Comparative Analysis of Change-Point Techniques for Nonlinear Photovoltaic Performance Degradation Rate Estimations

Marios Theristis, Andreas Livera, Leonardo Micheli, Julián Ascencio-Vásquez, George Makrides, George E. Georghiou, and Joshua S. Stein

***Abstract*—A linear performance drop is generally assumed during the photovoltaic (PV) lifetime. However, operational data demonstrate that the PV module degradation rate (*Rd*) is often nonlinear, which, if neglected, may increase the financial uncertainty. Although nonlinear behavior has been the subject of numerous publications, it was only recently that statistical models able to detect change-points and extract multiple *Rd* values from PV performance time-series were introduced. A comparative analysis of six open-source libraries, which can detect change-points and calculate nonlinear *Rd*, is presented in this article. Since the real *Rd* and change-point locations are unknown in field data, 960 synthetic datasets from six locations and two PV module technologies have been generated using different aggregation and normalization decisions and nonlinear degradation rate patterns. The results demonstrated that coarser temporal aggregation (i.e., monthly vs. weekly), temperature correction, and both PV module technologies and climates with lower seasonality can benefit the change-point detection and *Rd* extraction. This also raises a concern that statistical models typically deployed for *Rd* analysis may be highly climatic- and technology-dependent. The comparative analysis of the six approaches demonstrated *median* mean absolute errors (MAE) ranging from 0.06 to 0.26%/year, given a maximum absolute *Rd* of 2.9%/year. The *median MAE* in change-point position detection varied from 3.5 months to 6 years.**

***Index Terms*—Change-point analysis, modeling, nonlinear degradation, photovoltaics (PVs).**

Manuscript received May 4, 2021; revised July 15, 2021; accepted September 8, 2021. Date of publication September 29, 2021; date of current version October 21, 2021. The work of Marios Theristis and Joshua S. Stein was supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy under the Solar Energy Technologies Office Award Number 34366. The work of Andreas Livera, George Makrides, and George E. Georghiou was supported by the University of Cyprus through the PV-4LIFE project. *(Corresponding author: Marios Theristis.)*

Marios Theristis and Joshua S. Stein are with the Sandia National Laboratories, Albuquerque, NM 87185 USA (e-mail: mtheris@sandia.gov; jsstein@sandia.gov).

Andreas Livera, George Makrides, and George E. Georghiou are with the PV Technology Laboratory of the FOSS Research Centre for Sustainable Energy and the Department of Electrical and Computer Engineering, University of Cyprus, 1678 Nicosia, Cyprus (e-mail: livera.andreas@ucy.ac.cy; makrides.georgios@ucy.ac.cy; geg@ucy.ac.cy).

Leonardo Micheli is with the Centro de Estudios Avanzados en Tierra, Energía y Medio Ambiente, University of Jaén, 23071 Jaén, Spain (e-mail: lmicheli@ujaen.es).

Julián Ascencio-Vásquez is with 3E SA, B-1000 Brussels, Belgium (e-mail: julian.ascencio.vasquez@gmail.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JPHOTOV.2021.3112037.

Digital Object Identifier 10.1109/JPHOTOV.2021.3112037

I. Introduction

Degradation rates (*Rd*) inform photovoltaic (PV) lifetime energy yield predictions. Knowledge of this metric is therefore important for feasibility, reliability and financial analyses of PV systems. Simplistic assumptions may cause detrimental effects increasing PV financial uncertainties [1] and, hence, investment risk [2]. Such assumptions may include the usage of: single *Rd* values from literature; values reported from different climatic conditions; values reported for a specific technology but built with differing manufacturing quality or with different packaging materials; and the assumption of constant performance loss over time.

On the other hand, relatively simple statistical analyses can be performed to extract the *Rd* of a system based on its performance time-series. However, it is known that PV performance fluctuates due to a number of seasonally related factors such as temperature [3], spectrum [4], soiling [5], or even abrupt changes caused by failures [6], etc. Therefore, although the statistical tools are available and relatively easy to use, it is inherently challenging to extract reproducible and accurate PV *Rd*s [7], [8].

The way PV time-series data are handled and processed adds to the complexity and uncertainty [9]. For example, if temperature correction is applied and the reported temperature coefficients are biased, a seasonality due to changing temperature ranges will be introduced. Furthermore, spatial and temporal variability in the actual temperature of the array could lead to seasonality in the performance metric, especially if the wind direction varies seasonally and affects spatial patterns of array temperature. Moreover, with respect to data quality, integrity and processing, the *Rd* calculation can also be influenced by other factors such as missing data [10], sensor drift [11], filtering criteria [12], temporal aggregation [13], etc. Therefore, data handling and quality assurance is crucial in an *Rd* pipeline [10].

Ambitious targets for 50-year PV module lifetimes [14] tend to focus on design challenges, but they also require tools for predicting *Rd* at an early stage in order to inform service lifetime (e.g., [15]) and provide reassurance during operation. Of equal importance is the ability to detect subsequent changes in *Rd*, especially at the beginning and end of life, in order to help identify the mechanisms behind such behaviors. Furthermore, the ability to detect and quantify changes in *Rd* behavior is important for condition-based maintenance. For example, assuming an installation that does not experience inverter clipping, the energy recovery from cleaning is reduced with time due to degradation, therefore, cleaning frequency would vary depending on *Rd* patterns and other financial metrics [16]. However, in the case of high dc/ac ratios, significant levels of soiling can be tolerated in the early years of the PV plants due to clipping, while in later years soiling can actually have more impact as the system degrades and clips less often.

