New PV Performance Loss Methodology Applying a Self-Regulated Multistep Algorithm

Sascha Lindig, Atse Louwen, David Moser, and Marko Topic

***Abstract*—Analyses of performance loss rates in photovoltaic (PV) systems are not yet standardized, and are typically carried out by a linear regression of the evolution of a certain performance metric (e.g., performance ratio, yield, etc.) over time. In this article, we propose a novel methodology of advanced PV system performance loss rate modeling applying a self-regulated multistep algorithm. The developed algorithm automatically detects the number and positions of breakpoints in nonlinear performance time series and divides the performance trend into an adequate number of linear segments. Instead of calculating one linear performance loss rate, given in percentage per year, as is common practice, multiple linear performance loss values are determined, depending on the trend of the time series and subsequently the number of breakpoints. The algorithm is fully automated. We have applied our methodology on data of an experimental PV installation in Bolzano/Italy, which consists of 26 different PV systems. The overall linear performance loss rate of the facility's crystalline silicon systems is between −0.5 and −1.3%/year, whereas the thin-film PV systems experience values between −0.6 and −2.4%/year. Based on our results, the algorithm appears to be stable and accurate. The methodology is to be used as a fast and automated check of PV systems in operation to detect anomalies affecting performance (in an early stage). By building up a large database of detected issues in the field this algorithm will enable us to better understand the performance evolution of different PV system types in varying climates.**

***Index Terms*—Degradation models, photovoltaic (PV) performance, PV systems, performance loss rate, PV ageing.**

I. Introduction

RECENT studies show that worldwide installed photovoltaic (PV) capacity reached 580.16 GW by the end of 2019 [1]. This accounts for almost 23% of the total installed capacity of all renewable energy sources. During 2019 alone, an added amount of 97.1 GW was recorded, which equals more than 50% of the total addition of renewable capacity in that year. These numbers show that PV is one of the major drivers of a transition towards sustainable electricity generation.

Solar PV is no longer just a niche technology for specific applications or conditions but has established itself as a sustainable and competitive source of electricity. Because of this, it is expected that PV systems have a reliable, long-term operation, and that their operational behavior is well understood. PV systems are subject to permanent cycles of different potentially performance impairing influences, such as temperature, humidity, solar irradiation, soiling, wind or snow loads, and other mechanical stresses. The presence and interaction of these influences can lead to reversible or irreversible damage, which decreases the power output of a PV system and might even cause safety issues [2]. Most of the irreversible failures can be classified as a degradation of the PV module/cell and are subdivided into specific degradation modes. A thorough characterization of PV modules is necessary to detect and identify specific degradation modes [2]–[4]. Whether PV systems' exhibit performance degradation, and to what extend seems to be, among other things, climate dependent. Recent studies have shown that degradation rates can vary by up to 50% depending on the prevailing climate [5].

One way to assess the health status of a PV system is to calculate the performance loss rate (PLR). The PLR is a parameter, which assesses the performance evolution based on a performance metric. These are either electrical parameters, empirical metrics, or normalized values. Usually, when assessing the PLR of PV systems, a linear performance behavior is assumed, and many current studies primarily apply this approach [6]–[9]. This representation might be sufficient to get a first assessment of how a system is operating, but in reality the development of performance is more likely to be nonlinear. Köntges *et al.* [10] reported that power losses can take different shapes such as linear, stepwise, exponential or saturating curves while Virtuani *et al.* [11] reported polynomial degradation patterns of PV modules after 35 years of field exposure. Nonlinear behavior of performance evolution is first and foremost expected at the beginning of life, and during the wear-out phase of the PV plant, characterized by the well-established performance bathtub curve for PV systems [4]. Nevertheless, a nonlinear trend line, which would provide very detailed information about the system performance, is difficult to assess and evaluate.

In this article, we present an adapted PLR methodology, which aims to provide a compromise between giving a more detailed description of the performance, while still being easily understandable and applicable. This so-called multistep

Manuscript received January 26, 2021; revised March 29, 2021 and April 28, 2021; accepted May 2, 2021. Date of publication May 24, 2021; date of current version June 21, 2021. This work was supported in part by the European Union's Horizon 2020 Programme under Grant 721452 - H2020-MSCA-ITN-2016 and in part by the Slovenian Research Funding Agency under the research program P2-0197. *(Corresponding author: Sascha Lindig.)*

Sascha Lindig is with the Institute for Renewable Energy, EURAC Research, 39100 Bolzano, Italy, and also with the Faculty of Engineering, University of Ljubljana, 1000 Ljubljana, Slovenia (e-mail: sascha.lindig@eurac.edu).

Atse Louwen and David Moser are with the Institute for Renewable Energy, EURAC Research, 39100 Bolzano, Italy (e-mail: atse.louwen@eurac.edu; David.Moser@eurac.edu).

Marko Topic is with the Faculty of Engineering, University of Ljubljana, 1000 Ljubljana, Slovenia (e-mail: marko.topic@fe.uni-lj.si).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JPHOTOV.2021.3078075

performance loss (MS-PL) algorithm divides a nonlinear performance trend into two or more data subsets, which are evaluated independently. The algorithm provides several PLR values for one PV system, divided by breakpoints. This approach has two key advantages. First, possible common performance evolution patterns depending on certain parameters such as the chosen PV technology, the prevailing climate, the mounting type or the year of installation are more easily isolated and, thus, detectable. This knowledge might be very beneficial for optimizing PV systems under varying conditions. Second, the determination of breakpoints in the data offers insight into the piecewise evolution of PV system performance, with breakpoints possibly indicating changes in the dominant failure mode. Thus, this information can aid investigations that aim to identify specific failure modes. The methodology is validated by applying it to synthetically created PV time series, and used to evaluate the MS-PL algorithm of 26 experimental PV systems.

The rest of this article is organized as follows. First, the PV test site in Bolzano/Italy is introduced in Section II. Then, the MS-PL algorithm is described in greater detail and applied on synthetically created PV time series for validation and on 26 PV systems included in the test site. Finally, Section VI concludes this article.

II. Background

As discussed previously, many studies published in recent years have been proposing or evaluating linear performance loss rate algorithms. The most widely used approaches include the year-on-year model, developed by Sunpower [12] and further adapted by NREL [8], and PV time series decomposition followed by a linear regression. Different algorithms and decomposition models have been tested and compared in the literature [9], [13], [14]. The results of an international benchmarking exercise of linear performance loss rate calculation methodologies suggest that averaging across multiple calculation methods provides the most consistent results across different PV systems [15]. Based on these and similar activities, it is visible that the current state of the art focuses on a linear PLR representation. However, a study by Jordan *et al.* [4] pointed out that the understanding of nonlinear degradation paths of PV systems is essential for long term performance studies and lifetime estimations. Studies by Kyprianou *et al.* [16] and Meftah *et al.* [17] suggested approaches to calculate linear performance loss rates as an average of differences in power or performance ratio parameter between years. Another methodology to assess nonlinearities in PV system performance has been developed by Belluardo *et al.* [18], where a performance metric is proposed, which efficiently removes seasonality from the time series. Here, monthly trends of this metric are discussed in terms of nonlinearities in the time series. These approaches are a step towards evaluating nonlinear performance trends of PV systems, but they do not provide information about absolute performance trends or related changing points and are, therefore, difficult to assess.

