Machine Learning for Optimization of Mass-Produced Industrial Silicon Solar Cells

Hannes Wagner-Mohnsen[1,2] and Pietro P. Altermatt[3]

[1]Dep. Solar Energy, Inst. Solid-State Physics, Leibniz University of Hannover, Appelstr. 2, 30167 Hannover, Germany
[2]WAVELABS Solar Metrology Systems GmbH, Spinnereistr. 7, 04179 Leipzig, Germany
[3]Trina Solar, State Key Laboratory for Photovoltaic Science and Technology (SKL), No 2 Trina Road, Xinbei District, Changzhou, Jiangsu Province, China 213031

Abstract— **We present a methodology where we combine numerical TCAD device modeling, machine learning and advanced statistics for getting a deeper understanding of how process variations influence device performance in mass produced crystalline silicon solar cells. For this, we use seven model input parameters that affect the mainstream solar cell design (PERC) and its performance the most and perform about a couple of hundred numerical TCAD device simulations in an expected range of these parameters. As such detailed numerical simulations take long time, we train and validate machine learning models on these simulations, which serve to describe ten thousands of fabricated PERC cells. The method gives concrete information for improving PERC cells with a modest amount of numerical modeling and hence in a very short time. This approach is not limited to a specific solar cell design or product.**

Keywords—machine learning, mass-production, TCAD, PERC, Solar Cells

I. Introduction

In mass fabrication of silicon solar cells, statistical fluctuations and variations in the manufacturing tools lead to scattering of resulting cell performance. Analyzing how these relate to each other [1, 2], gives two useful insights. First, a better process control by evaluating which variations in which manufacturing tools lead to the most performance degradation, and second, indications of how to increase the median cell performance. Generally, process control can be improved by either reducing entropy, for example by tracking each wafer through the production line, or by maximizing entropy, for example by shuffling the wafers so a statistically identical set of wafers goes through the fabrication lines, and underperforming tools can then be located by statistical analysis. Most cell manufacturers nowadays do not track the wafers through their production lines because this is regarded an expensive way of process control [3]. Instead, each manufacturing tool is monitored separately to keep the tool parameters within specified limits. This gives some extent of process control, but it is not known at what time which wafer went through which tool. This means that no direct learning is possible for how to reduce scattering of cell performance and how to improve the median cell performance.

In this paper, we create a digital twin of produced cells by numerical TCAD device simulation, and we find reasons for underperforming or best performing cells in terms of *device* parameters rather than parameters in manufacturing *tools*. Once the responsible device parameters are known by the analysis, it is straight forward to conclude to the behavior of fabrication tools because PERC fabrication consists of only 9 steps. Because a specific device performance can be attained with different sets of device parameters, we apply a machine learning algorithm for categorizing these possible sets. From such categories, we deduce which manufacturing tools are the most likely reasons for underperforming or best performing cells.

II. Methodology

We use Sentaurus TCAD to model the PERC solar cell. We vary seven simulation input parameters that effect device performance most.

Each detailed Sentaurus simulation take approx. 30 min, too long to describe massive data sets. For example, a single work shift produces about 37’000 cells on a line, and Trinasolar produced 2.5 billion cells last year. We therefore use only about 500 Sentaurus simulations as training for a machine learning algorithm. The trained machine learning algorithm is then very quick to calculate almost the same output (e.g. Efficiency η or open-circuit voltage V_{oc}) with the same seven input parameters. To test the approximation quality of these models, we perform a 5-fold cross validation and observe the smallest root-mean-square error (RMSE). In our case a Gaussian process regression model (with rational quadratic kernel function) has the lowest RMSE values of 0.026 mA/cm2 for the short-circuit current density J_{sc}, 0.372 mV for V_{oc}, 0.017% for the fill factor of the IV curve, FF, and 0.018% for the energy conversion efficiency η.

The trained machine learning model is then used to generate a massive data set of 7 x 10^9 random combinations of the seven input parameters. From this data-set, optimal parameters are derived to describe each solar cell by using Euclidean distances.

III. RESULTS

Figure 1 shows the energy conversion efficiency distribution from a typical, new PERC cell factory in China. The blue symbols are the IV parameters of about 32,000 cells from a single labour shift in a mass production line. The machine learning (ML) results of the 10% cells with the lowest, medium and best efficiency η are shown in the traffic light colours.

