A non-destructive method for crack quantification in photovoltaic backsheets under accelerated and real-world exposures

Addison G. Klinke, Abdulkerim Gok, Silas I. Ifeanyi, Laura S. Bruckman, * 

Abstract

The long-term durability of photovoltaic modules is paramount for the continued growth of the industry. Polymer backsheets are of particular concern since they provide electrical insulation and an environmental barrier. In this study, 23 freestanding, multilayer backsheets with nine unique material combinations underwent four different weathering exposures under accelerated and real-world conditions. Besides changes in color and gloss, the induced degradation included parallel or mud cracks on 11 backsheets, sometimes in combination with delamination or blistering. Similar degradation has been observed in previous studies and is concerning since cracks compromise the mechanical integrity and electrical safety of backsheets. Quantitative parameters are desirable to reliably classify categories of cracks and supply unbiased features for statistical analysis in predictive lifetime models. We developed an analysis technique that utilizes surface profilometry data to quantify the depth, width, area, spacing, and number of cracks. Parameters are automatically extracted from the raw data by an algorithm running on a high performance distributed computing cluster. Our algorithm excelled at characterizing parallel cracks with minimal de-adhesion, and only an estimated 4% of crack detections were false positives. The addition of humidity and temperature variation formed up to three times as many cracks on a photodose basis compared to dry, constant temperature exposures. Cracks in real-world and accelerated exposures propagated to similar depths with equivalent photodoses; however, the number of cracks formed in accelerated exposures was far greater on a photodose basis. Of samples that cracked, the best performing backsheet configuration was polyvinyl fluoride/poly (ethylene-terephthalate)/polyethylene (PVF/PET/PE) while the least durable was PET/PET/ethylene-vinyl acetate. None of the six PVF/PET/PVF backsheets cracked in any of the exposures.

1. Introduction

With rapidly dropping prices, the photovoltaics (PV) industry is becoming an increasingly important contributor to the global energy market [1,2]. PV energy costs have been approaching price parity with fossil fuels resulting in an increased emphasis on large-scale, industrial PV power plants [3,4]. To continue this positive trend and promote wide-scale adoption of PV, it is imperative to understand the forms of degradation that can jeopardize a module’s long-term profitability [2,5–7].

Our approach to this lifetime and degradation science [8] challenge utilizes data science and statistics to extract unbiased correlations between environmental stressors and degradation responses. Identifying these statistically significant relationships requires large data sets [7]. Knowledge of physical models then enables further exploration of the indicated correlations to identify active and quiescent degradation pathways. Cross-correlation between accelerated and real-world models also helps develop accelerated lifetime qualification standards and tests [9] which accurately reflect real-world PV degradation [10–13]. As additional data becomes available, models can be iteratively updated and improved, ultimately leading to predictive lifetime models that are validated against a wide range of experimental data.

Backsheets are typically three layer, thin polymer laminate films [14] and provide an environmental barrier as well as insulation...
against electrical shock due to the high dielectric breakdown strength of poly (ethylene-terephthalate) (PET) which is commonly used for the core layer. Premature backsheet degradation has been correlated with losses in overall module performance [15–19].

Backsheet cracks increase the likelihood of ground faults and current leakage [9], result in insufficient electrical insulation, and create pathways for further degradation via electrochemical corrosion [18,19]. This form of degradation is well-documented in field retrieved modules [19–21] and a variety of accelerated tests [19–22]. The outer layers of backsheets were constructed of either PET [19–21], polyvinylidene fluoride (PVDF) [19], polyamide (PA) [21], or various undisclosed commercial backsheets [22]. PET and PA backsheets sustained substantial losses in mechanical properties including reduced elongation at break [21], lower molecular weight, and broader molecular weight distribution [19]. Severe cracks in a fielded PET backsheet exposed the tabbing ribbon, creating a potential path for electrical leakage and compromising wet insulation resistance [9,20]. Thermal shrinkage tests conducted on samples degraded under accelerated exposures demonstrated greater contraction in the direction of crack propagation, suggesting that polymer chains become aligned in this direction as a result of extrusion flow during manufacturing [22,23]. Consequently, intermolecular forces in the transverse direction are weaker which allows crack propagation.

Degrading polymers exhibit a wide variety of crack patterns when exposed to either field or laboratory conditions [19–22,24,25]. For backsheets in particular, these include micro cracks [21,26], parallel or directional cracks [19,20,22], wide longitudinal cracks with transverse branches [20], and networks of branching diagonal cracks (referred to here as mud flat cracks) [19]. These patterns occur on both the outer [19,20] and inner layers [19] of backsheets. Qualitative observations show increases in the width and number of cracks over time [22].