The most commonly used assumption to statistically extract the *Rd* is of a linear performance drop over time. Operational data demonstrate that this is unrealistic in some cases [17] mainly due to the initial and wear-out degradation that may occur [18]. For example, light and elevated temperature induced degradation [19] and light induced degradation [20] may occur initially in the PV module lifetime and then cease once an equilibrium has been reached, directly affecting the *Rd*. Nonlinear PV degradation behavior and its corresponding financial impacts have been reported in numerous studies [11], [18], [21]. However, it was only recently that change-point methodologies were introduced to analyze nonlinear degradation [17], [22]–[24] and soiling [25], [26] in PV systems. Such methodologies can automatically detect changes in PV time-series/profiles and quantify the loss rates for different time periods. Some change-point methods were also applied on real PV performance data verifying that some systems may exhibit nonlinear performance loss [17].

Although these methods make use of open-source libraries from Python and *R* (which are commonly used for statistical analysis) they have never been extensively compared "side-by-side" against PV time-series of known degradation behavior. This can only be achieved by leveraging synthetic PV performance datasets, because it is impossible to know the "real" values of *Rd* in field data. This work expands on a previous study where three methods were compared using five synthetic datasets from a single location and PV module technology [22]. Here, 960 synthetic datasets are generated for six different locations and two PV technologies in order to compare the performance of six different open-source libraries. The reason for using a much larger amount of datasets is to test the algorithms against different conditions (two- and three-step degradation profiles, seasonality in different climate zones and module technology differences) and analysis methodologies (i.e., aggregation, temperature correction) enabling a robust comparative investigation. It should be noted that only degradation is synthesized in this article, and therefore the term *Rd* is used. The presented models, however, can also be applied on field data to extract the performance loss rate (*PLR*); a metric used to represent all performance losses.

II. Methodology

In statistics, a change-point (or switch-point, or break-point) refers to a change in time-series properties (e.g., mean, variance, correlation, etc.) [17]. Such changes can be either continuous or discontinuous and in the case of nonlinear degradation, the change is considered as continuous since the segments have the same *Rd* value at the change-point [18].

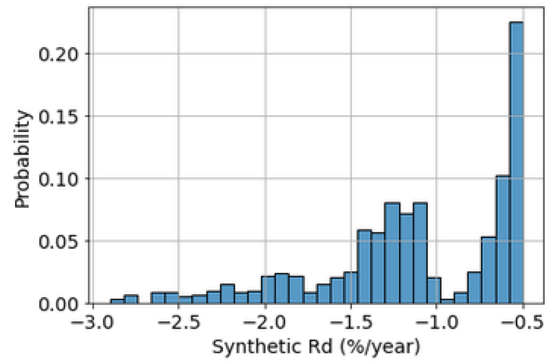

Fig. 1. Probability histogram of randomly sampled degradation rates. The minimum and maximum absolute *Rd* were 0.5%/year and 2.9%/year, respectively.

Trend extraction and change-point detection methods are frequently used statistical tools and several open-source algorithms are available in Python and *R*. Six of these algorithms were selected in this study based on the hypothesis that they can be applied to PV performance data.

A. Generation of Synthetic Datasets

Six locations (i.e., Bangkok, Phoenix, Albuquerque, Berlin, Moscow, Minneapolis) were selected to include weather classification indices from all Koppen–Geiger-PV (KGPV) climate zones except polar [27].

Long-term meteorological data (hourly over 30 years) were sourced from the global reanalysis ERA5 of the European Centre for Medium-Range Weather Forecasts [28]. These data were used as inputs to PV performance models of a monocrystalline silicon (c-Si) and cadmium telluride (CdTe) modules using the Sandia PV array performance model [29] from *pvlib-python* [30]. Nonlinear *Rd*s were then applied randomly (see Fig. 1) to the performance data based on two- and three-step profiles (i.e., one change-point [cp1] and two change-points [cp2]). Although any number of change-points can be synthesized, in this article, the *Rd*s and change-point locations were selected in a way to emulate the reliability bathtub curve with higher absolute *Rd*s in the beginning- and end-of-life. To ensure realistic scenarios, the following boundaries were set: total performance loss between 20% and 30% in 30 years of operation; *Rd* from −0.5 to −3%/year depending on segment (i.e., Rd_1 and Rd_3 between −0.5 and −3%/year and Rd_2 between −0.5 and −1.5%/year); and avoid change-points in the first and last year due to the difficulty of such models to detect change-points in those locations.

The "raw" power time-series were then normalized using the performance ratio [31] and aggregated monthly/weekly, with/without temperature correction. Only a nighttime filter was applied to the "raw" time-series (plane-of-array irradiance from 100–1200 W/m^2). Ten different *Rd*s were applied for every possible combination of location, technology, degradation pattern, aggregation and temperature correction resulting in 960 synthetic datasets. The six open-source libraries were then applied and their performance and effectiveness was evaluated

TABLE I
LIST OF LOCATIONS, LIBRARIES, AND VARIED PARAMETERS

Locations*	Package-Language	Varied parameters
Bangkok (AH)	*fbp-py*	Technology (c-Si, CdTe)
Phoenix (BK)	*NumPy-py*	Degradation pattern (cp1, cp2)
Albuquerque (CK)	*cpm-R*	Aggregation (weekly, monthly)
Berlin (DL)	*segmented-R*	Temperature correction (yes, no)
Moscow (EL)	*strucchange-R*	Degradation rates**
Minneapolis (EM)	*Rbeast-R*	

*First letter of KGPV classification [27] represents temperature and precipitation (A: Tropical, B: Desert, C: Steppe, D: Temperate, and E: Cold); the second letter is based on solar irradiation (L: Low, M: Medium, H: High, and K: Very high).
**Ten different nonlinear degradation rates for each combination.

based on the ability to detect the number and position(s) of change-point(s) and also the precision errors in estimating the *Rd* values. The list of locations, libraries and varied parameters are given in Table I.

