Evaluation of variations in solar cell production lines using time-series electroluminescence images

Gaia Maria Javier, Rhett Evans, Priya Dwivedi, Yoann Buratti, Thorsten Trupke, Ziv Hameiri

The University of New South Wales, Sydney, Australia

**E-mail address: g.javier@unsw.edu.au*

Context and Motivation

Fabrication of solar cells requires several process steps.[1] Manufacturers must ensure the consistency of each of these steps so that each solar cell satisfies a design specification.[2] However, variations between solar cells are inevitable. These variations can originate from the quality of raw materials, processing under slightly different conditions, and changes in the external environment.[3] Ideally, the mean performance of solar cells fabricated in a production line is constant, thus, the cell-to-cell variance is expected to be *minimal* and *random*. Minimal and random variations are preferred as they do not have apparent long-term effects on production.

Electroluminescence (EL)[4] and photoluminescence (PL)[5] imaging are common end-of-line characterisation tools in assessing the performance of solar cells. These techniques capture spatial information from solar cells to highlight defective areas.[6] Previous studies have also used luminescence images to predict the critical electrical parameters of the solar cells.[7–10] This study presents a new capability based on these images. We demonstrate a novel technique to determine production line variations using time-series EL images.

Proposed Methodology

The random variability of the time-series EL images is determined through lag-sequential analysis. The concept of this technique is illustrated in Figure 1. For example, a dataset where x_0, $x_1 \ldots x_n$ correspond to measurements of the 0th, 1st…nth solar cell is presented. The original dataset (denoted as lag-0) is shifted to generate lag datasets as shown by the blue circles. Lag-1 corresponds to the original dataset shifted by one timestamp. Similarly, lag-2 is shifted by two timestamps, and lag-n is shifted by n timestamps. In this study, n ranges from 1 to 2,000. These lag datasets are then subtracted from the original dataset to produce the lag-differenced datasets. As an example, in Figure 1, the lag-1 differenced dataset is represented by the orange circles. Through the lag-differenced datasets, the cell-to-cell variance between various timestamps can be determined.

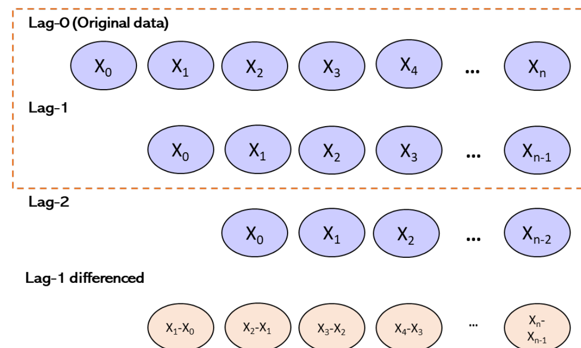

Figure 1. Visual illustration of lag-sequential analysis

We then define a variable that we named the ‘Rhett Factor’ (RF) to quantify the random variability of the dataset:

$$\mathrm{RF(n)} = \frac{\sigma^2(X_{\mathrm{lag_n,diff}})}{2*\sigma^2(X)}, \quad (1)$$

where $\sigma^2(X_{lag_n,diff})$ is the variance of the lag-differenced dataset at lag-n and $\sigma^2(X)$ is the variance of the original (lag-0) dataset. Note that $\sigma^2(X)$ is used in the denominator because it is assumed that the maximum possible variance in the datasets is the population variance. The data is statistically independent at a given lag if its RF is equal to unity. In this case, the lag's mean is the same as the mean of lag-0 and the variations at this lag are random.

Proof of Concept

To demonstrate the method, we generated 10,000 time-series measurements of cells with a mean efficiency of 23% and a standard deviation of 0.15%. Figure 2(a) shows the randomly generated Gaussian distribution of the data as a function of time. The resulting RF-vs-lag is shown in Figure 2(d). The lag where RF reaches unity (in this presented case, lag-1) is the 'statistical batch point'. The statistical batch point defines the batch size where samples are statistically independent of one another. A lower statistical batch point is preferred in a production line as it indicates that the variance among cells is more random. Since the data is completely random in the simulation, RF reaches unity already from lag-1 (and stays constant).