Although it is common to measure mechanical properties related to backsheet cracking (i.e. elongation at break, molecular weight, etc.), a well-established methodology for quantifying the precise dimensions, number, and spacing of cracks does not currently exist. Laser scanning confocal microscopy (LSCM) has been used in other fields [27,28], but was only recently employed for backsheets by Lin et al. [29]. Their study used LSCM to quantify crack density for PET backsheets under uni-axial fragmentation testing. The resulting relationship between crack density and normalized applied strain matched a theoretical film/substrate model developed by Hsueh and Yanaka [30,31]. Cracks were not observed until rectangular samples were stretched in a pair of grips post-exposure to generate enough strain to cause cracking. It is not clear whether this fracture mechanism would have eventually occurred naturally with enough UV exposure if the external strain had not been applied.

Other fields where crack quantification is crucial, such as pavement maintenance in civil engineering [32], coating durability in the automotive industry [33], detection of enamel cracks in dentistry [34,35], and fracture of poly (methyl-methacrylate) (PMMA) Fresnel lenses in concentrating PV [36,37] can provide insights for developing an approach to measure backsheets. For instance, modern pavement maintenance often employs a combination of cameras and laser line scanners [38–40]. X-ray computer tomography (XCT) has also been used to follow the evolution of microstructural changes in materials, including crack formation [41–43].

A widely-applicable method for crack quantification is necessary to enhance the performance of PV module predictive lifetime models. Our paper addresses this problem with the introduction of a data analysis method for quantifying the width, depth, area, spacing, and number of backsheet cracks. Here, we report on the development and validation of this algorithm and insights gained from its application on a sample dataset.

2. Weathering study

2.1. Backsheets

Each of the 23 three layer backsheet samples were cut into 64, 38.1 x 50.8 mm rectangular, freestanding film replicates such that the long and short edges corresponded to the machine and transverse film extrusion directions, respectively (Table 1). Materials were selected to cover a variety of those commonly used in industry.

2.2. Exposures

Replicates were split into one of four exposures—two accelerated exposures with 8 steps of 500 h and two real-world exposures with 6 steps of 2 months. Backsheets were oriented with their outer layer (junction box side) towards the irradiance source. Two replicates per exposure of each backsheet sample were left as unexposed controls equating to 8 controls per sample. After each completed step, two replicates per exposure were removed and saved in a retained sample library. Our study’s analysis was conducted with this library.

Accelerated replicates were exposed to two different ASTM G154-04 cycle 4 exposures [44] using UVA-340 fluorescent ultraviolet lamps (wavelengths 280–400 nm). These lamps closely match the spectral distribution of sunlight up to 365 nm and were set to an intensity of 1.55 W/m²/nm at 340 nm (approximately three times the intensity of the air mass (AM) 1.5 spectrum at 365 nm).

Table 1

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>Layers</th>
<th>Supplier</th>
<th>Thickness (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPE1</td>
<td>PVF/PET/EVA</td>
<td>4</td>
<td>18/140/70</td>
</tr>
<tr>
<td>PPE2</td>
<td>PVF/PET/EVA</td>
<td>8</td>
<td>46/118/114</td>
</tr>
<tr>
<td>PPE3</td>
<td>PVF/PET/EVA</td>
<td>6</td>
<td>44/118/120</td>
</tr>
<tr>
<td>PPE4</td>
<td>PVF/PET/EVA</td>
<td>2</td>
<td>25/256/45</td>
</tr>
<tr>
<td>PPE5</td>
<td>PVF/PET/EVA</td>
<td>2</td>
<td>29/258/45</td>
</tr>
<tr>
<td>PPF1</td>
<td>PVF/PET/PVF</td>
<td>4</td>
<td>19/220/28</td>
</tr>
<tr>
<td>PPF2</td>
<td>PVF/PET/PVF</td>
<td>4</td>
<td>19/232/40</td>
</tr>
<tr>
<td>PPF3</td>
<td>PVF/PET/PVF</td>
<td>5</td>
<td>18/238/40</td>
</tr>
<tr>
<td>PPF4</td>
<td>PVF/PET/PVF</td>
<td>2</td>
<td>35/228/49</td>
</tr>
<tr>
<td>PPF5</td>
<td>PVF/PET/PVF</td>
<td>3</td>
<td>49/248/49</td>
</tr>
<tr>
<td>PPF6</td>
<td>PVF/PET/PVF</td>
<td>1</td>
<td>64/251/49</td>
</tr>
<tr>
<td>PPF1</td>
<td>PVF/PET/PE</td>
<td>12</td>
<td>11/64/203</td>
</tr>
<tr>
<td>PPF2</td>
<td>PVF/PET/PE</td>
<td>1</td>
<td>34/226/26</td>
</tr>
<tr>
<td>PPF3</td>
<td>PVF/PET/PE</td>
<td>7</td>
<td>23/261/49</td>
</tr>
<tr>
<td>PPF1</td>
<td>PET/PET/EVA</td>
<td>10</td>
<td>51/150/53</td>
</tr>
<tr>
<td>PPF2</td>
<td>PET/PET/EVA</td>
<td>10</td>
<td>33/189/68</td>
</tr>
<tr>
<td>PPF1</td>
<td>PET/PET/PE</td>
<td>11</td>
<td>33/104/95</td>
</tr>
<tr>
<td>PPF2</td>
<td>PET/PET/PE</td>
<td>11</td>
<td>24/74/166</td>
</tr>
<tr>
<td>THV1</td>
<td>THV/PET/EVA</td>
<td>9</td>
<td>112/153/74</td>
</tr>
<tr>
<td>THV2</td>
<td>THV/PET/EVA</td>
<td>9</td>
<td>143/139/52</td>
</tr>
<tr>
<td>AAA1</td>
<td>PA/PA/PA</td>
<td>5</td>
<td>51/265/41</td>
</tr>
<tr>
<td>DPD1</td>
<td>PVDF/PET/PVDF</td>
<td>3</td>
<td>31/229/23</td>
</tr>
<tr>
<td>DPE1</td>
<td>PVDF/PET/EVA</td>
<td>3</td>
<td>38/248/96</td>
</tr>
</tbody>
</table>
Both accelerated exposures used Q-Lab QUV accelerated weathering testers [46]:

1. **Hot QUV**: Continuous irradiance and a chamber temperature of 70°C. 
2. **Cyclic QUV**: Cyclic exposure of 8 h of Hot QUV followed by 4 h of darkness and condensing humidity at 50°C.

Real-world exposures were conducted in Cleveland, Ohio between July 2, 2013 and October 7, 2014 on two-axis trackers. The Köppen-Geiger climate zone [47] for this precise location is Cfa, but it closely borders Dfa and Dfb [48].

For ease of comparison with accelerated exposures, real-world photodose values were subset to include only wavelengths 360–246 nm and below (referred to here as photodose values were subset to include only wavelengths 360–246 nm and below). Irradiance data was collected in 1 min intervals from a fixed Kipp & Zonen CMP6 full-spectrum (285–2800 nm) pyranometer [49] supplementing missing values (from instrument and server outages) with 15 min interval Solar GIS data [50].

1. **Real-World 1x**: Natural, full-spectrum solar irradiance and exposure to all weather conditions.
2. **Real-World 5x**: Identical to Real-World 1x with the addition of front surface aluminum mirrors to increase the direct irradiance approximately five times. Depending on the season, time of day, and weather, the measured increase in total irradiance ranged between 1.5 and 4.1.

### 3. Methods

#### 3.1. Measurements

Images of each backsheets’ layers were obtained from unexposed replicates with a PAX-it PAXcam camera [51] and Nikon Epiphot optical microscope at 100x magnification leading to a resolution of 0.675 µm per pixel. Layer thicknesses were extracted from the images using open-source ImageJ software from the National Institute of Health [52,53]. Each layer was measured 10 times and the average is reported in Table 1. The largest standard error among all measurements was 3.25% and the average was 0.94 µm. Fourier transform infrared (FTIR) spectroscopy was used to confirm the inner and outer layer materials [54,55].

A Nanovea ST400 [56] was used to gather surface profilometry data as input to the crack quantification algorithm. The instrument’s lens has a well-characterized axial chromatic aberration designed to focus light of different wavelengths at different z-axis depths [57–59]. Reflected light passes through a spatial filter that only allows the in-focus wavelength through with high efficiency (Fig. 1). The received wavelength can be matched with the known depth of focus to determine the z-coordinate at the current (x, y) location. This particular profilometry technique is insensitive to surface reflectivity, covers a wide range of depths, and can measure steep surface angles [60].

The lens used for measurements had a 1100 µm depth range, 0.09 µm maximum linearity error, and 41 nm static noise at the center of measurement range. Freestanding backsheets are thin and prone to curling and warping at the edges during exposure. A custom metal test jig was machined to hold samples flat against the Nanovea’s central mounting plate (Fig. 2).

#### 3.2. Statistics

Local regression is a non-parametric statistical method where the fitted model at any point is weighted towards the closest data [61,62]. This approach is applicable when the relationship between variables cannot be described by a single functional form and values of the dependent variable are best predicted by their neighbors [63]. Among others, two common variations include LOESS [64] and Friedman’s Super Smoother [65–67]. Both rely on a tuning parameter called the span which specifies the percentage of data used in each local fit. This process is repeated across the range of the independent variable much like a moving average filter in time series analyses [64]. Large spans result in low variance and high bias while small spans have the opposite effect [61,65]. The added flexibility of lower spans makes them more effective at capturing high frequency components in the data.