A way to assess the performance of PV systems in great detail is to transport individual PV modules from an operating plant in periodic intervals (e.g., yearly) to the laboratory and perform a set of standardized indoor measurements [3]. This method gives detailed information of which electrical parameters are subject to performance losses but fails to provide an accurate PLR assessment over a larger portfolio of modules and is only recommended for research purposes due to high related costs.

Fig. 1. Experimental PV plant - Bolzano.

Recently, Theristis *et al.* [19] published a paper reporting an approach similar to the one proposed in this publication demonstrating the rising scientific interest in the topic on nonlinear degradation rates. The originality of our work lies in the foundations built in a previous work presented by Lindig *et al.* [20] and in the application of our methodology in more real cases [21]. Comparing our work to that by Theristis *et al.*, we have analyzed PV systems operating in a different climate, and have applied our algorithm to a larger set of systems, covering more technologies with slightly longer datasets. Instead, we are using with seasonal and trend decomposition using loess (STL) [22] a different decomposition approach. By maximizing a parameter developed specifically for this article, we use a different approach to identify breakpoints. The approach presented by Theristis *et al.* accounts for a number of breakpoints, but in their work, a maximum of only one breakpoints is presented, whereas in our study up to three breakpoints have been detected and reported. Finally, our methodology is presented in greater detail including every individual calculation step. A comparative analysis of different proposed methodologies will be the focus of future publications. The novel algorithm proposed in this article provides a compact way of presenting nonlinear performance loss rates with a meaningful temporal distribution. Thereby, time series of PV system performance are automatically divided into subsets, which are evaluated individually.

III. Test Site

The experimental PV plant under evaluation was installed at the airport of Bolzano/Italy (ABD) in 2010 and consists of 26 different subsystems. They are arranged in two rows of PV arrays, visible in Fig. 1. According to the Koeppen–Geiger classification, the climate in Bolzano is categorized as a temperate climate with warm summers and without dry seasons [23].

Ascencio-Vásquez *et al.* [24] proposed a new PV sensitive climate classification that includes irradiance. This categorization defines the climate in Bolzano as temperate with medium irradiation.

The installed capacities of the individual PV systems range from 1 to 4 kWp per system for eight different technologies: micromorphous silicon (micro-c-Si), amorphous silicon (a-Si) ranging from one to three junctions, cadmium telluride (CdTe), copper indium (gallium) selenide (CIGS), silicon solar cells made out of a heterojunction with an intrinsic thin layer (HIT), mono-crystalline silicon (mc-Si), poly-crystalline silicon (pc-Si), and poly-crystalline silicon string ribbon (ribbon). The systems are indicated with the abbreviation of the solar cell technology and a running number. All systems are ground mounted with a fixed tilt of 30° and an orientation of 8.5° west of south. Due to self-shading of the rows some modules experience local shading in December. Dc-side electrical parameters of each system are measured every 15 min by commercial inverters that provide a high level of accuracy in current and voltage with an average difference of less than 5% and 2%, respectively. Both, dc and ac values are measured.

Additionally, a weather station is located in close proximity to the test side. Here, ambient temperature, various irradiance parameters, such as the plane-of-array (POA) irradiance, and wind speed are recorded. On the rear side of each system the module temperature is measured. The sensors are systematically cleaned and periodically calibrated in order to comply with the IEC61724-1:2017 standard [25].

Weather data are recorded with a measurement frequency of one minute. Since the electrical parameters have a resolution of 15 min, all values are averaged to the longer time interval. A period of nine years is evaluated ranging from February 2011 until January 2020. It is important to mention that the time of observation is not equal to the operation time. The systems began operating in August 2010, roughly six months before the observation time starts. The delayed start of observation was set to exclude initial degradation effects and due to a delayed start of recording weather data.

IV. Methodology

In this section, the MS-PL algorithm is explained by evaluating the PLR of system pc-Si2 from the ABD plant, which has a nominal power of 4.14 kW. Fig. 2 presents the necessary calculation steps, which are subdivided into three main parts. These are input data assessment as well as treatment, the application of the MS-PL algorithm and finally the evaluation of the results. In the following, the individual steps are discussed in detail.

A. Input Data Assessment and Data Treatment

First, the raw data, which are usually power and climate data of irradiance and temperature, are loaded and a quality check is performed. A recommended way to do so is to look at a density plot of instantaneous system power, as shown in Fig. 3(a). Here, normalized 15-min power data of pc-Si2 are plotted in dependency of the time of the day and the day of the year. Normalization is performed by dividing system power

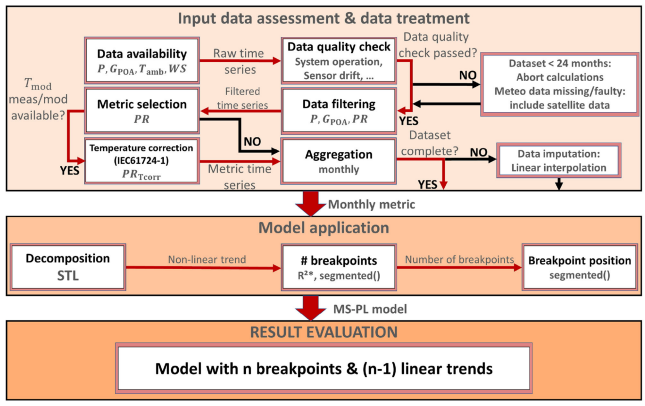

Fig. 2. Diagram showing calculation steps of multistep performance loss algorithm (P-PV power; G_{POA}-irradiance in the plane-of-array; T_{amb}-ambient temperature; WS-wind speed; PR-performance ratio; T_{mod}-module temperature; STL-seasonal and trend decomposition using LOESS [22]; R^{2*}-regression parameter; MS-PL-multistep performance loss).

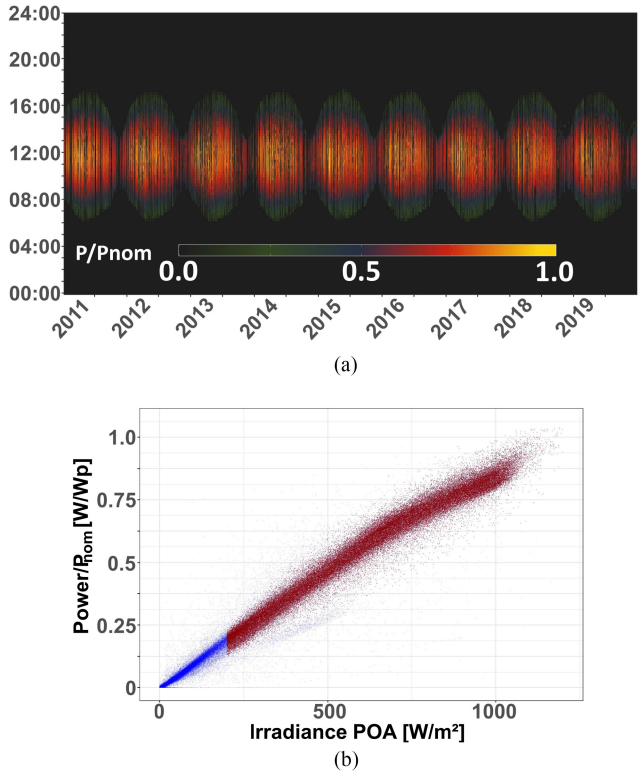

Fig. 3. Data quality check of 15 min values of PV system pc-Si2. (a) Power-density plot. (b) Power versus irradiance in POA. Pnom is the nominal power of the plant.

and the nominal power of the plant. Higher power values for a longer duration of the day are detected in summer because of higher irradiation and longer days in summer time. From such a plot one can detect system outages, timestamp issues, shading instances and also exceptionally strong degradation. The data of pc-Si2 are converted to UTC to remove daylight-saving time shifts in the density plot. It is visible that the data follow a fairly stable pattern and can be considered being of high quality.