Besides knowing the "real" *Rd* values and change-point positions, synthetic data also have the advantage of being independent of sensor drift, temperature uncertainty, soiling, maintenance issues etc., that may affect the accuracy and/or uncertainty of the calculations. Typically, when *Rd* or *PLR* methodologies are compared using field data, the "true" value is considered to be the mean of all estimates; assuming that this is the closest to the "real" value [8]. However, caution is recommended in the presence of outliers, which can dominate the mean value; in this case, the median should be used.

B. Description of Modeling Approaches

Breakpoint (*bkp*) is a function within the *R* package "*strucchange*" for testing, monitoring, and dating structural changes in linear regression models [32]. This function implements the algorithm described by Bai and Perron [33] for the simultaneous estimation of multiple break-points in time-series regression models. The *bkp* algorithm estimates the break-points by minimizing the residual sum of squares of the regression model. A dynamic programming approach based on the Bellman's principle of optimality [34] is followed for estimating the optimum break-points, given the number of breaks (which was the only input parameter requirement for the execution of this algorithm).

In this article, the given time-series were initially decomposed into seasonal, trend and irregular components through the seasonal and trend decomposition by Loess [35] method and then the *bkp* algorithm was applied to the extracted trend component.

Sequential and Batch Change Detection Using Parametric and Nonparametric Methods (cpm) [36]: Different change-point models for performing both parametric and nonparametric change detection on univariate data are implemented in the *R* package "*cpm*." In this study, the "*detectChangePointBatch*" function was used to detect a single change-point in the given time-series and estimate its location using the Mann–Whitney test statistic. This detects location shifts in a stream with a possibly unknown distribution by testing a hypothesis. The data are processed in one batch and information regarding the existence of a change-point are returned.

The "*processStream*" function was initially tested for detecting multiple change-points resulting in large errors. For this reason, the *cpm* algorithm was only used for detecting one change-point. The algorithm requires as input the desired value of "alpha", which was set to 0.05 in this study. The null hypothesis of no change is rejected when the length of the sequence (D_n) is greater than the upper alpha percentile of the test statistic distribution (h_n).

Facebook prophet algorithm (*fbp*) [37] is an open-source library, available in Python and *R,* used to forecast time-series based on an additive decomposition model, which combines trend, seasonality and holidays (not considered). Once the trend is extracted for the *PR* time-series, change-point analysis is performed to identify the number and location(s) of change-point(s) by capturing statistical changes in the slopes of predefined segments of the time-series. In order to detect the most significant change-points in an automated manner, the *SciPy.signal.find_peaks* [38] function was implemented. Once the nonlinear trend is "sliced," the methodology treats each segment in a linear manner. This is achieved by applying the ordinary least squares method [39] in order to compute the different *Rd*s for each segment, accordingly.

In order to setup *fbp* to provide meaningful results for PV behavior at the lowest possible computational burden, the flexibility of the extracted trend, number of potential change-points, and range had to be adjusted. With respect to the monthly aggregation, similar settings as in Theristis *et al*. [17] were used with a *changepoint_prior_scale* $= 0.07$ whereas for the weekly aggregation, the *changepoint_prior_scale* was modified to 0.04 in order to reduce the trend's flexibility (weekly aggregation fluctuates more than monthly). The number of potential change-points was set to 100 for both weekly and monthly aggregation to lower computational burden; when not batch analyzing a large number of datasets, a higher number of potential change-points could be used to potentially improve the results even further.

*Piecewise regression (*pw-p *[python version*] *[40] and pw-R [R version] [41])*: In piecewise linear regression (or segmented regression), the data points are fitted through a sequence of linear functions. Each function is assigned to a specific time period. In this case, each pair of consecutive functions was fitted to have the same value at the change-point in order to model a continuous trend. Compared to other change-point algorithms, this approach allows one to simultaneously identify change-points and corresponding slopes (i.e., *Rd*s).

This *pw-p* model was implemented using the *NumPy's piecewise* function [40] and the *SciPy's curve_fit* function [38] whereas the *pw-R* utilized the *segmented* function implemented in the *R* package "*segmented"* [41]. One of the drawbacks in the formulation of *pw-p* is that it is only able to identify a predetermined number of change-points whereas *pw-R* can utilize the "*selgmented: selecting number of breakpoints in segmented regression"* function, which automatically selects the number of change-points according to a sequential hypothesis testing (via the Score test). In this study, however, *pw-p* was also automated and customized by examining the standard deviation returned for the date of each detected change-point. For each time-series, the cp_x and cp_{x+1} scenarios were iteratively ran, starting from $x = 1$ and using incremental steps of 1 (x represents the number of change-points). The iteration stopped and the number of change-points was set to x when cp_{x+1} exhibited standard deviations higher than one year on at least one of the detected change-point dates. It is worth mentioning that no seasonal decomposition was employed in *pw-R* and *pw-p*, which were directly fed with the raw synthetic datasets.

The main difference between the *pw-p* and *pw-R* is that the latter implements a bootstrap restarting algorithm. The *pw-R* implementation requires the class of the model (set to standard linear) and the number of bootstrap samples (set to 10), given that the segmented model fitting process was set to "true."

Bayesian estimation of abrupt change, seasonality and trend (*RBeast*) [42] applies the Bayesian ensemble time-series decomposition algorithm. By utilizing the ensemble modeling technique, the results of the multiple fitted models are determined and incorporated into the final averaged Bayesian model. *Rbeast* enables the determination of abrupt changes (i.e., change-points), cyclic variations (e.g., seasonality) and nonlinear trends in time-series observations by decomposing the time-series data into three components: abrupt changes, trends and cyclic/seasonal variations. The algorithm also quantifies the likelihood of the detected changes.