Unless the variance of the error terms and the frequency spectrum of the underlying data are constant, the optimal span will be different for each x-value [65]. To overcome this issue, Friedman’s variant of local regression uses cross-validation to choose the optimal span parameter for every x-value. This creates more accurate models for data with highly variable frequency spectra and maintains an unsupervised nature by avoiding the need to manually set the span parameter.

Iterative application of local regression can also be used for

---

**Fig. 1.** The operating theory of the Nanovea ST400 profilometer.

**Fig. 2.** Left: The measurement lens and coordinate system for the mounting stage of the Nanovea ST400. Right: An up close view of the metal test jig pinning down a backsheet sample for measurement. The 10 green lines indicate the path of the lens’ light beam. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
outlier detection [68]. After creating an initial trend line with local regression, outliers are identified and replaced with interpolated values. Then, a second, more robust trend line is estimated. This procedure is repeated until no new outliers are found (the green/middle box in Fig. 4), indicating convergence of the trend lines to a final unsupervised, non-parametric model.

4. Crack quantification

4.1. Acquisition of profilometry data

The Nanovea ST400 permits any user-specified grid of \((x, y)\) coordinate measurements provided each axis does not exceed its 150 mm range limit. A good balance between detailed measurements and runtime was obtained using 10 equally spaced “lines” in the y-direction running perpendicular to the primary direction of cracking (Figs. 2 and 3) with each line collecting data every 1 \(\mu\)m in the x-direction. By selecting 10 lines, the spacing in the y-direction was approximately 3.5 mm which allowed the methodology to be robust to local variations in cracking. Approximately half a million data points were measured for each backsheet replicate with each collection process taking 9.5 min. Additional parameters such as averaging multiple measurements for each xy-coordinate or making each x-step less than 1 \(\mu\)m were explored, but the improvement in data quality was not significant enough to justify the additional time required.

4.2. Data ingestion

The algorithm requires raw data in the form of \((x, y, z)\) coordinates as input. This was satisfied with text files from the Nanovea ST400. Files utilized a specific naming convention to help automate the data analysis and streamline ingestion into CRADLE, our Hadoop/HBase/Spark distributed computing cluster [69]. Any z-coordinate beyond the focal range of the lens was represented by a \(-1.0\) in the data file. Each line deliberately started and ended on the jig surface in order to create several series of \(-1.0\)'s in the text file. These series were used to parse the lines from each other so they could be analyzed individually. At this point, y-coordinates were discarded since each line’s value is constant.

4.3. Statistical analysis

The analysis method was written in the programming language R [67,70] with the aid of the following packages: stats [67], tidyverse [71], dplyr [72], ggplot2 [73], stringr [74], cgwtools [75], kgc [48], and forecast [76,77]. Case Western Reserve University’s high performance computing cluster was used to run fleets of SLURM jobs in parallel. The data analysis pipeline is illustrated in Fig. 4.

4.3.1. Outlier removal and surface model

A model of the backsheet’s surface (referred to here as the surface model) is a prerequisite for determining quantitative crack parameters. However, giving weight to data points below the surface (like those from cracks or delamination) artificially biases the model lower than the true surface. Dust particles on the backsheet or blistering can also cause data points to be measured above the surface.

Outlier detection and modeling can be combined into a single step with local regression, thereby decreasing computational time. Friedman’s Super Smoother was selected for the local regression portion of the procedure. After each iteration of the Super Smoother, points that were three interquartile ranges outside the 25 th and 75 th percentiles were marked as outliers [76,77]. Outlier values were linearly interpolated, and iteration continued until no new outliers were detected or the number of iterations exceeded 10. Then, the residuals of each data point relative to the final surface model were calculated.

4.3.2. Crack detection

A preliminary search flags all points with residuals at least 10 \(\mu\)m and two standard deviations below the surface model. The 10 \(\mu\)m threshold was chosen since measurements of unexposed samples (which should have no cracks) commonly had noise of this magnitude. Since each crack typically consisted of multiple data points, they were aggregated into continuous blocks to discern where one crack stopped and another one started.

Manual inspection of the raw data revealed oscillations in the z-values near the edges of some cracks as the Nanovea’s light beam moved across measurement line 2 of 10. Dashed lines in the positive y direction represent future measurement lines. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Fig. 3. Conversion of 3D cracks into 2D profilometry data. The blue xz-plane represents the path of the profilometer’s light beam as it moves across measurement line 2 of 10. Dashed lines in the positive y direction represent future measurement lines. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Fig. 4. A flowchart describing the steps in the crack detection algorithm. The column of solid color boxes with the dashed vertical line shows the major inputs and outputs while the white boxes on the right represent detailed actions manipulating these inputs. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
entered the crack (Fig. 5). Evaluating the same surface at an extremely fine (and time-intensive) x-step value did not produce similar behavior but rather showed one continuous crack. This implies these oscillations are likely an optical effect from multi-reflection and divergence of the light due to the extreme local surface slope. To avoid double counting these oscillations, any neighboring cracks within 10 μm of each other were combined into a single crack.