We then simulated measurements with a varying mean for every 100 cells as shown in Figure 2(b). The corresponding RF-vs-lag is presented in Figure 2(e). RF reaches unity only at around lag-100, hence, in this case, the statistical batch point is 100. Further analysis is provided by dividing the graph into two: (1) a yellow region before the statistical batch point, and (2) a red region after the statistical batch point. The yellow region shows the degree of random variability at smaller lags while the red region provides a measure of stability in the data. At lag-50 (yellow region), the RF is around 0.75. This means that for every batch of 50 cells, 75% of the variance is random whereas 25% (the remainder) of the variance is related to mean shifts. Higher RF (>0.9) in the yellow region is preferred as it indicates that a lower percentage of variance is related to mean shifts. Above lag-100 (red region), large deviations from unity are observed, thus, indicating high instability in the data. This is expected since this dataset was generated by randomly shifting the mean for every batch of 100. Small fluctuations in the red region are preferred as they represent a more stable production line (fewer changes in the mean).

Lastly, Figure 2(c) illustrates simulated measurements with outliers. As can be seen in Figure 2(f), although RF reaches unity at lag-1, the outliers introduce clear variations ('dips'). These dips could affect the interpretation of RF. Thus, to avoid misleading interpretations, it is recommended to discard outliers in further analysis, as we are interested in the randomness of the main distribution.

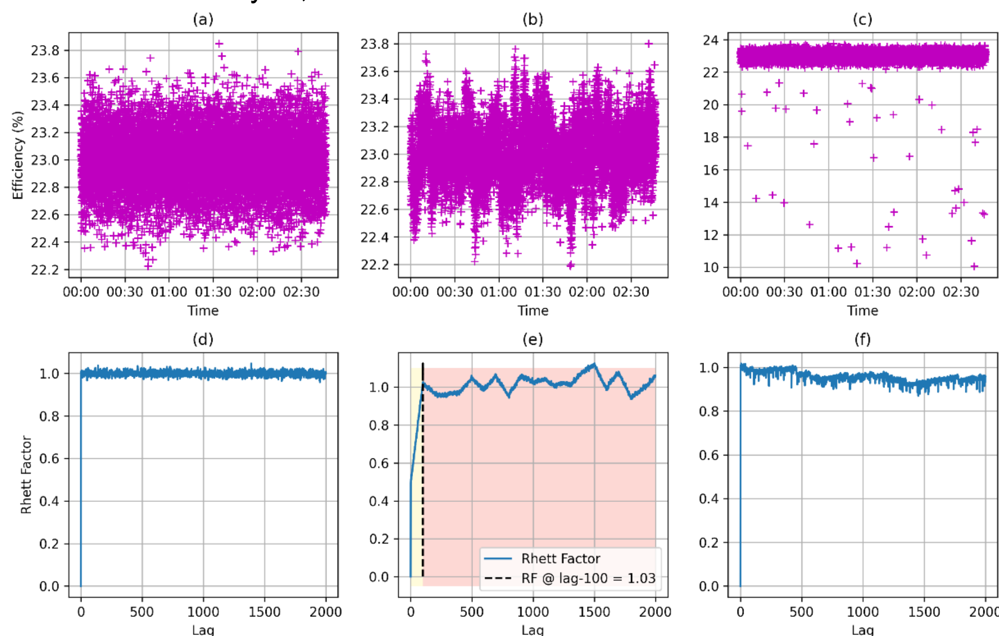

Figure 2. Simulation results: (a) Randomly generated efficiency data vs time, (b) efficiency data with a varying mean for every 100 cells vs time, (c) efficiency data with outliers vs time, (d-f) corresponding Rhett Factors vs lag of graphs (a-c), respectively

Experimental Results

A total of 17,000 EL images of multi-crystalline silicon (mc-Si) cells were collected from an industrial production line. These images are ordered according to their timestamps and are quantified based on two metrics: (1) the mean of the pixels, and (2) the standard deviation of the pixels. For simplicity, the outliers (identified via the modified z-score method[11]) were excluded in the computation of the RF.