A similar approach can be used to assess the quality of the irradiance dataset. A minimum of 24 months of data is required to calculate PLR, although a time series of at least five years is desirable to decrease interannual effects. If no on-site irradiance measurements are available or subject to severe problems such as sensor alignment issues, satellite/reanalysis data can be retrieved and used for the calculations. Additional information on input data quality evaluation can be found in [26].

The next steps involve data filtering and the choice of a performance metric for the application of the subsequent algorithm. Filtering is performed to remove outliers and nonrepresentative values. Power values out of the range of 1–120% of the nominal power are excluded as well as irradiance values lower than 200 and higher than 1200 W/m^2 to capture sufficiently stable measurement conditions. This wide irradiance filter was chosen to capture losses due to both, series and shunt resistance. Performance losses based on shunt resistance have a higher impact at low irradiance levels [27]. Additionally, a performance ratio filter is applied, which removes values out of the range of two times the standard deviation of the monthly mode, which is the most occurring value in each particular month. The effect of the applied filter can be seen in Fig. 3(b). Here, normalized 15-min power values are plotted against the corresponding POA irradiance (G_{POA}) values. The blue dots are the raw data and the remaining red dots are the filtered values. This graph additionally rates the quality of the irradiance dataset.

After filtering, a performance metric was selected and calculated. In this article, we use the performance ratio (PR). The PR is a good indicator for the quality of a PV installation. It compares the measured array yield Y_{a} of a system with its reference yield Y_{ref} [25]. The yields are ratios of energy (E) or irradiation (H) measurements with values obtained under standard test conditions (STC - module tested indoor under T_{STC} = 25 °C, G_{STC} = 1000 W/m^2, air mass 1.5). The PR on the dc side is calculated by

$$\mathrm{PR}_{\mathrm{dc}} = \frac{Y_{\mathrm{a}}}{Y_{\mathrm{ref}}} = \frac{E_{\mathrm{dc}}/P_{\mathrm{nom}}}{H_{\mathrm{POA}}/G_{\mathrm{STC}}} \tag{1}$$

where P_{nom} is the nominal power of the system under STC, E_{dc} is the measured energy on the dc side and H_{POA} is the incoming irradiation on the PV panel. First, the 15-min power and irradiance data points are averaged to hourly values, thereby converting power to energy values ([W] to [Wh]). Afterwards, energy and irradiation measurements are aggregated to monthly values before calculating the PR. The calculations could also be carried out using ac values, thereby including possible inverter-level effects.

It is recommended to use the temperature corrected PR for the calculations if reliable module temperature values are available. E_{dc} and thereby PR is temperature corrected according to [25]

$$\mathrm{PR}_{\mathrm{Tcorr}} = \frac{E_{\mathrm{dc}}/(P_{\mathrm{nom}}(1+\gamma_{\mathrm{T}}(T_{\mathrm{mod}} - 25^\circ\,\mathrm{C})))}{H_{\mathrm{POA}}/G_{\mathrm{STC}}} \tag{2}$$

where γ_{T} denotes the power temperature coefficient and T_{mod} is the module temperature. Unfortunately, the back-of-module temperature measurements at the ABD test site are incomplete during the investigated period. Hence, module temperature was estimated using a well-known model proposed by Sandia [28]

$$T_{\mathrm{mod}} = G_{\mathrm{POA}} * (e^{a+b*\mathrm{WS}}) + T_{\mathrm{amb}} \tag{3}$$

where G_{POA} is the solar irradiance in the POA, WS is the measured wind speed, T_{amb} is the ambient temperature, and a and b are parameters depending on module type as well as mounting configuration. The values for a and b can be found in [28]. System pc-Si2 is categorized as "glass/cell/polymer - Open rack" and the parameters are $a = -3.56$ and $b = -0.075$.

The effect of temperature correction on the monthly PR of pc-Si2 is visible in Fig. 6. Here, the PR and the temperature corrected PR ($\mathrm{PR}_{\mathrm{Tcorr}}$) are depicted next to further calculation steps. Comparing PR and $\mathrm{PR}_{\mathrm{Tcorr}}$, it is visible that the weather dependent variability in the dataset is decreasing substantially. The flattening of the curve is beneficial for the accuracy of the subsequent calculation steps. If single data-points are missing after the aggregation step due to filtering along the data treatment chain, they are imputed using linear interpolation.

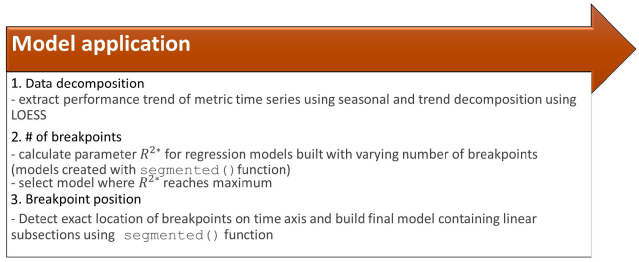

Fig. 4. Calculation steps for multistep performance loss algorithm [22], [29].

B. Model Application

The actual multistep performance loss algorithm contains three steps: data decomposition, determining the optimal amount of breakpoints, and building the final multistep performance loss model. In Fig. 4, these three steps are described.

First, the long-term trend of the performance metric $\mathrm{PR}_{\mathrm{Tcorr}}$ was extracted, by applying the decomposition methodology STL [22]. STL uses locally weighted regression to separate the seasonal part and a remainder from the nonlinear performance trend line. The trend is represented by the dark-red straight line in Fig. 6. Separating the long-term trend from the overall variation in the dataset is performed to prevent the relatively constant seasonal variation from affecting the final results.

In the next step, we utilize the performance trend to calculate the MS-PL values. The idea behind the MS-PL algorithm is to create and rate different regression models based on the data of the performance trend. The individual models differ in the number of breakpoints. The model, which best represents the performance trend without being overfit, will be selected and considered individually. This step is summarized in Fig. 4 and involves the application of a model parameter to find the optimal number of breakpoints for which we build a multistep regression model. This parameter R^{2*} is based on the coefficient of determination R^2, a parameter used in multivariate regression to rate regression models using different amounts of

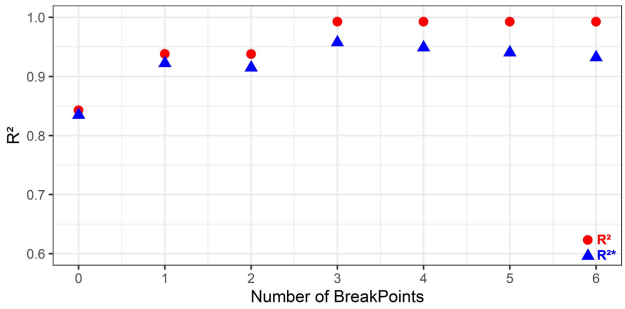

Fig. 5. R^2 and R^{2*} for 7 different regression models including 0 to 6 breakpoints for PV system pc-Si2 considering 95 months of operation.