In this study, the period of the cyclic/seasonal component of the time-series was either set to 12 or 52 depending on the aggregation level (i.e., monthly or weekly). The optimal number of change-points was selected by iteratively distributing 0 to 60 change-points on the trend; once all change-point combinations are executed, the one with the highest probability is selected.

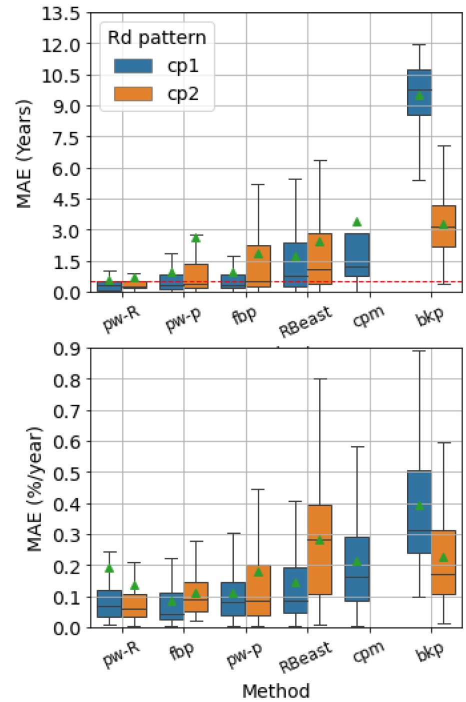

Fig. 2. *MAE* comparison of change-point detection in Years (top) and nonlinear degradation rate estimation (bottom) using all methods and two- and three-step degradation patterns; *cpm* was applied for one change-point only. The green triangles indicate mean values and the red dashed horizontal line indicates the 0.5 Year as a reference. The methods are sorted by their median values.

III. Results and Discussion

The application of the six approaches on the 960 synthetic datasets resulted in 5760 simulations revealing different slopes, mutually verifying the presence of nonlinear, two- and three-step behaviors. The comparison utilizes the mean absolute errors (*MAEs*) of change-point location in Years (i.e., mean of absolute error of $|CP_{\iota,true} - CP_{\iota,calculated}|$ where ι is 1, 2) and *Rd*s at each segment in %/year (i.e., mean of absolute error of $|Rd_{\kappa,true} - Rd_{\kappa,calculated}|$ where κ is 1, 2, 3 depending on the segment). The boxplots do not include outliers for clarity; instead, the mean values are indicated by the green triangles.

A. Two-Step and Three-Step Nonlinearity

With respect to the ability of these models to locate the change-point(s) location(s) in two- (i.e., one change-point) and three-step (i.e., two change-points) nonlinear patterns (see Fig. 2, top), the mean *MAE* varied from 0.55 to 9.50 Years whereas the median *MAE* varied from 0.26 to 9.76 Years. The *pw-R* model seems to perform equally with either one or two change-points whereas the *bkp* exhibits a significantly better performance with two change-points, although the median is still high (around 3.2 Years).

The most critical part of the performance comparison in this study is the prediction error in quantifying the *Rd* at each segment. As shown in Fig. 2 (bottom), the *Rd MAE* ranged from 0.08 to 0.39%/year (mean) and 0.04 to 0.31%/year (median). Overall, all methods except *bkp* exhibited mean and median *MAE* within 0.3%/year. The piecewise linear regression models (both *pw-R* and *pw-p*) and *fbp* outperformed the others in both detection of change-point locations (median within 6 Months) and *Rd* extraction, independently of the number of change-points in the time-series.

B. Aggregation Level

When comparing the impact of aggregation (i.e., weekly or monthly) the mean and median *MAE* ranged from 0.59 to 6.43 Years and 0.25 to 6.06 years, respectively (see Fig. 3, top). Overall, the majority of methods seem to perform slightly better when time-series are aggregated monthly with up to ∼1.6 year improvements in the mean values (in the case of *cpm*) as compared to weekly. This is contradictive because conventional statistical models for linear *Rd* extraction perform better with larger amounts of data and hence, lower-than-monthly aggregation levels. However, since change-point techniques try to detect statistical changes in the time-series, they are not favored by the higher oscillations caused by lower aggregation levels.

With respect to the *Rd* estimations (see Fig. 3, bottom), the influence of aggregation seems to be relatively low with median differences up to 0.04%/year (in the case of *cpm*). However, mean differences of up to 0.12%/year were exhibited by the *pw-R*, which was due to the influence of outliers. The mean and median *MAE* ranged from 0.09 to 0.31%/year and 0.06 to

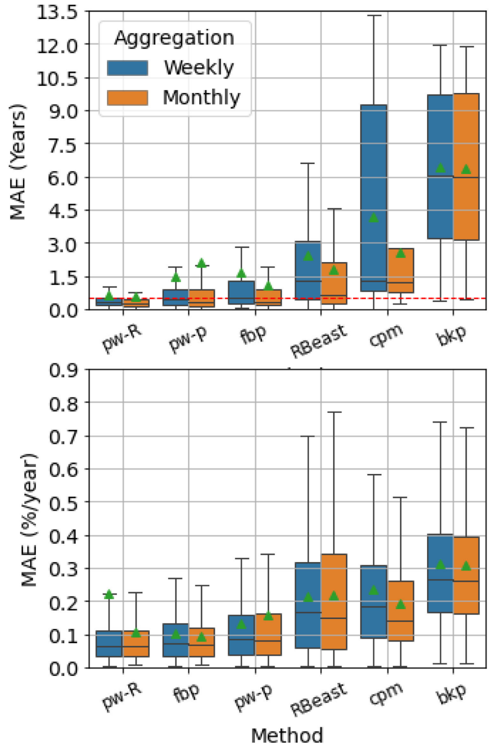

Fig. 3. *MAE* comparison of change-point detection in years (top) and nonlinear degradation rate estimation (bottom) using all methods for weekly and monthly aggregation. The green triangles indicate mean values and the red dashed horizontal line indicates the 0.5 year as a reference. The methods are sorted by their median values.