Figure 3(a) presents the normalised EL mean intensity (transformed to mean=0 and variance=1) vs time. Around 88% of the cells are within the main distribution (0±1.5 region). Approximately 8% of the cells fall below the main distribution indicating that these cells are relatively darker (with lower intensity). Figure 3(b) shows the computed RF as a function of lag. Here, lag-600 (where RF reaches 1) is the statistical batch point. Hence, the variance is completely random within batches of 600 cells and each batch is statistically independent of one another. At lag-250, the RF is around 0.92 meaning 92% of the variance is random for batches of 250 cells while 8% of the variance reflects shifts in the mean. From lag-600 to lag-1250, the RF increases while it decreases from lag-1250 to lag-2000. This unstable RF indicates shifts in the mean EL counts, indicating possible shifts in the efficiency (or at least the obtained voltage).

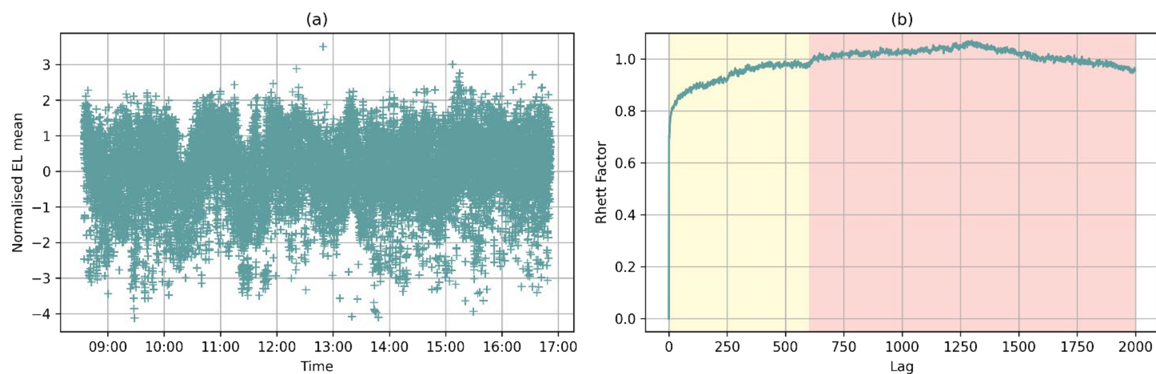

Figure 3. (a) Normalised EL mean intensity vs time, (b) Rhett Factor vs lag

Figure 4(a) shows the normalised open-circuit voltage (V_{oc}) as a function of time of the same dataset. Around 8% of the cells fall below the 0±1.5 region, indicating that these cells have a lower V_{oc} than the main distribution. Figure 4(b) shows the corresponding RF-vs-lag graph. A similar trend as the RF-vs-lag graph for the EL mean intensity can be observed. The statistical batch point is also at lag-600 and instability is noticed in the red region. Further analysis shows that the EL mean intensity is 95% correlated with the measured V_{oc} of the cells. Hence, the EL mean intensity can be used to investigate the V_{oc} variability in production lines.