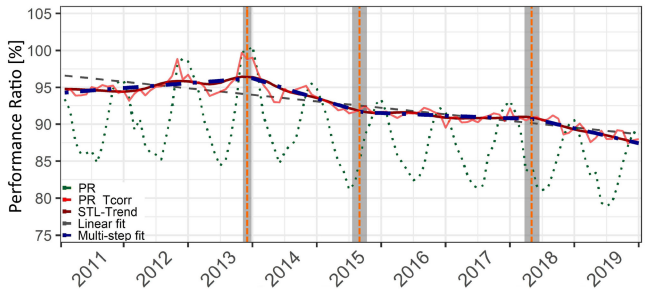

Fig. 6. Multistep performance fit for pc-Si2 system over nine years of observation: PR (dark-green - dotted), PR_{Tcorr} (red - straight), trend-line of PR_{Tcorr} (dark-red - straight), Linear fit (gray - dashed), multistep fit (dark-blue - dotdashed), and vertical breakpoints (orange - dashed).

predictor variables. R^2 is calculated by dividing the residual sum of squares $\text{RSS} = \sum(y_i - \hat{y}_i)^2$ with the total sum of squares $\text{TSS} = \sum(y_i - \overline{y})^2$ and subtracting the result from 1

$$R^2 = 1 - \frac{\text{RSS}}{\text{TSS}} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \overline{y})^2}. \tag{4}$$

Hereby, $\hat{y}_i$ is the modeled, or predicted, value of y_i, which is the corresponding value in the real dataset, and $\overline{y}$ is the average of all values of y_i. R^2 runs from 0 to 1. The closer R^2 of a model is to 1, the better this model describes the underlying data. In other words, R^2 indicates the ability of a model to represent the real dataset.

The problem of R^2 is that it converges towards a value close to 1 with an increasing number of breakpoints, visible in Fig. 5. Here, the model parameter R^2 and R^{2*} are shown for different models including an increasing amount of breakpoints. The reason for this steady increase can be understood when looking at (4). The better a model fits the real dataset, the higher is R^2. The usage of a maximized R^2 results, therefore, inevitably in an overfit. That means that the model is too complex, inhibits a possible error from the original dataset and, therefore, may not represent the underlying trend very well. If we would select the number of breakpoints according to R^2, we would overfit the model. We would not detect the real amount and position of breakpoints in the time series, which may have a physical meaning based on performance loss processes in the PV system.

Thus, we developed R^{2*} on the basis of R^2_{adjusted}. The parameter R^2_{adjusted} penalizes a regression model (and, therefore, R^2) with a high number of predictors, if the addition of predictors does not improve the accuracy of the model. We use this basic idea and adapt R^2_{adjusted} to make it useful for time series and an increasing amount of breakpoints according to the following equation:

$$R^{2*} = R^2 * \frac{n-1}{n+p^*-1}. \tag{5}$$

The amount of time steps (here months) considered is n and p^* is the number of breakpoints included in the model. That means a nearly constant R^2 together with an increasing number of breakpoints p^* results in a decrease of R^{2*}, visible in Fig. 5. An additional, data-driven, penalty is in effect, which determines if the maximum R^{2*} is at least 1.2% higher compared to values expressing models with a lower amount of breakpoints. This extra step is performed to ensure that the model is not overfitting. The optimal model for the performance trend of pc-Si2 has a maximum R^{2*} at three breakpoints, visible in Fig. 5. A higher number of breakpoints decreases R^{2*}.

The calculations are performed in the software environment `R`. We build each model (with 0 to n breakpoints) for an individual performance time series with the `segmented` function [29]. It is first used to determine the optimal number of breakpoints. The input is a linear model of the performance trend against temporal data together with the trend time series of PR_{Tcorr} and a suggested number of breakpoints. The function tests all possible breakpoint positions and estimates the optimal output by minimizing the residual sum of squares RSS. It also delivers R^2, which goes into (5). Based on this results, the optimal number of breakpoints is selected and the final model is built. The optimized number of breakpoints is then used to determine the exact locations together with the fitted linear trend lines. The final model of pc-Si2 can be seen in Fig. 6 (multistep fit). The performance trend of pc-Si2 is divided into four individual sections by three breakpoints, illustrated with vertical orange dashed lines.

C. Result Evaluation

In this section, the MS-PL results are collected and evaluated. In Table I, the results for applying the MS-PL algorithm to the performance trend of all systems of the experimental PV plant are depicted. Under pc-Si2, we find the values corresponding to Fig. 6. The breakpoints are the end-, respectively, starting points of time periods in which a single linear PLR is detected. PLR are presented either in relative or absolute terms. The absolute PLR is calculated by

$$\text{PLR}\,[\%/a] = (\beta_1 t)100 \tag{6}$$

and relative PLR by

$$\text{PLR}\,[\%/a] = \left(\beta_1 \frac{t}{\beta_0}\right)100 \tag{7}$$

β_1 is the gradient and β_0 the y-intercept of the linear trend line. t is a time scaling parameter, which converts the performance metric to a yearly scale (12 for monthly). For the MS-PL algorithm, the absolute PLR is used. The absolute PLR (sometimes abbreviated as performance loss PL) is independent of the initial starting value of the chosen metric. The absolute PLR gives an

indication of the absolute loss rate. The root-mean-square error (RMSE) describes the deviation between the performance trend and the multistep fit. Since the PLR here is an absolute value, which does not take into account the intercept corresponding to PR_{Tcorr} at the beginning of operation, it is important to evaluate PLR values in relation to the intercept. Looking at the results in Fig. 6, it is visible that the modeled intercept computed with the MS-PL algorithm is very close to the real intercept of the performance trend, which is not the case for the linear performance model. This is another important benefit of this algorithm compared to a linear evaluation. The results for pc-Si2 in Table I show that we have a slight initial performance gain for almost three years, followed by a decrease in performance. After four and a half years of operation the performance settles at a lower PLR rate before it decreases again for the last one and a half years. It is suspected that the initial gain is based on a combination of PV technology behavior and favorable weather conditions. In the summer of 2013 and winter of 2013 to 2014 high temperature corrected PR values were recorded before a measurable reduction in performance was detected. As efficiency of crystalline systems decreases under low irradiance conditions due to higher related resistance losses, high average irradiance conditions under constant temperature yields high efficiency values and, therefore, an increase in performance [30], [31]. These high performance values can be connected to a very sunny July, in which the highest monthly yield for this and most other systems under investigation was recorded, and a mild and sunny following winter where the PR reaches very high values of almost 100%. Additionally, the uncertainties of the individual breakpoints and segments have been calculated. The 95% confidence interval around each individual breakpoint ranges from ± 1 to 1.5 months, shown with the grey interval bars around the breakpoints in Fig. 6. The individual segment uncertainties range from ± 0.03 to 0.06%.