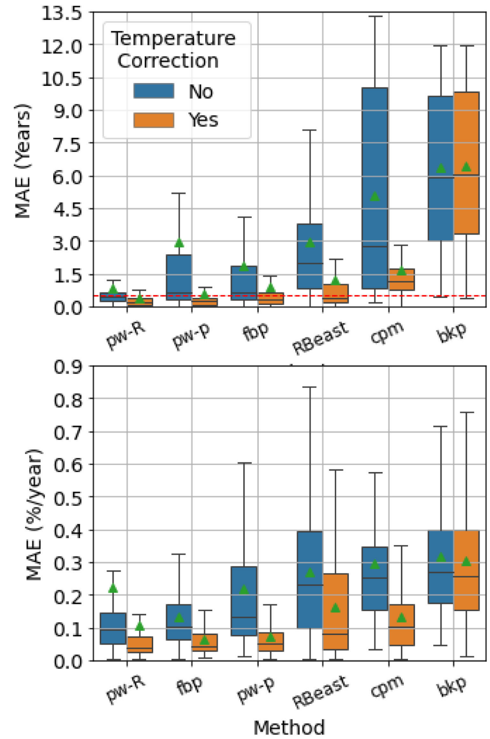

Fig. 4. Comparison of *MAE* of change-point detection in years (top) and nonlinear degradation rate estimation (bottom) using all methods with and without temperature correction. The green triangles indicate mean values and the red dashed horizontal line indicates the 0.5 year as a reference. The methods are sorted by their median values.

0.27%/year, respectively and, again, there is a clear differentiation of best- and worst-performing methods with piecewise linear regression models and *fbp* outperforming *Rbeast*, *cpm,* and *bkp*.

C. Temperature Correction

Temperature correction in time-series provided great improvements in both change-point detection (see Fig. 4, top) and *Rd* extraction (see Fig. 4, bottom). This aligns with the conventional *Rd* models and data pipelines since temperature correction reduces signal fluctuations. Specifically, with respect to change-point detection, the mean and median *MAE* varied from 0.39 to 6.42 years and 0.21 to 6.02 years, respectively. All methods except the *bkp* improved significantly with temperature correction with up to a 3.37 Years reduction in mean *MAE* in the case of *cpm*. The best performing models, *pw-R*, *pw-p*, *fbp,* exhibited median *MAE* improvements of 0.22, 0.44, and 0.35 Years, respectively.

The *Rd* estimation was also improved with temperature correction as shown in Fig. 4 (bottom). Mean and median *MAE* ranged from 0.07 to 0.32%/year and 0.04 to 0.27%/year, respectively. Maximum absolute differences between temperature corrected and uncorrected time-series were exhibited by the *cpm* model with a 0.16%/year decrease in mean *Rd* estimation.

D. Different PV Technologies

With respect to the influence of different PV technologies (e.g., c-Si and CdTe) on the change-point detection ability of the investigated methods, the mean and median *MAE* values varied from 0.31 to 6.45 years and 0.23 to 6.01 years, respectively, (see Fig. 5, top). Although all methods exhibited lower median *MAE* values in the case of CdTe, the differences were marginal for *bkp* and *cpm* with 0.03 and 0.07 years, respectively whereas the largest improvements of 0.64 Years occurred with *Rbeast*.

As can be seen in Fig. 5 (bottom), the prediction of the corresponding *Rd* values at each segment varied from 0.08 to 0.32%/year (mean *MAE*) and 0.04 to 0.26%/year (median *MAE*). As expected, *Rd* in c-Si PV modules is more difficult to be estimated compared to CdTe. This is due to the different seasonal behavior (e.g., magnitude of temperature coefficient) of c-Si modules compared to CdTe. Specifically, all methods besides *bkp* were more accurate in predicting *Rd* by up to 0.14%/year absolute difference in median *MAE*. This indicates that seasonal fluctuations also affect the prediction ability of such models, which raises a concern when applying methods and data analysis pipelines that are not technology-independent and require additional training.

E. Different Geographic Locations

Perhaps the most important criterion for an *Rd* extraction model in general, is whether it is location-independent or not.

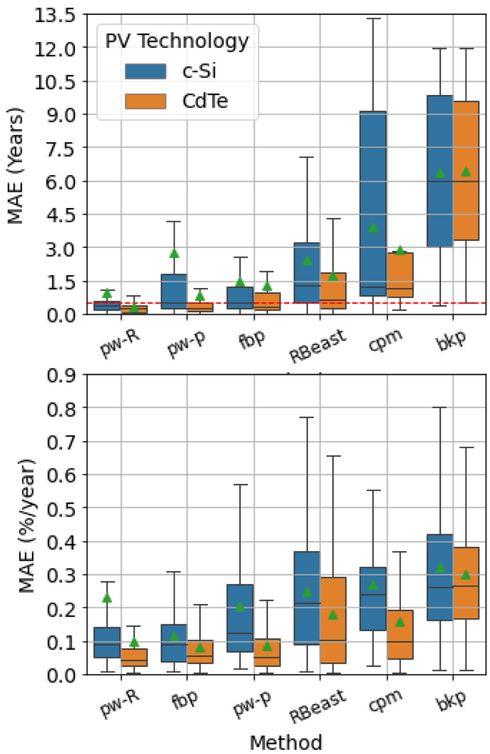

Fig. 5. Comparison of *MAE* of change-point detection in years (top) and nonlinear degradation rate estimation (bottom) using all methods for different PV technologies: c-Si and CdTe. The green triangles indicate mean values and the red dashed horizontal line indicates the 0.5 year as a reference. The methods are sorted by their median values.

