS

The authors are thankful to the National Institute of Solar Energy (NISE), Gurgaon, for providing the resources and data that are used for this study.

REFERENCES

[1]See https://2rjrmf33rccw3lrxgi3x82yy-wpengine.netdna-ssl.com/wpcontent/uploads/2018/02/MercomIndiaSolarMarketUpdateQ42017ExecSumm.pdf for the contribution of solar power in India's energy portfolio; accessed 20 March 2018.
[2]A. Ndiaye, A. Charki, A. Kobi, C. M. F. Kébé, P. A. Ndiaye, and V. Sambou, Sol. Energy **96**, 140 (2013).
[3]R. Dubey, S. Chattopadhyay, V. Kuthanazhi, J. J. John, B. M. Arora, A. Kottantharayil, C. S. Solanki, V. Kuber, and J. Vasi, in *All-India Survey of Photovoltaic Module Degradation* (2013), p. 271.
[4]S. Chattopadhyay, R. Dubey, V. Kuthanazhi, J. J. John, C. S. Solanki, A. Kottantharayil, B. M. Arora, K. L. Narasimhan, V. Kuber, J. Vasi, A. Kumar, and O. S. Sastry, IEEE J. Photovoltaics **4**, 1470 (2014).
[5]See http://files.sma.de/dl/7680/Perfratio-TI-en-11.pdf for detail study of performance ratio; accessed 20 April 2018.
[6]A. Phinikarides, N. Kindyni, G. Makrides, and G. E. Georghiou, Renewable Sustainable Energy Rev. **40**, 143 (2014).
[7]A. Phinikarides, G. Makrides, and G. E. Georghiou, in *32nd EU-PVSEC* (2016), p. 1754.
[8]J. Y. Ye, T. Reindl, A. G. Aberle, and T. M. Walsh, IEEE J. Photovoltaics **4**, 1288 (2014).
[9]A. Kyprianou, A. Phinikarides, G. Makrides, and G. E. Georghiou, in *29th EU-PVSEC* (2014), Vol. 58, p. 2939.
[10]S. Shrestha and G. Tamizhmani, in *IEEE 42nd Photovoltaic Specialists Conference (PVSC)* (2015), p. 8.
[11]D. C. Jordan, M. G. Deceglie, and S. R. Kurtz, in *Conference Recordings of the IEEE Photovoltaic Specialists Conference* (2016), p. 273.
[12]D. C. Jordan, C. Deline, S. R. Kurtz, G. M. Kimball, and M. Anderson, IEEE J. Photovoltaics **8**, 525 (2018).
[13]D. C. Jordan, B. Sekulic, B. Marion, and S. R. Kurtz, IEEE J. Photovoltaics **5**, 744 (2015).
[14]P. Rajput, G. N. Tiwari, O. S. Sastry, B. Bora, and V. Sharma, Sol. Energy **135**, 786 (2016).
[15]P. Rajput, O. S. Sastry, and G. N. Tiwari, Sol. Energy **144**, 267 (2017).
[16]P. Rajput, G. N. Tiwari, and O. S. Sastry, Sol. Energy **139**, 569 (2016).
[17]V. Sharma, A. Kumar, O. S. Sastry, and S. S. Chandel, Energy **58**, 511 (2013).
[18]C. Cañete, J. Carretero, and M. Sidrach-de-Cardona, Energy **65**, 295 (2014).
[19]D. Micheli, S. Alessandrini, R. Radu, and I. Casula, Energy Convers. Manage. **80**, 436 (2014).
[20]K. Padmavathi and S. A. Daniel, Energy Sustainable Dev. **17**, 615 (2013).
[21]I. Santiago, Energy Convers. Manage. **81**, 144 (2014).
[22]A. Likas, N. Vlassis, and J. Verbeek, Pattern Recognit. **36**, 451 (2003).
[23]A. Phinikarides, G. Makrides, and G. E. Georghiou, in *28th EU-PVSEC* (2013), p. 3973.