V. Validation and Testing

A. MS-PL Algorithm Applied to Validation Systems

Since the exact performance loss rate of a PV-system as a function of time is unknown and difficult to asses, it is complicated to validate a performance loss model. A way to do so is to model a theoretical PV time series, alter the performance with defined linear performance loss rates over certain time periods and apply the MS-PL algorithm in order to see if we can detect the fed-in breakpoints and values.

A 20 year long power time series based on input parameter of PV system pc-Si2 was modeled for that purpose. The modeled time series is based on ERA5 reanalysis irradiance and temperature values [32] for the location of Bolzano, used as inputs for the De Soto "five-parameter" model [33]. Several steps are necessary to model the final dataset.

First, the POA irradiance is modeled using Erbs model [34] for decomposition and the isoptropic model [35] for transposition. The module temperature is modeled with the Sandia module temperature model [28]. The five parameters of the specific PV module type at reference conditions (diode reverse saturation current, light-generated current, series and shunt resistance, diode ideality factor) are taken from the CEC module database [36]. The values, together with modeled weather data are input for the `deSoto` function [33], to model the five parameters at operating conditions for each timestamp, which are required as input for the `singlediode` function. The power in the maximum power point is predicted based on the input parameter according to the single diode equation. All functions as well as the database entries are taken from the `PVLIB` library [37]. The resulting power time series depends on the module input parameters, as well as the operating conditions, and does not exhibit degradation of performance over time. In order to create more realistic data, which exhibits measurement errors, noise is added to the modeled power time series. Comparing measured power data with the results from the De Soto model, we determined the distribution of the De Soto model's errors. Subsequently, we performed random sampling from this distribution to generate the noise. Next, the hourly power data series is altered using random number generators for the number (0 to 3) and position of breakpoints as well as for the sublinear performance loss values (0 to -4%/a in 0.5%/a steps) between two individual breakpoints to create 15 performance datasets. Subsequently, for each individual dataset, filters are applied, PR_{Tcorr} is calculated and the datasets are aggregated to monthly resolution.

For all 15 datasets, the number of breakpoints, the position and the individual PLR values are calculated using the MS-PL approach as explained in Section IV. Overall, 26 breakpoints and 41 PLR values were generated. All breakpoints were found with an average deviation of ± 1.4 months and accompanying PLR values with an average deviation of ± 0.04%/a. Overall, we deem the results to be satisfactory as the modeled breakpoints and PLR values have been found reliably. The deviations between the induced breakpoint positions - PLR values and the calculated ones root from the fact that degradation is induced on raw hourly noisy power data. Afterwards, the data are subject to filtering, irradiance and temperature correction, aggregation, and decomposition. These steps, although crucial for the calculations as to provide clean data trends, do induce a certain error in the raw data and have an effect of their own when carrying out performance studies. An increasing deviation of breakpoint positions is observed if the difference in PLR values on both sides of the breakpoint is too low, i.e., below 0.5%/a. Nevertheless, all breakpoints of the validation systems have been detected within an accuracy range of 0 to 3 months and an overall precise PLR estimation. Based on this results, the algorithm seems to be stable and accurate.

B. MS-PL Algorithm Applied to Time Series of PV Systems at the Airport of Bolzano

The MS-PL algorithm was subsequently applied to 26 experimental PV systems installed at the airport Bolzano (see Section III). The results are shown in Table I. For each system, the intercept of the segmented PLR, the linear PLR with the associated uncertainty, the MS-PL with their associated uncertainties, the respective breakpoints and the RMSE are presented.

TABLE I
LINEAR PLR AND MS-PL VALUES WITH ASSOCIATED UNCERTAINTIES AND BREAKPOINTS FOR PV SYSTEMS IN OPERATION AT THE AIRPORT OF BOLZANO FROM 02/2011 UNTIL 01/2020

	Intercept	PLR_{lin}		PLR_1		PLR_2		PLR_3		PLR_4	
	[%]	[%/a]	$RMSE_{lin}$ [%]	[%/a]	BP1	[%/a]	BP2	[%/a]	BP3	[%/a]	$RMSE_{MS\text{-}PL}$ [%]
HIT1	94.54	-1.13±0.05	1.66	0.88±0.03	05/13	-1.53±0.07					0.56
mc-Si1	95.95	-0.52±0.05	1.30	1.12±0.03	12/13	-2.21±0.06	10/15	-0.04±0.01	07/18	-1.48±0.02	0.20
mc-Si2	91.09	-0.50±0.05	1.21	1.06±0.03	12/13	-2.12±0.07	09/15	-0.09±0.01	06/18	-1.67±0.01	0.19
mc-Si3	94.53	-1.25±0.07	2.19	0.85±0.04	12/13	-1.92±0.01					0.71
mc-Si4	95.12	-1.24±0.07	2.13	0.97±0.04	01/14	-3.25±0.04	01/16	-1.03±0.02			0.51
mc-Si5	95.18	-1.03±0.06	1.97	0.89±0.04	12/13	-2.45±0.07	12/16	1.62±0.02	01/18	-2.49±0.03	0.37
pc-Si1	94.09	-1.56±0.09	2.42	0.15±0.04	10/13	-4.55±0.05	11/15	-0.13±0.05			0.50
pc-Si2	94.32	-0.89±0.05	1.61	0.66±0.03	12/13	-2.68±0.05	09/15	-0.36±0.06	05/18	-2.01±0.04	0.21
pc-Si3	92.32	-1.25±0.06	1.89	0.34±0.04	01/14	-4.39±0.04	05/15	-0.71±0.04			0.47
pc-Si6	91.98	-1.07±0.06	1.84	0.79±0.03	01/14	-3.06±0.08	06/15	-1.08±0.03			0.33
pc-Si7	95.52	-1.23±0.05	1.81	0.46±0.04	08/13	-1.64±0.07					0.57
pc-Si8	94.70	-0.76±0.05	1.84	1.06±0.03	12/13	-2.67±0.04	08/15	-0.34±0.05	06/18	-2.00±0.04	0.18
pc-Si9	93.24	-1.05±0.06	1.40	0.74±0.03	08/13	-1.50±0.04					0.52
ribbon1	94.67	-1.28±0.06	2.00	0.33±0.03	01/14	-4.47±0.07	05/15	-0.76±0.02			0.50
micro-c-Si1	88.53	-1.16±0.06	1.63	-0.20±0.05	02/14	-3.62±0.08	06/15	-0.62±0.01			0.41
micro-c-Si2	84.29	-1.30±0.08	1.88	-0.71±0.05	02/14	-3.53±0.07	08/15	-0.51±0.03			0.58
micro-c-Si3	84.03	-1.09±0.07	1.50	-0.39±0.05	02/14	-3.29±0.08	06/15	-0.54±0.02			0.43
1j-a-Si1	89.61	-1.14±0.07	1.78	-0.31±0.06	03/14	-3.43±0.008	07/15	-0.64±0.02			0.42
1j-a-Si2	93.09	-1.36±0.07	1.81	-0.86±0.06	03/14	-3.89±0.07	07/15	-0.55±0.02			0.57
2j-a-Si1	111.40	-1.23±0.16	4.31	-8.78±0.17	05/13	16.51±0.88	12/13	-3.55±0.66	08/15	-0.38±0.08	1.11
3j-a-Si1	81.56	-0.72±0.10	2.23	4.05±0.16	05/12	0.64±0.15	01/14	-3.60±0.14	09/15	-0.27±0.02	0.35
3j-a-Si2	81.18	-0.64±0.09	2.19	3.705±0.12	06/12	0.53±0.13	01/14	-3.36±0.11	10/15	-0.12±0.09	0.34
CdTe2	91.69	-1.73±0.05	1.44	-0.76±0.04	09/13	-2.00±0.07					0.70
CIGS2	91.38	-1.59±0.10	2.32	-2.67±0.02	02/16	0.04±0.02					0.64
CIGS3	81.05	-2.42±0.19	3.20	-18.42±0.10	12/11	-3.00±0.02	10/15	-0.75±0.05			0.38
CIGS4	85.77	-2.02±0.14	2.72	-12.05±0.07	05/12	-1.39±0.04					0.82