Here, the performance of the models under investigation was compared at six different KGPV climates. In Fig. 6 (left), most of the models exhibited different climatic dependence with mean and median *MAE* ranging from 0.24 to 6.52 Years and 0.17 to 7.31 Years, respectively. More specifically, the most climate-dependent methods defined as the ones that exhibit the maximum mean and median *MAE* differences are the *cpm* and *bkp*, respectively with 3.44 and 1.78 Years. Piecewise regression models (*pw-R* and *pw-p*) and *fbp* proved to be the most robust with respect to their climate-dependence with median *MAE* differences from 0.12 (*fbp*) to 0.30 (*pw-p*) Years. However, mean *MAE* differences increase to 1.28 Years for the *fbp* model when comparing Bangkok and Phoenix. Overall, Minneapolis (EM, cold with medium irradiation) was the most challenging location for all models with a mean and median *MAE* values of 3.19 and 1.90 years, respectively. On the other hand, locations with lower seasonal fluctuations such as Bangkok (AH, tropical with high irradiation) and Phoenix (BK, desert with very high irradiation), were the least challenging locations with a mean *MAE* of 2.17 Years in Bangkok, and a median *MAE* of 1.47 Years in the case of Phoenix.

With respect to the *Rd* quantification, the mean and median *MAE* values varied from 0.07 to 0.34%/year and 0.04 to 0.28%/year, respectively (see Fig. 6, right). All mean and median *MAE* fluctuations (i.e., absolute difference of maximum and minimum *MAE* for each method) were within 0.15%/year. Furthermore, Moscow (EL, cold with low irradiation) and Minneapolis were the most challenging locations/climates with mean *MAE* of 0.21%/year whereas Phoenix was as low as 0.12%/year. Nevertheless, *MAE* within 0.2%/year should be considered as acceptable since this is within the commonly reported statistical uncertainty.

TABLE II
Qualitative Classification of Execution Time

Computational burden		
Low	**Medium**	**High**
pw-R	*bkp*	*fbp*
pw-p		*Rbeast*
cpm		

F. Overall Comparison and Discussion

The models under investigation were naturally grouped based on their performance into two categories: best- and worst performing models. Best-performing models include the piecewise linear regression models in both *R* (*pw-R*) and Python (*pw-p*) as well as *fbp* whereas the list of worst-performing models includes *cpm*, *Rbeast* and *bkp*. In Fig. 7 (top), all models are summarized with respect to their ability to detect the change-point locations using all 960 synthetic datasets. It can be seen that the mean and median *MAE* ranged from 0.62 to 6.41 years and 0.29 to 6 years, respectively. The best-performing models exhibited median *MAE* values below 0.5 year, whereas the worst-performing model (*bkp*) reached median values of up to 6 Years. Furthermore, the mean values provide information on the influence of outliers that cause more skewed distributions, especially in the case of *pw-p* and *cpm*, whereas the *pw-R* proves to be the most robust in this case.

On the other hand, with respect to the *Rd* estimation (see Fig. 7, bottom), most methods demonstrated relatively low prediction errors, even when the change-point detection error was relatively high. Mean and median *MAE* ranged from 0.10 to 0.31%/year and 0.06 and 0.26%/year. The best-performing models exhibited median *MAE* within 0.1%/year, whereas the worst-performing models were between 0.16 (*Rbeast*, *cpm*) and 0.26%/year (*bkp*).

The computational burden of each model is qualitatively categorized in Table II. The reason for not quantifying the execution time is because the assessment is based on the execution of 960 datasets in a loop using different machines, whereas in real applications a much lower number of datasets, or more optimized loops or higher performance computers may be used. Nevertheless, the piecewise linear regression models and *cpm* provided the fastest results whereas *fbp* and *Rbeast* suffer from high computational burden.

Originally, all libraries required pre-knowledge of the number of change-points, which is not a reasonable approach for measured PV data. For this reason, the best four among the six methods (*pw-R*, *pw-p*, *fbp*, and *Rbeast*) were customized in this study to automatically detect the number of change-point(s). The remaining two methods (*cpm* and *bkp*) were not customized because even in the perfect scenario where they were fed with the

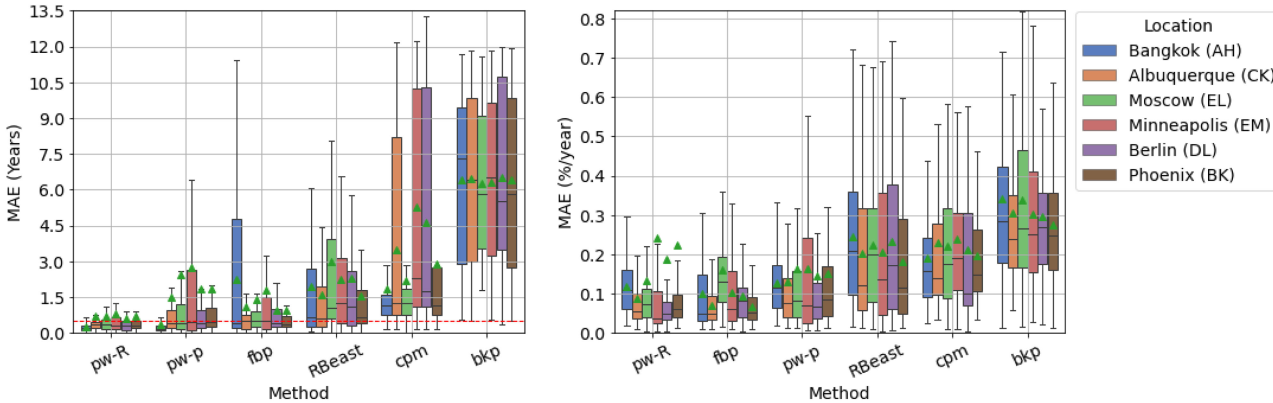

Fig. 6. Comparison of *MAE* of change-point detection in years (left) and nonlinear degradation rate estimation (right) at different locations. The green triangles indicate mean values and the red dashed horizontal line indicates the 0.5 year as a reference. The methods are sorted by their median values.