The weight of the test jig often caused the surface profile at the end of each line to curve downwards sharply. These points typically fell below the surface model and therefore had negative residuals. To prevent them from being incorrectly counted as cracks, only points between the first and last positive residual were considered for crack detection.

4.3.3. Crack parameters

The width, depth, and area of each individual crack can be calculated from the start and end indices determined in the aggregation process. Depth was calculated as the mean depth of all points between the first and last positive residual were also tracked. Once the elementary crack parameters were known for all replicates, the following summary parameters were calculated for each measurement line and written into CRADLE:

- Average crack depth, width, and area (d_{avg}, w_{avg}, and a_{avg}, respectively)
- Minimum, maximum, and average crack spacing (s_{min}, s_{max}, and s_{avg}, respectively)
- Number of cracks (c)

The number of observations used to calculate the average parameters varied depending on the number of cracks detected in each measurement line. To compare backsheet performance between samples and exposures, the normalized crack depth (D_n) and normalized number of cracks (C_n) were also defined (Equations (1) and (2))

\[
D_n = \frac{d_{avg}}{d_UVA_{<360}}
\]

\[
C_n = \frac{c_{avg}}{UVA_{<360}}
\]

where \(d_U\) is the thickness of the backsheet’s inner layer and \(D_n\) and \(C_n\) are calculated for all replicates of a sample with photodose \(UVA_{<360}\). \(C_n\) and \(D_n\) approach the most concerning aspects of cracking from two different angles—its overall extent vs. the damage from a single crack, respectively.

Reference plots of the raw depth data and identified cracks (Fig. 6a, b, and c) are automatically generated for each measurement line and can aid in exploratory data analysis—for instance, closer examination of questionable data points. The surface model residuals can also be plotted instead of depth (Fig. 7, referred to here as a residual plot). This allows easier visual comparison of crack depths and their location within the layered structure of the backsheet (i.e., whether propagation has reached the inner, core, or outer layer).

5. Results

5.1. Case study: sample FPE2

The cracking observed for sample FPE2 was similar between accelerated and real-world exposures. Reference plots generated by the crack quantification algorithm (Fig. 6a and b) illustrate two of the more complex degradation phenomena encountered in this study. Namely, cracking combined with some form of inner layer de-adhesion—either delamination where large-scale loss of adhesion causes pieces of the inner layer to fall off or blistering where localized loss of adhesion permits the inner layer to bubble upwards (or curl if a crack is present). In Fig. 6a, most of the data points associated with the blistering are correctly identified as outliers leading to an accurate surface model. Thus, the calculated crack parameters will match their physical counterparts. In comparison, the delamination between 23 < x < 30 mm in Fig. 6b results in a large trough which is incorrectly handled by the outlier detection procedure. Even though a surface model can be constructed, cracks measured relative to it will not be physically meaningful since part of the “surface” now starts at the inner/core layer interface.

Most observed cracking was substantially less complex than the previous examples and successfully handled by the crack quantification algorithm. For instance, the Real-World 5x exposure of sample FPE2 resulted in thin, parallel cracks without any delamination (Fig. 6c).

5.2. Capabilities and limitations

251 sample replicates were selected for detailed measurements on the optical profilometer which in total took approximately 50 h. From this raw data, the algorithm detected 52,139 cracks, generated the reference and residual plots for each line, and automatically computed the seven quantitative crack parameters—all in only 6 min with the use of CRADLE’s high performance and distributed computing cluster. Thus, parameters were measured at a rate of

![Fig. 5. Zoomed view of a single crack showing how the depth and width parameters were determined. The circled data point belongs to one of the oscillations described in Section 4.3.2.](image-url)
17.3 cracks per minute when accounting for all time involved (i.e. both human supervised and automated). Our algorithm’s combined outlier detection and surface model procedure converged quickly, taking an average of 2.7 iterations and a maximum of 5. Using an automated algorithm kept crack detection unbiased and was able to utilize 99.0% of the raw data on average (points not used were removed for reasons mentioned in Section 4.3.2).

Certain forms of cracking and degradation were able to be distinguished from parallel cracks based on their profilometry data but could not accurately be quantified by the algorithm.

1. **De-adhesion:** For a typical sample without de-adhesion (Fig. 6c), the density distribution of measured depths has a single, prominent peak related to points on the surface since the cumulative width of all cracks is much smaller than the total measured width. Both forms of de-adhesion lead to more pronounced deviations from this unimodal distribution. As peak magnitude values become closer to each other, the outlier detection algorithm approaches a threshold value where it can no longer identify which peak corresponds to the true surface.