PLR, Performance Loss Rate; u, uncertainty; BP, breakpoint; RMSE, root-mean-square error.

The uncertainties are evaluated from the variance of the linear model coefficients of the performance metric, returning a 95% confidence interval [9]. To calculate the uncertainties, the remainder is added back to the trend line of the performance metric. The RMSE is provided both for linear models and the multistep fits and describes the deviation from the models and the performance trend. Additionally, the uncertainty intervals of the individual breakpoints have been calculated using a 95% confidence interval window. The values range from ±0.2 to 3.2 months with an average of 1.6 months. The individual values are omitted from the table to ensure a good readability of the results. Additionally, the results, averaged per technology, are shown in Fig. 7.

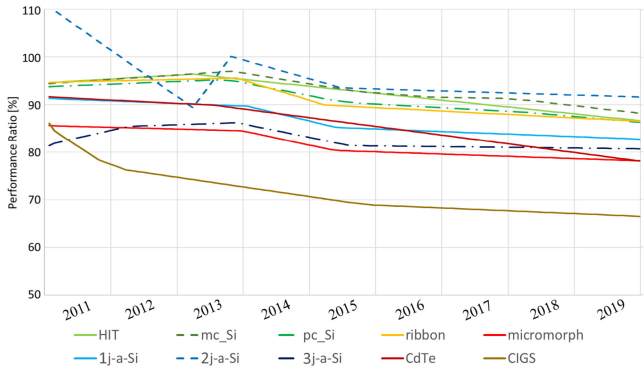

Fig. 7. Average multistep performance of PV systems of ABD test site per technology.

The first four technologies are all types of crystalline silicon whereas the last four belong to thin-film technologies, which are in general subject to higher PLR rates. For these systems, whose performance was evaluated for nine years, a maximum of three breakpoints and four PLR values were calculated. This indicates a performance evolution over time without too many performance changes, especially when looking at the crystalline systems. The intercept with the y-axis of PR_{Tcorr}, which expresses the value at the beginning of lifetime, is very similar for most crystalline systems with a mean of 94.09%. mc-Si2 has the lowest intercept among the crystalline systems with 91.09%. The thin-film systems have a much wider range of intercept values from 81% up to 111%. This observation can be attributed to a more stable and standardized material composition of crystalline silicon systems, which are additionally more similar amongst one another. Also in terms of a linear PLR the values of the crystalline systems have a smaller range, whereby pc-Si1 has the highest value with −1.56%/a. Hot-spots within the cells of this particular system were detected and the front glass has an unusual textured surface with deep pyramids where soiling is likely to adhere. These circumstances might be responsible for its high linear PLR. Except of both triple junction a-Si systems, all thin-film systems are subject to PLR_{lin} larger than −1%/a. CIGS3 and CIGS4 are subject to very high losses. A comparison between electroluminescence images of the systems and new modules of the same type show that the active PV material

is not operating well anymore. Major system degradation was observed through the MS-PL algorithm within the first year of observation with values of PLR_1 equal to -18.42%/a and -12.05%/a, respectively. Unfortunately, it is not possible to trace back the reasons, but it is assumed that a strong degradation of the active PV material might have caused high losses at such an early stage of operation. Furthermore, a very strong initial performance loss can be seen for 2j-a-Si1, expressed in $PLR_1 = -8.78\%$/a. This loss can be attributed to a disconnection of the junction box of one module within the system, which occurred between October 2012 and July 2013 and thereby causing a partial electrical disconnection of the modules of the PV group. Because of that, the performance of this system is subject to major fluctuations. This additionally results in the highest RMSE of all systems stemming from a highly nonlinear PR_{Tcorr}-trend evolution and, therefore, a high deviation with the multistep fit.

In Section IV, where the model was applied to pc-Si2, a slight initial performance gain was observed. This is the case for all crystalline systems under observation, with values ranging from 0.15 to 1.12%/a. That strengthens the assumption that the increase might be caused by a combination of crystalline silicon technology and the weather effects which were described in Section IV-C. Furthermore, the performance evolution of the crystalline systems seem to follow a certain pattern. After the detected performance gain ($PLR_1 > 0$), most systems are subject to an elevated loss (PLR_2) for roughly three years before they settle at a lower PLR rate (PLR_3). Here, the performance of mc-Si5 increases again for one year in 2017 before it drops again. Furthermore, the PLR of another four crystalline systems increases from the middle of 2018 to values above -1%/a. The observed period is too long to assign this high PLR to being a seasonal effect. It is visible that the performance of crystalline systems follow a certain trend. Thereby, no distinct differences between the individual crystalline systems have been detected. It would be interesting to investigate similar crystalline systems installed in different climate zones to see if these effects are dependent on the prevailing outdoor/climate conditions or could possibly be generalized. If that would be the case, a generalized performance pattern of crystalline PV systems could be developed and implemented in PV performance modeling software.

Looking at the thin-film systems a certain pattern is recognizable, although not as clear. It is visible that the systems start with a performance loss. An exception are both 3j-a-Si systems, which, just as crystalline systems, are subject to an initial performance gain. It is assumed that this initial gain can be correlated to thermal annealing. The modules of both 3j-a-Si systems are film laminated with a flexible polymer glued to an aluminum-zinc plate, which is installed onto a corrugated metal sheet. This kind of mounting systems do not allow effective module cooling via rear-side ventilation, resulting in higher operating temperatures. It has been shown that annealing mechanisms in 3j-a-Si modules are more effective at higher temperatures close to 80 °C [38]. It is expected that nonoptimal system cooling caused higher operating temperatures, which in turn led to boosted thermal annealing affects, represented in positive PLR_1 values.

The high yields due to favorable weather conditions reported for the crystalline systems are not visible for most thin-film systems because they seem to be offset by high initial degradation. PLR_2 is characterized by performance losses of most systems. When evaluating the results of the thin-film systems, it is obvious that a generalization of the performance trend is inaccurate and a further subdivision into the individual PV module types is necessary.