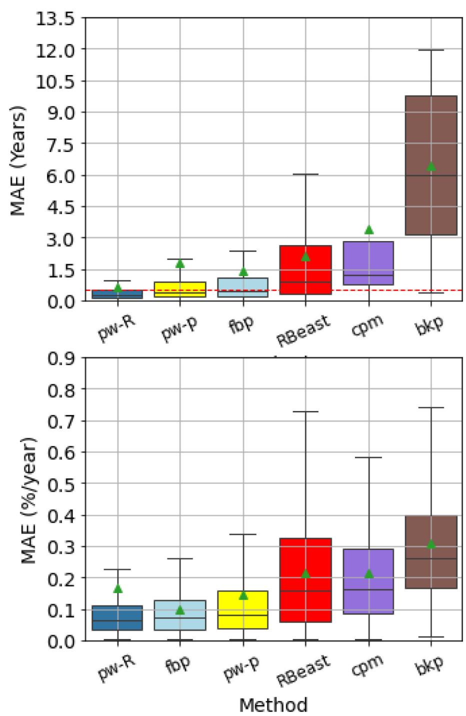

Fig. 7. Overall comparison of *MAE* for change-point detection in years (top) and nonlinear degradation rate estimation (bottom) using all 960 synthetic datasets. The green triangles indicate mean values and the red dashed horizontal line indicates the 0.5 year as a reference. The methods are sorted by their median values.

real number of change-point(s), they failed to accurately detect change-point(s) and/or calculate the corresponding *Rd*s. Furthermore, although the ability to detect only one change-point might be limiting, the *cpm* library was included in this comparative analysis due to its high speed and simplicity in its implementation (only time-series are required as input). Therefore, such libraries may be beneficial for applications of e.g., real-time monitoring where speed might be more important than accuracy (given that the accuracy is reasonable). In this case, the median *MAE* of *cpm* was within 0.2%/year for all locations except Moscow, which might be reasonable for some applications.

This study focused on irreversible effects characterized by the synthesized degradation. With real field data, degradation is masked by other effects, which can also be reversible, such as soiling, snow, partial shading due to vegetation, etc. Since reversible and irreversible effects vary or fluctuate differently in PV performance time-series (e.g., soiling fluctuates differently and at different magnitudes compared to snow and/or degradation), change-point models would require separate tuning in order to differentiate and quantify each individual loss. Otherwise, the same models can be applied to field data to extract the *PLR* metric, which considers all losses as a whole. However, existing pipelines and trend-based performance loss differentiation algorithms (e.g., [43]) can be coupled to the investigated models for detecting and quantifying nonlinear *Rd*s.

IV. Conclusion

A comparative analysis was performed to investigate different open-source methodologies for detecting and quantifying nonlinear PV *Rd*s. Since the real change-point location(s) and *Rd*s are unknown, synthetic datasets from six locations were generated using different analysis pipelines.

Based on this analysis and considering accuracy in change-point detection, *Rd* extraction and computational time, piecewise linear regression libraries in *R* and Python are clearly the winners. Other models with additional functionalities, such as forecasting in the case of *fbp* can also be used with the cost of high computational burden.

Overall, it was shown that monthly aggregation, temperature correction and locations/climates and PV module technologies with lower seasonality favor the change-point algorithms. These outcomes raise concerns when applying methods and data analysis pipelines that are not technology- and climate-independent or extensively verified. Since such large-scale datasets are not readily available, and considering that the real *Rd* values and change-point locations are unknown, implementing synthetic datasets proved to be a good solution.

Future work will examine the robustness of the top performing algorithms against noisy datasets.

Acknowledgment

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

References

[1] B. Stridh *et al.*, "Uncertainties in Yield Assessments and PV LCOE," Report IEA-PVPS T13-18: 2020.

[2] D. C. Jordan and S. R. Kurtz, "Photovoltaic degradation rates—An analytical review," *Prog. Photovolt. Res. Appl.,* vol. 21, pp. 12–29, 2013.

[3] E. Skoplaki and J. A. Palyvos, "On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations," *Sol. Energy,* vol. 83, pp. 614–624, 2009.

[4] C. Stark and M. Theristis, "The impact of atmospheric parameters on the spectral performance of multiple photovoltaic technologies," in *Proc. 42nd IEEE Photovolt. Specialist Conf.*, 2015, pp. 1–5.

[5] L. Micheli and M. Muller, "An investigation of the key parameters for predicting PV soiling losses," *Prog. Photovolt. Res. Appl.,* vol. 25, pp. 291–307, 2017.

[6] A. Livera *et al.*, "Recent advances in failure diagnosis techniques based on performance data analysis for grid-connected photovoltaic systems," *Renewable Energy,* vol. 133, pp. 126–143, 2019.

[7] 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.

[8] 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, pp. 573–602, 2021.

[9] D. C. Jordan and S. R. Kurtz, "The dark horse of evaluating long-term field performance-data filtering," *IEEE J. Photovolt.,* vol. 4, no. 1, pp. 317–323, Jan. 2014.

[10] A. Livera *et al.*, "Data processing and quality verification for improved photovoltaic performance and reliability analytics," *Prog. Photovolt., Res. Appl.,* vol. 29, pp. 143–158, 2021.

[11] 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.

[12] A. J. Curran *et al.*, "Performance loss rate consistency and uncertainty across multiple methods and filtering criteria," in *Proc. 46th IEEE Photovolt. Specialists Conf.*, 2019, pp. 1328–1334.