---

**Fig. 6.** Several of the reference plots generated for sample FPE2 showing the variety of observed cracking (blue squares) and de-adhesion. (a) The crack quantification program successfully handled combined blistering and cracking (most prominent between $7 < x < 22$ mm). (b) Combined delamination and cracking ($23 < x < 30$ mm) lead to poor outlier detection and an error in the surface model (blue line). (c) The parallel cracks and consistent surface were easily detected by the algorithm. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
2. Mudflat Cracks: When cracks are parallel, it is possible for each measurement line to run perpendicular to them. However, this is no longer feasible for mudflat cracks whose random orientation can cause the light beam to remain inside and parallel to one crack for an extended amount of time. Although the root cause is different, mudflat cracks result in similar crack depth density distribution issues as those observed for de-adhesion.

5.3. Validation of methodology

Two common situations for false detections were cracks with a positive depth (i.e. above the backsheets' surface) and cracks detected on unexposed samples. Out of the 52,139 cracks detected, these two scenarios occurred only 0.004% and 1.1% of the time, respectively. These cases were removed using automated filters for cracks with a photodose of zero or a positive depth. Other issues such as errors in surface model are best observed by manually examining the reference plots generated by the R code. For those cases, 2.9% of the 2501 measurement lines were removed. Therefore, the total percentage of false cracks and surface model errors was approximately 4%. If the backsheets measured in this study are a representative population, this suggests over 96% of the cracks detected do in fact physically exist and are not simply artifacts of the measurement or analysis technique.

As a separate measure of validity, cracks were manually counted for each replicate of Hot QUV samples FPP1 and FPP3 and compared against the algorithm’s output (Fig. 8). These samples were most conducive to manual counting since cracks were spaced reasonably far apart and had minimal branching. Some variation in results is observed since each line from the algorithm crosses branching cracks at different locations.

The error in individual crack depth and width was estimated by repeatedly measuring the same location on a given backsheet. Samples FPP1, FPP3, and PPE1 were chosen since they included cracks with a wide range of depth, width, and spacing. Three replicates were randomly chosen from each sample to represent backsheets from various stages of exposure. In total, 220 different cracks were measured 10 times each. The mean depth and width of each crack was determined along with its associated standard error. Then, the raw and percent error of each individual measurement relative to its expected mean was calculated. Averaging across all 220 cracks, depth and width measurements had errors of \( \pm 3.4 \, \mu m \) (4.4%) and \( \pm 2.7 \, \mu m \) (9.6%), respectively. Both errors were higher for narrower cracks, likely because fewer data points were available for the algorithm to determine the crack depth and width.

5.4. Cracking by material and exposure

Of the 23 backsheets in this study, 11 showed some form of
cracking (Table 2). All cracks occurred on the inner (cell side) layers and had varied depth (10–292 μm), width (10–1322 μm), and inter-crack spacing (10–46.238 μm). For some samples, cracks were highly localized to a particular corner of the backsheet that had presumably curled closer to the UVA-340 lamps during exposure. These non-uniformities were still able to be characterized by virtue of measuring multiple lines on each replicate. PPP and THV samples were measured in the early stages of developing the crack quantification methodology. However, comprehensive measurements of all replicates were not made for those materials after discovering the algorithm’s difficulty in characterizing mudflat cracks and delamination.

This methodology successfully quantified changes in crack parameters throughout exposure. The Cyclic QUV exposure of sample PPE2 illustrates this capability in Table 3. Between 500 and 4000 h of exposure the average number of cracks increased from 4 to 63. Additionally, Table 4 lists the ranked Dn and Cn values across all measured samples at the final step of exposure (either 4000 h or 12 months).

Crack depth is of primary interest as a parameter since deeper cracks—in particular those penetrating the core PET layer—lower the backsheet’s effectiveness as an electrical and environmental barrier. The residual plots generated by the crack quantification program (Fig. 7) are a useful diagnostic tool in determining the severity of cracking. The replicate of sample PPE1 shown displays cracks at two distinct depths relative to the layers of the backsheet. Similar behavior was also observed for other samples including the PPE2 reference plots shown earlier. For instance, the 4 cracks below the delaminated portion of the inner layer (23 < x < 30 mm) in Fig. 6b and the second crack from the left in Fig. 6a.

6. Discussion

6.1. Algorithm performance

The case study of sample FPE2 shows that our crack quantification algorithm is most applicable for characterizing parallel cracks with minimal de-adhesion (Fig. 6c). In those cases, the algorithm proved to be a fast and accurate method for measuring a variety of quantitative variables. Compared to a manual counting procedure, less bias is introduced and far more data can be generated—both of which make this method particularly valuable for a data science approach to studying polymer degradation.

This analysis routine could in theory be applied to data from any instrument that records high resolution (x, y, z) coordinates. The materials studied could be any component whose characteristic surface roughness is substantially smaller than the deviations caused by cracks. This method shows promise to generalize across laboratories with different equipment and fields of research.