A summary of the overall results are shown in Fig. 7. Here, the multistep performance trend averaged by technology is shown. The discussed system related performance trends are visible; for instance, the initial performance gain for crystalline systems, the strong decline of PLR_1 of 2j-a-Si as well as the CIGS systems, and the very similar evolution for the crystalline systems.

VI. Summary and Outlook

In this article, a new performance loss rate methodology for the calculation of a self-regulated MS-PL is presented. Instead of simplifying the performance evolution of a PV system to a linearized performance loss rate, this algorithm provides a detailed and easily comprehensible performance evaluation. Thereby, the performance trend of the temperature corrected performance ratio of a PV system is first extracted using time series decomposition and subsequently divided into a number of segments with different gradients. These time series parts are separated by a self-regulated breakpoint analysis and evaluated independently. The final result is a linear multistep model of the performance evolution of a PV plant. The MS-PL algorithm identifies an adequate number and position of breakpoints individually. A major benefit of a MS-PL compared to a linear PLR is the fact that the performance analysis is presented in greater detail. The MS-PL algorithm is a tradeoff between a simple, but too generalized, linear PLR and a very precise, but also very poorly interpretable, nonlinear PLR. One possible limitation of the approach lies in the fact that breakpoints are sensitive to maintenance activities/repairs and possibly even cleaning. On one hand these can be very useful information, for example in technical due-diligence. On the other hand, such breakpoints are not reporting on anything technology- or climate-related and thereby do not increase the understanding of correlations between degradation and system characteristics. Furthermore, we believe a minimum duration for each segment should be set to six months in order to find breakpoints reliably.

In Section V-A, the algorithm is tested on synthetically created datasets. In total, 15 20-year, initially degradation less, time series were created using functions of the `PVLIB` library and ERA5 temperature and irradiance data. Noise, extracted from measured data, was added to the datasets before a distinct degradation, with a random number of breakpoints and random performance loss values, was added to the time series. By applying the MS-PL algorithm all breakpoints were detected with an average deviation of ± 1.4 months and the PLR values calculated within acceptable margins of $\pm 0.04\%$/a.

Additionally, the algorithm is applied to a number of PV systems of an experimental PV plant installed in Bolzano/Italy. The plant includes 26 different systems, which are divided into eight

different technologies and the observation time amounts to nine years. By dividing the systems into crystalline and thin-film PV systems, it was seen that the performance evolution differs. Most crystalline systems show a distinct pattern in the performance trend. The performance experiences a slight gain for a little less than three years. Afterwards, a stable performance loss sets in or a strong one-to-two year decrease settling at a moderate PLR value. Some of the systems experience another higher loss at the end of observation. The overall linear performance loss rate of most crystalline systems reaches values between -0.5% and -1.3%/a. The thin-film PV systems under observation experience a higher overall PLR. The values for micro-c-Si systems range from -1.1 to -1.3%/a, for a-Si systems from -0.6 to -1.4%/a, the CdTe system is subject to a linear PLR of -1.7%/a and the CIGS systems have the highest losses ranging from -1.6 to -2.4%/a. For six of twelve systems, an initial lower PLR increases after three years. The other six systems experience distinctly different performance patterns. Looking at the results, it is important to subdivide the thin-film systems into their corresponding technologies to avoid false generalization conclusions.

It is foreseen to create further validation datasets and test the algorithm under a variety of different conditions in terms of number of breakpoints, observed time period, climate, and PV technology. It will be interesting to gain results of the MS-PL algorithm applied to longer PV system time series, especially to see if the performance resembles a reliability bathtub curve with high losses at the beginning and end of operation, as it is expected and often presumed in the literature. An important application could be the usage in due-diligence of companies who want to take over the operation of PV plants from the secondary market. This algorithm allows a quick health analysis of a PV system. One could reduce the amount of data to analyze by having just to evaluate the time series data close to detected breakpoints. It can be expected to find performance anomalies close to the breakpoints, and by looking at corresponding maintenance tickets, if available, it might be possible to detect certain issues the system experienced in the past. Also operation and maintenance companies might benefit from the application of this methodology. By running the algorithm in regular intervals, PV performance degrading effects could be detected early in an automatized way. By having full access to the tickets of a well tracked PV system, it could be possible to connect the detected anomalies with recorded issues a PV system had to endure, or even specific degradation modes. By building up a database containing the frequency and severity of degradation modes together with the position of the breakpoints, the impact of certain degradation modes and other failures could be predicted more accurately. By applying the algorithm to a great number of systems, it might be possible to see performance trend pattern depending on different parameter. Such results could lead to certain optimization possibilities in terms of climate specific module designs and when planning a PV plant under specific known circumstances. The mentioned application possibilities clearly show the benefit of the MS-PL algorithm compared to a linear PLR evaluation. Last but not least, anomaly detection possibilities go far beyond the scope of a linear PLR.

ACKNOWLEDGMENT

The authors thank the Department of Innovation, Research and University of the Autonomous Province of Bozen/Bolzano for covering the Open Access publication costs.

REFERENCES

[1] *IRENA Renewable Capacity Statistics 2020*, International Renewable Energy Agency, Abu Dhabi, United Arab Emirates, 2020.

[2] M. Köntges, G. Oreski, M. Herz, U. Jahn, P. Hacke, and K.-A. Weiss, "Assessment of photovoltaic module failures in the field," Photovolta. Power Syst. Programme, Int. Energy Agency, Paris, France, Rep. IEA PVPS T13-09:2017, 2017.

[3] J. S. Stein, C. Robinson, B. King, C. Deline, S. Rummel, and B. Sekulic, "PV lifetime project: Measuring PV module performance degradation: 2018*Indoor flash testing results*," in *Proc. IEEE 7th World Conf. Photovolt. Energy Convers.*, Jun. 2018, pp. 0771–0777.

[4] D. C. Jordan, T. J. Silverman, B. Sekulic, and S. R. Kurtz, "PV degradation curves: Non-linearities and failure modes," *Prog. Photovolt.: Res. Appl.*, vol. 25, no. 7, pp. 583–591, 2017.

[5] S. W. Adler, M. S. Wiig, A. Skomedal, H. Haug, and E. S. Marstein, "Degradation analysis of utility-scale PV plants in different climate zones," *IEEE J. Photovolt.*, vol. 11, no. 2, pp. 513–518, Mar. 2021.

[6] A. Phinikarides, G. Makrides, N. Kindyni, A. Kyprianou, and G. E. Georghiou, "ARIMA modeling of the performance of different photovoltaic technologies," in *Proc. 39th IEEE Photovolt. Specialists Conf.*, Jun. 2013, pp. 0797–0801.

[7] A. Phinikarides, N. Kindyni, G. Makrides, and G. E. Georghiou, "Review of photovoltaic degradation rate methodologies," *Renewable Sustain. Energy Rev.*, vol. 40, pp. 143–152, Dec. 2014.

[8] D. C. Jordan, C. Deline, S. R. Kurtz, G. M. Kimball, and M. Anderson, "Robust PV degradation methodology and application," *IEEE J. Photovolt.*, vol. 8, no. 2, pp. 525–531, Mar. 2018.