[13] B. Paudyal *et al.*, "PV plant performance loss rate assessment: Significance of data filtering and aggregation," in *Proc. 46th IEEE Photovolt. Specialists Conf.*, 2019, pp. 0866–0869.

[14] J. S. Stein, "Technical opportunities for improving PV modules and systems? The path to a 50 yr module," Sandia Rep. No. SAND2019-14779C, 2019.

[15] I. Kaaya *et al.*, "Photovoltaic lifetime forecast model based on degradation patterns," *Prog. Photovolt., Res. Appl.,* vol. 28, pp. 979–992, 2020.

[16] L. Micheli *et al.*, "Photovoltaic cleaning frequency optimization under different degradation rate patterns," *Renewable Energy,* vol. 166, pp. 136–146, 2020.

[17] M. Theristis *et al.*, "Nonlinear photovoltaic degradation rates: Modeling and comparison against conventional methods," *IEEE J. Photovolt.,* vol. 10, no. 4, pp. 1112–1118, Jul. 2020.

[18] D. C. Jordan *et al.*, "PV degradation curves: Non-linearities and failure modes," *Prog. Photovolt., Res. Appl.,* vol. 25, pp. 583–591, 2017.

[19] F. Kersten *et al.*, "A new mc-Si degradation effect called LeTID," in *Proc. 42nd IEEE Photovolt. Specialist Conf.*, 2015, pp. 1–5.

[20] B. Sopori *et al.*, "Understanding light-induced degradation of c-Si solar cells," in *Proc. 38th IEEE Photovolt. Specialists Conf.*, 2012, pp. 1115–1120.

[21] D. C. Jordan *et al.*, "Compendium of photovoltaic degradation rates," *Prog. Photovolt., Res. Appl.,* vol. 24, pp. 978–989, 2016.

[22] M. Theristis *et al.*, "Modeling nonlinear photovoltaic degradation rates," in *Proc. 47th IEEE Photovolt. Specialists Conf.*, 2020, pp. 0208–0212.

[23] 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.*, 2020, pp. 443–0448.

[24] S. Lindig, A. Louwen, D. Moser, and M. Topic, "New PV performance loss methodology applying a self-regulated multistep algorithm," *IEEE J. Photovolt.,* vol. 11, no. 4, pp. 1087–1096, Jul. 2021.

[25] L. Micheli *et al.*, "Improved PV soiling extraction through the detection of cleanings and change points," *IEEE J. Photovolt.,* vol. 11, no. 2, pp. 519–526, Mar. 2021.

[26] L. Micheli, M. Muller, E. F. Fernandez, and F. Almonacid, "Segmentation of deposition periods: An opportunity to improve PV soiling extraction," in *Proc. 47th IEEE Photovolt. Specialists Conf.*, 2020, pp. 0595–0598.

[27] J. Ascencio-Vásquez *et al.*, "Methodology of Köppen-Geiger-photovoltaic climate classification and implications to worldwide mapping of PV system performance," *Sol. Energy,* vol. 191, pp. 672–685, 2019.

[28] H. Hersbach *et al.*, "The ERA5 global reanalysis," *Quart. J. Roy. Meteorol. Soc.,* vol. 146, pp. 1999–2049, 2020.

[29] D. L. King *et al.*, "Photovoltaic array performance model," Sandia Rep. No. SAND 2004-3535*,* 2004.

[30] W. F. Holmgren, R. W. Andrews, A. T. Lorenzo, and J. S. Stein, "PVLIB python 2015," in *Proc. 42nd IEEE Photovolt. Specialists Conf.*, 2015, pp. 1–5.

[31] M. Theristis *et al.*, "Chapter II-1-B - Energy yield in photovoltaic systems A2 - Kalogirou, S. A.," in *McEvoy's Handbook of Photovoltaics*, 3rd ed., Cambridge, MA, USA: Academic, 2018, pp. 671–713.

[32] A. Zeileis, F. Leisch, K. Hornik, and C. Kleiber, "strucchange: An R package for testing for structural change in linear regression models," *Journal of Statistical Software*, vol. 7, no. 1, pp. 1–38, 2002.

[33] J. Bai and P. Perron, "Computation and analysis of multiple structural change models," *J. Appl. Econometrics,* vol. 18, pp. 1–22, 2003.

[34] D. E. Kirk, *Optimal Control Theory: An Introduction*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1970.

[35] R. B. Cleveland *et al.*, "STL: A seasonal-trend decomposition," *J. Official Statist.,* vol. 6, pp. 3–73, 1990.

[36] G. J. Ross, "Parametric and nonparametric sequential change detection in R: The cpm package," *J. Stat. Softw.,* vol. 66, pp. 1–20, 2015.

[37] S. J. Taylor and B. Letham, "Forecasting at scale," *Amer. Statistician,* vol. 72, pp. 37–45, 2018.

[38] E. Jones, T. Oliphant, and P. Peterson, "SciPy: Open Source Scientific Tools for Python," 2001, www.scipy.org.

[39] S. Makridakis *et al.*, *Forecasting Methods and Applications*. Hoboken, NJ, USA: Wiley, 2008.

[40] C. R. Harris *et al.*, "Array programming with NumPy," *Nature,* vol. 585, pp. 357–362, 2020.

[41] V. M. Muggeo, "Segmented: An r package to fit regression models with broken-line relationships," *R News,* vol. 8, pp. 20–25, 2008.

[42] K. Zhao *et al.*, "Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm," *Remote Sens. Environ.,* vol. 232, 2019, Art. no. 111181.

[43] A. Skomedal and M. G. Deceglie, "Combined estimation of degradation and soiling losses in photovoltaic systems," *IEEE J. Photovolt.,* vol. 10, no. 6, pp. 1788–1796, Nov. 2020.