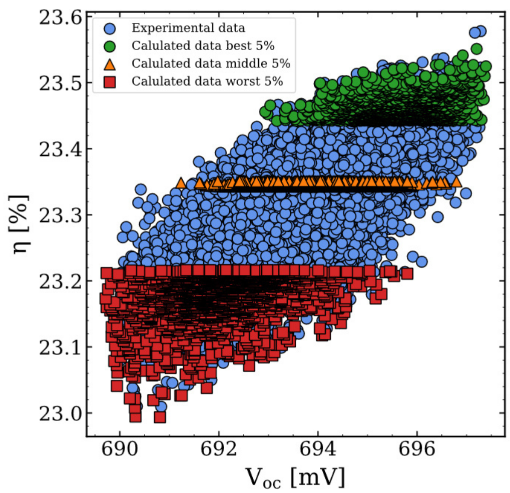

FIG. 1. 32’000 experimental IV parameters of a typical new PERC factory in China (blue symbols), with the machine learning (ML) results of the 10% cells with the lowest, medium and best efficiency η are shown in the traffic light colors.

Figure 2 shows that the most efficient cells (green) can of course only be achieved with high lifetime (low interstitial iron concentration, Fe_i), but interestingly also with the lowest emitter sheet resistivity (R_{sheet}). This implies that for further efficiency improvements, the front finger pitch must be reduced (more fingers added per cell), as confirmed by a cross-check with low FF values (not shown here).

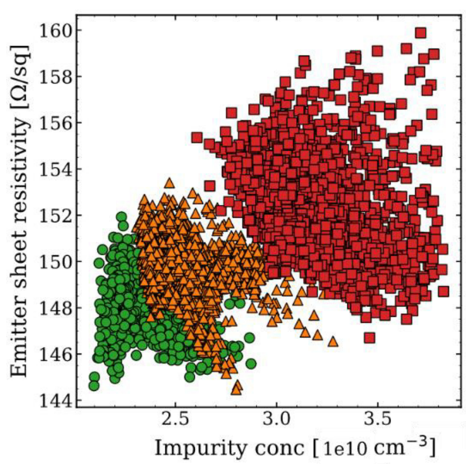

FIG. 2. Corresponding simulation input parameters to modelled data points in Figure 1, from machine learning, of the emitter sheet resistivity over the concentration of interstitial iron, limiting the excess carrier lifetime. Same colors as in Fig. 1.

Both parameters influence for collection efficiency of minority carriers, and therefore the cells quantum efficiency.

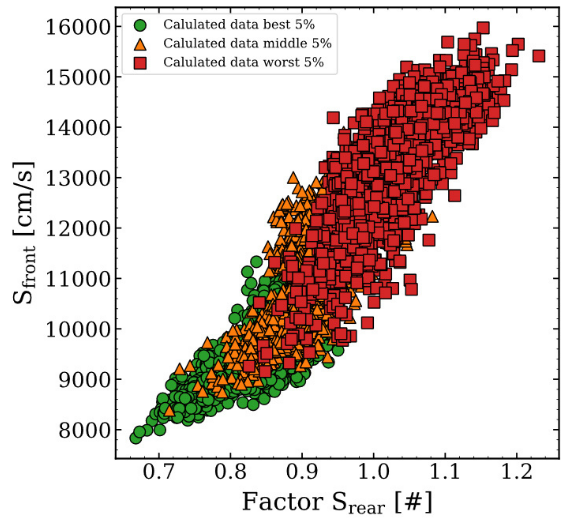

FIG. 3. Corresponding simulation input parameters to modelled data points in Figure 1, from machine learning, of the SRH recombination velocity at the front surface over the rear surface. Same colors as in Fig. 1.

It may follow from Fig. 2 that emitters with higher R_{sheet} may not have sufficient surface passivation. This is refuted in Figure 3, which indicates that good front (and rear) surface passivation is achieved in high-efficiency cells.

IV. CONCLUSION AND OUTLOOK

A methodology is presented to understand and minimize the efficiency variations in the mass production of PERC solar cells. We do this with end-of-line IV parameter data alone, i.e. without knowing at what point the cells went through which manufacturing tools, which in turn have their own variations. In mass production, it is important to monitor whether the front finger pitch and R_{sheet} of the emitter are well tuned, and to what extent variations in R_{sheet} cause variations in efficiency. Equally important is to monitor that the front and rear passivation is maintained at high levels, even when the R_{sheet} of the emitter is high. Without this methodology, experimental series within the mass production would need to be performed to find this out.

REFERENCES

[1] S. Wasmer and B. Klöter, “Interpretable machine learning from production optimization”, 36th European Photovoltaic Solar Energy Conference, Marseille, France (2019), p. 272 – 274.

[2] R. Evans and M. Boreland, "Multivariate Data Analytics in PV Manufacturing - Four Case Studies Using Manufacturing Datasets," IEEE Journal of Photovoltaics, vol. 8, no. 1, pp. 38-47, 2018.

[3] P. P. Altermatt, “A method for optimizing PERC cells in industrial production lines using final IV parameters, statistical procedures and numerical device modeling”, 8th Silicon PV, AIP conference proceedings, 1999, 11000