[9] S. Lindig, I. Kaaya, K.-A. Weiss, D. Moser, and M. Topic, "Review of statistical and analytical degradation models for photovoltaic modules and systems as well as related improvements," *IEEE J. Photovolt.*, vol. 8, no. 6, pp. 1773–1786, Nov. 2018.

[10] M. Köntges *et al.*, "Review of failures of photovoltaic modules," Photovolta. Power Syst. Programme, Int. Energy Agency, Paris, France, Rep. IEA PVPS T13-03:2014, 2014.

[11] A. Virtuani *et al.*, "35 years of photovoltaics: Analysis of the TISO-10-kW solar plant, lessons learnt in safety and performance-Part 1," *Prog. Photovolt.: Res. Appl.*, vol. 27, no. 4, pp. 328–339, 2019.

[12] E. Hasselbrink *et al.*, "Validation of the PVLife model using 3 million module-years of live site data," in *Proc. IEEE 39th Photovolt. Specialists Conf.*, Jun. 2013, pp. 0007–0012.

[13] A. J. Curran, C. B. Jones, S. Lindig, J. Stein, D. Moser, and R. H. French, "Performance loss consistency and uncertainty across multiple methods and filtering criteria," in *Proc. IEEE 46th Photovolt. Specialists Conf.*, Jun. 2019, pp. 1328–1334.

[14] D. C. Jordan *et al.*, "Reducing interanalyst variability in photovoltaic degradation rate assessments," *IEEE J. Photovolt.*, vol. 10, no. 1, pp. 206–212, Jan. 2020.

[15] S. Lindig *et al.*, "International collaboration framework for the calculation of performance loss rates: Data quality, benchmarks, and trends (towards a uniform methodology)," *Prog. Photovolt.: Res. Appl.*, vol. 29, no. 6, pp. 573–602, 2021.

[16] A. Kyprianou, A. Phinikarides, G. Makrides, and G. E. Georghiou, "Robust principal component analysis for computing the degradation rates of different photovoltaic systems," in *Proc. 29th Eur. Photovolt. Solar Energy Conf. Exhib.*, Jan.2014, pp. 2939–2942.

[17] M. Meftah, E. Lajoie-Mazenc, M. Van Iseghem, R. Perrin, D. Boublil, and K. Radouane, "A less environment-sensitive and data-based approach to evaluate the performance loss rate of PV power plants," in *Proc. 36th Eur. Photovolt. Solar Energy Conf. Exhib.*, Sep. 2019, pp. 1554–1559.

[18] G. Belluardo, P. Ingenhoven, W. Sparber, J. Wagner, P. Weihs, and D. Moser, "Novel method for the improvement in the evaluation of outdoor performance loss rate in different PV technologies and comparison with two other methods," *Sol. Energy*, vol. 117, pp. 139–152, Jul. 2015.

[19] M. Theristis, A. Livera, C. B. Jones, G. Makrides, G. E. Georghiou, and J. Stein, "Nonlinear photovoltaic degradation rates: Modeling and

comparison against conventional methods," *IEEE J. Photovolt.*, vol. 10, no. 4, pp. 1112–1118, Jul. 2020.

[20] S. Lindig, D. Moser, and M. Topic, "Multi-step performance loss rates of photovoltaic systems," Presented at *9th SOPHIA Workshop PV-Module Rel.*, May 2019.

[21] S. Lindig, D. Moser, B. Müller, K. Kiefer, and M. Topic, "Application of dynamic multi-step performance loss algorithm," in *Proc. 47th IEEE Photovolt. Specialists Conf.*, Jun. 2020, pp. 0443–0448.

[22] R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. Terpenning, "STL: A seasonal-trend decomposition procedure based on loess," *J. Official Statist.*, vol. 6, no. 1, pp. 3–33, Mar. 1990.

[23] M. Kottek, J. Grieser, C. Beck, B. Rudolf, and F. Rubel, "World map of the Köppen-Geiger climate classification updated," *Meteorologische Zeitschrift*, vol. 15, pp. 259–263, May 2006.

[24] J. Ascencio-Vasquez, K. Brecl, and M. Topic, "Methodology of Köppen-Geiger-photovoltaic climate classification and implications to worldwide mapping of PV system performance," *Sol. Energy*, vol. 191, pp. 672–685, Oct. 2019.

[25] International Electrotechnical Commission, *Photovoltaic System performance, Part 1: Monitoring.* Geneva, Switzerland, IEC 61724-1:2017, 2017.

[26] S. Lindig, A. Louwen, D. Moser, and M. Topic, "Outdoor PV system monitoring-input data quality, data imputation and filtering approaches," *Energies*, vol. 13, no. 19, Sep. 2020, Art. no. 5099.

[27] C. S. Ruschel, F. P. Gasparin, E. R. Costa, and A. Krenzinger, "Assessment of PV modules shunt resistance dependence on solar irradiance," *Sol. Energy*, vol. 133, pp. 35–43, 2016.

[28] D. L. King, W. E. Boyson, and J. A. Kratochvil, "Photovoltaic array performance model," Sandia Nat. Lab., Livermore, CA, USA, Tech. Rep. SAND2004-3535, Dec. 2004.

[29] V. M. R. Muggeo, "Segmented: An R package to fit regression models with broken-line relationships," *R News*, vol. 8/1, pp. 20–25, 2008. [Online]. Available: https://cran.r-project.org/doc/Rnews/

[30] F. Mavromatakis, F. Vignola, and B. Marion, "Low irradiance losses of photovoltaic modules," *Sol. Energy*, vol. 157, pp. 496–506, 2017.

[31] I. de la Parra, M. Muñoz, E. Pigueiras, M. García, J. Marcos, and F. Martínez-Moreno, "PV performance modelling: A review in the light of quality assurance for large PV plants," *Renewable Sustain. Energy Rev.*, vol. 78, pp. 780–797, Oct. 2017.

[32] Copernicus Climate Change Service (C3S), "ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. copernicus climate change service climate data store (CDS)," 2017. [Online]. Available: https://cds.climate.copernicus.eu/cdsapp/home

[33] W. De Soto, S. Klein, and W. Beckman, "Improvement and validation of a model for photovoltaic array performance," *Sol. Energy*, vol. 80, no. 1, pp. 78–88, 2006.

[34] D. Erbs, S. Klein, and J. Duffie, "Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation," *Sol. Energy*, vol. 28, no. 4, pp. 293–302, 1982.

[35] H. Hottel and B. Woertz, "Evaluation of flat-plate solar heat collector," *Trans. Amer. Soc. Mech. Eng.*, vol. 64, pp. 64–91, 1942.

[36] A. P. Dobos, "An improved coefficient calculator for the california energy commission 6 parameter photovoltaic module model," *J. Sol. Energy Eng.*, vol. 134, no. 2, Mar. 2012, Art. no. 021011.

[37] W. Holmgren, C. Hansen, and M. Mikofski, "pvlib python: A python package for modeling solar energy systems," *J. Open Source Softw.*, vol. 3, Sep. 2018, Art. no. 884.

[38] I. Pola, D. Chianese, and A. Bernasconi, "Flat roof integration of a-Si triple junction modules laminated together with flexible polyolefin membranes," *Sol. Energy*, vol. 81, no. 9, pp. 1144–1158, 2007.