The principal advantage of our approach over more traditional methods like XCT or LSCM is the ability to gather high resolution information over large areas and simultaneously analyze tens of thousands of cracks. XCT and LSCM achieve higher resolution, but at the cost of extremely small physical sample sizes [41]. This limits their analyses to small numbers of cracks which may not provide sufficient statistical power for predictive lifetime models. However, XCT and LSCM could prove valuable as secondary tools to investigate the findings from our algorithm in greater detail.

6.2. Materials cracking performance

Four different backsheet configurations did not crack in any exposure: FPF, AAA, DPD, and DPE. Since FPF was the most common backsheet configuration used in our study, its resistance to cracking is considered the highest. It is particularly noteworthy that sample AAA1 did not crack in this study since field retrieved modules with polyamide backsheets have consistently shown this issue [21,78]. Polyamide backsheets have consistently shown this issue [21,78].

Physical degradation as evidenced by decrease in strain at break has also been noted for polyamide and PET backsheets [79,80]. Polyamide absorbs and releases large amounts of moisture depending on the humidity of the surrounding environment. Although the precise moisture retention varies with the type of polyamide (i.e., nylon 6, 6-6, or 12), this behavior induces a cyclic stress. Therefore, exposures with changing humidity levels (i.e. Cyclic QUV and Real-World 1x and 5x) would be expected to cause cracking. The lack of physical constraint from a module frame could also explain why the freestanding configuration used in this study did not crack. In a PV module, thermal expansion of the backsheet is primarily constrained by the glass frontsheet with additional influence from the cells and interconnects.

Of materials that did crack and were measured, the Hot QUV sample of FPE1 performed the best in terms of both Dn and Cn (Table 4). In contrast, the Cyclic QUV sample of PPE1 had the poorest performance. The mudflat cracks on PPP and THV samples...
were unable to be quantified, however this is still considered poor performance. Using material-wide averages of $D_n$ and $C_n$ gives an overall material performance rating, again with lower numbers indicating better performance. Based on $D_n \text{ avg}$ the order from best to poorest performance was FPP ($1.43 \times 10^{-3}$ MJ/m$^2$), FPE ($1.56 \times 10^{-3}$ MJ/m$^2$), and PPE ($2.08 \times 10^{-1}$ MJ/m$^2$). The order of performance remains the same with $C_n \text{ avg}$: FPP ($3.88 \times 10^{-1}$ MJ/m$^2$), FPE ($5.04 \times 10^{-1}$ MJ/m$^2$), and PPE ($6.94 \times 10^{-1}$ MJ/m$^2$).

### 6.3. Effect of different exposures

All samples which cracked in a real-world exposure also cracked in both accelerated exposures (Table 2). The only sample to crack in all four exposures was PFE2. For the three samples which cracked in the Real-World 5x exposure, $D_n$ values were closely mirrored by both accelerated exposures with average differences of 0.26 and 0.37 (10$^{-2}$ MJ/m$^2$) for Cyclic and Hot QUV, respectively. However, accelerated $C_n$ values typically far exceeded those of real-world exposures with average differences of 9.48 and 18.2 (10$^{-2}$ MJ/m$^2$) for Cyclic and Hot QUV, respectively. This discrepancy highlights the importance of exposing backsheets to real-world weather conditions encompassing multiple seasons, full spectrum irradiance, and simultaneous environmental stressors.

Averaging $D_n$ and $C_n$ values from Table 4 by exposure instead of material shows that Cyclic QUV formed more cracks per UVA<360 dose ($8.84 \times 10^{-2}$ MJ/m$^2$) than Real-World 1x ($4.08 \times 10^{-2}$ MJ/m$^2$), Hot QUV ($2.94 \times 10^{-2}$ MJ/m$^2$), and Real-World 5x (1.92). When ranked by $D_n \text{ avg}$, Real-World 1x was substantially more damaging: Real-World 1x (3.99 (10$^{-3}$ MJ/m$^2$)), Cyclic QUV (1.76 (10$^{-3}$ MJ/m$^2$)), Real-World 5x (1.63 (10$^{-3}$ MJ/m$^2$)), and Hot QUV (1.21 (10$^{-3}$ MJ/m$^2$)). Since outdoor photodose data came from fixed sensors while samples were on two-axis trackers, the true difference in $D_n \text{ avg}$ may be lower.

The degradation caused by Real-World 1x (measured by $D_n \text{ avg}$ and $C_n \text{ avg}$) was greater than Real-World 5x which is opposite of expected. Most likely, this is because the only Real-World 1x sample measured was PFE2 which was among the top four poorest performing samples. In comparison, Real-World 5x caused three times more samples to crack and would therefore have a more statistically valid average. It took samples FPP3 and PPE2 12 months to crack in Real-World 5x. If irradiance is the only contributing stressor, we would expect cracks to also occur in Real-World 1x if those replicates were exposed further. However, the temperature increase from the concentrating mirrors in Real-World 5x might also play a role, which could explain why cracks develop slower in Real-World 1x.