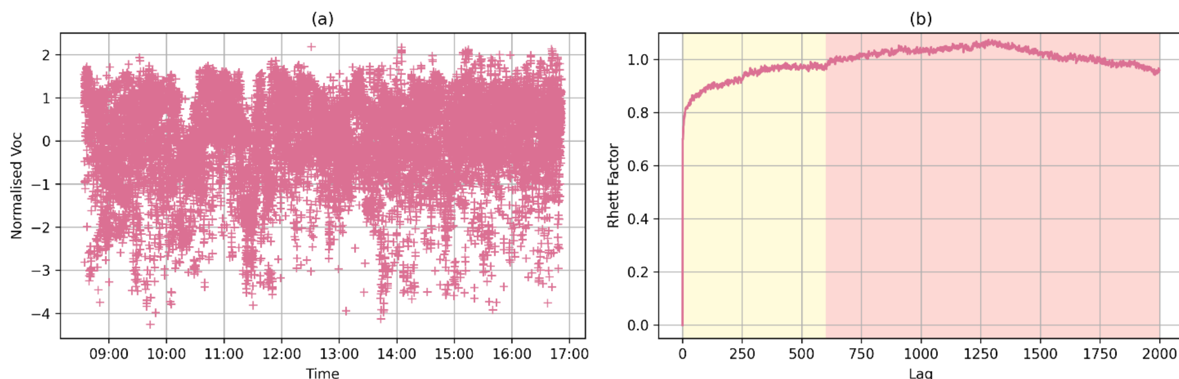

Figure 4. (a) Normalised V_{oc} vs time, (b) Rhett Factor vs lag

We also explored the standard deviation of the EL images as a possible metric. Although the EL mean intensity and EL standard deviation are two different image metrics, the corresponding RF-

vs-lag graphs are comparable. Hence, the EL standard deviation could also be used to investigate production lines.

Outlook

In this study, we proposed a simple and novel technique to investigate variations that can aid manufacturers in assessing and improving the quality of their production lines. Time-series EL images are used to identify random changes in an actual production line. Spatial information from EL images is translated to one-dimensional metrics and lag-sequential analysis is implemented. Using the Rhett Factor, a statistical metric based on variances, the degree of randomness and mean shifts in production lines are determined. Detailed analysis of additional EL image metrics related to solar cell performance will be presented at the conference. Future work will also include identifying components of variation to help pinpoint the sources of non-randomness in production.

Acknowledgement

This work was supported by the Australian Government through the Australian Renewable Energy Agency [ARENA Project 2020/RND016]. The views expressed herein are not necessarily the views of the Australian Government, and the Australian Government does not accept responsibility for any information or advice contained herein.

References

1. “Solar cell production: From silicon wafer to cell.” https://sinovoltaics.com/solar-basics/solar-cell-production-from-silicon-wafer-to-cell/.
2. S. Kurtz *et al.*, “Marrying quality assurance with design engineering – A winning partnership! But, a cultural divide?,” in *IEEE 44th Photovoltaic Specialist Conference (PVSC)*, Washington, DC, pp. 1275–1279, 2017.
3. M. N. Sinha and W. W. Willborn, *The management of quality assurance*, vol. 1. Wiley, 1985.
4. T. Fuyuki, H. Kondo, T. Yamazaki, Y. Takahashi, and Y. Uraoka, “Photographic surveying of minority carrier diffusion length in polycrystalline silicon solar cells by electroluminescence,” *Appl. Phys. Lett.*, vol. 86, no. 26, p. 262108, 2005.
5. T. Trupke, R. A. Bardos, M. C. Schubert, and W. Warta, “Photoluminescence imaging of silicon wafers,” *Appl. Phys. Lett.*, vol. 89, no. 4, p. 044107, 2006.
6. H. R. Parikh *et al.*, “Solar cell cracks and finger failure detection using statistical parameters of electroluminescence images and machine learning,” *Appl. Sci.*, vol. 10, no. 24, p. 8834, 2020.
7. B. Hallam, B. Tjahjono, T. Trupke, and S. Wenham, “Photoluminescence imaging for determining the spatially resolved implied open circuit voltage of silicon solar cells,” *J. Appl. Phys.*, vol. 115, no. 4, p. 044901, 2014.
8. T. Trupke, E. Pink, R. A. Bardos, and M. D. Abbott, “Spatially resolved series resistance of silicon solar cells obtained from luminescence imaging,” *Appl. Phys. Lett.*, vol. 90, no. 9, p. 093506, 2007.
9. Y. Augarten *et al.*, “Calculation of quantitative shunt values using photoluminescence imaging,” *Prog. Photovolt. Res. Appl.*, vol. 21, no. 5, pp. 933–941, 2013.
10. Y. Buratti, A. Sowmya, R. Evans, T. Trupke, and Z. Hameiri, “Half and full solar cell efficiency binning by deep learning on electroluminescence images,” *Prog. Photovolt. Res. Appl.*, vol. 30, no. 3, pp. 276–287, 2022.
11. V. Jain, “Anomaly detection by modified z-score,” *Analytics Vidhya*, 2020. https://medium.com/analytics-vidhya/anomaly-detection-by-modified-z-score-f8ad6be62bac.