### 7. Conclusions

A non-destructive analysis technique to quantify cracks from surface profilometry data has been developed. The algorithm is robust to local variations in cracking and can accurately process much higher volumes of data than any manual inspection. Particular value is realized in the PV industry since backsheets with parallel cracks and minimal de-adhesion are well-suited for this approach. FPP samples had the best resistance to cracking while PPE samples had the least. On a photodose basis, the addition of humidity and temperature variation caused more cracks to form by a factor of 3. Cracks in real-world and accelerated exposures propagated to similar depths with equivalent photodoses; however, the number of cracks formed in accelerated exposures was far greater on a photodose basis.

Wider use of this analysis technique can help address the current shortage of quantitative data on backsheet cracking. Additional quantitative data will enable more accurate predictive models, thus helping to prolong the lifetime of PV modules. Insights from the algorithm’s crack parameters could also narrow the scope of future degradation studies, paving the way for more in-depth analyses of particular aspects of backsheet cracking. On the whole, developing quantitative analysis methods is of vital importance for the PV scientific community.

### Acknowledgments

This research made use of the following core facilities and research centers at Case Western Reserve University: the Solar Durability and Lifetime Extension (SDLE) Research Center (Ohio Third Frontier, Wright Project Program Award Tech 12-004), the Materials for Opto/electronics Research and Education (MORE) Center (Ohio Third Frontier grant TECH 09-021), and the High Performance Computing Resource in the Core Facility for Advanced Research Computing. The authors would also like to acknowledge support from the Department of Energy’s SunShot PREDICTS2 program (DE-EE0007143).

### Table 4

Comparing the performance of different measured backsheet samples (non-italic samples from Table 2) at the final step of each exposure. A rank of 1 represents the best performance, while 17 is the worst.

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>Exposure</th>
<th>$D_n$ ($10^{-3}$ MJ/m$^2$)</th>
<th>$C_n$ ($10^{-3}$ MJ/m$^2$)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>FPE1</td>
<td>Cyclic QUV</td>
<td>1.32</td>
<td>8.00</td>
<td>6</td>
</tr>
<tr>
<td>FPE2</td>
<td>Real-World 1x</td>
<td>3.99</td>
<td>4.08</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Real-World 5x</td>
<td>1.17</td>
<td>4.23</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Cyclic QUV</td>
<td>1.48</td>
<td>10.5</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Hot QUV</td>
<td>0.98</td>
<td>3.00</td>
<td>2</td>
</tr>
<tr>
<td>FPE3</td>
<td>Cyclic QUV</td>
<td>1.46</td>
<td>4.98</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Hot QUV</td>
<td>0.55</td>
<td>0.44</td>
<td>1</td>
</tr>
<tr>
<td>FPP1</td>
<td>Cyclic QUV</td>
<td>1.49</td>
<td>2.19</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Hot QUV</td>
<td>0.59</td>
<td>0.57</td>
<td>3</td>
</tr>
<tr>
<td>FPP3</td>
<td>Real-World 5x</td>
<td>1.67</td>
<td>0.85</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Cyclic QUV</td>
<td>1.84</td>
<td>12.9</td>
<td>12</td>
</tr>
<tr>
<td>PPE1</td>
<td>Cyclic QUV</td>
<td>2.41</td>
<td>12.5</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Hot QUV</td>
<td>1.91</td>
<td>5.40</td>
<td>13</td>
</tr>
<tr>
<td>PPE2</td>
<td>Real-World 5x</td>
<td>2.07</td>
<td>0.68</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Cyclic QUV</td>
<td>2.35</td>
<td>10.8</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Hot QUV</td>
<td>1.63</td>
<td>5.30</td>
<td>10</td>
</tr>
</tbody>
</table>


J.H. Friedman, A Variable Span Smoother, S. Stanford Univ CA Lab for Computational Statistics, Nov. 1984,


H. Wickham, RStudio, tidyverse: easily install and load ‘tidyverse‘ packages, RStudio, r package version 1.1.1. https://cran.r-project.org/web/packages/tidyverse/index.html,

H. Wickham, R. Francois, L. Henry, K. Miller, RStudio, dplyr: a Grammar of Data Manipulation, RStudio, r package version 0.7.3. https://cran.r-project.org/web/packages/dplyr/index.html,


H. Wickham, RStudio, stringr: simple, consistent wrappers for common string operations, RStudio, r package version 1.2.0. https://cran.r-project.org/web/packages/stringr/index.html, 2017,

C. Wirthroff, cgplotools: miscellaneous Tools, r package version 3.0. https://cran.r-project.org/web/packages/cgplotools/index.html,


