# About the Relevance of Defect Features in As-Cut Multicrystalline Silicon Wafers on Solar Cell Performance

Aditya Kovvali, Matthias Demant, Theresa Trötschler, Jonas Haunschild and Stefan Rein

*Fraunhofer Institute for Solar Energy Systems ISE, Heidenhofstrasse 2, 79110 Freiburg, Germany, Phone +49 761/4588-5488*

Corresponding author: aditya.kovvali@ise.fraunhofer.de

**Abstract.** Recombination-active defects e.g. dislocations in multicrystalline silicon (mc-Si) wafers impact the quality of solar cells. These defects can be quantified during the incoming control of silicon wafers with Photoluminescence (PL) imaging and used to rate the solar cell quality. In this work, we analyze the relevance of defect patterns in PL images and grain-boundary (GB) data for current-voltage (IV) prediction by means of image processing algorithms. Based on a large set of empirical data of passivated emitter and rear cells (PERC), a sparse prediction model is trained for each IV-parameter. Our results include both, the prediction of different quality parameters and the relevance of the features extracted from PL images and GB images. We achieve mean absolute prediction errors as low as 2.72 mV and 0.18 $mA/cm^2$ for open circuit voltage ($V_{oc}$) and short circuit current density ($J_{sc}$) respectively, and 0.18% for efficiency as combined parameter. In this evaluation, the wafer data set is split into training group and test group. Therefore the results show the prediction of unknown material. This makes the prediction more challenging but represents a realistic use case for production. The comparative overview of the relevant feature set shows a difference between the prediction of short-circuit current and open-circuit voltage prediction.

## INTRODUCTION

The quality of as-cut multicrystalline silicon wafers in solar cell production has significant influence on the efficiency of the finished solar cell. Also the costs for solar cell process takes up a major part in total cost of manufacturing [1]. Allowing low quality wafers in PV production will lead to lower efficient modules and wastage of expenditure. Therefore, it is necessary to have a quality control of parameters in wafer production and cell fabrication stage thereby reducing the overall cost. In recent years, passivated emitter and rear cell (PERC) has emerged as a promising technology with higher efficiency. PERC based solar cell technology has been successfully transferred into the industry and it is expected to grow in future [2]. Our approach to rate multicrystalline silicon wafers can be applied to any cell technology and is evaluated for PERC technology in this work. In semiconductors like silicon, the lifetime of minority charge carriers is of particular interest in solar cells due to its effect on potential cell efficiency. And the effective lifetime is indirectly proportional to the different recombination effects due to the crystal defects in lattice. These crystallization-related defect structures are visible in PL images [3] of as-cut wafers. These defect structures are extracted as features using image processing methods and quantified to rate the solar cell parameters which is shown in [4]. Due to its faster data acquisition rates and ability to preserve the spatial information in the as-cut wafers, PL-imaging can be used for inline wafer characterization [5]. Earlier studies were successful in predicting the solar cell quality in terms of its open-circuit voltage (e.g. [6]). In order to transfer the existing methodology onto inline industrial environment, it is necessary to optimize algorithms. Moreover, to deduce improvement strategies from the prediction results, it is important for the crystal grower to understand which specific features influence the quality. We follow the approach in [6] and extend it to the prediction of $J_{sc}$ by means of PL images and GB data as input. In addition to the prediction results, most important features for the prediction of $J_{sc}$ and $V_{oc}$ are compared. This provides a deeper understanding of defect features and cell processes.

# APPROACH

## Feature Extraction

Usually dense dark-line structures in PL images relate to dislocation clusters. Grain boundaries and isolated dislocations appear as thin dark-lined structures. Regions with reduced PL signal appear in the edges of the wafers and are usually formed when impurities diffuse into the wafer from the crucible walls. Based on this, we separate the homogeneous and in-homogeneously illuminated wafer regions into regular and contaminated regions. In FIGURE 1(b) the blue coloured contour is the contaminated region and the rest of the wafer is considered as regular region. We quantify crystallization defects and intensity values and distinguish different degrees of defect density and intensity, respectively. To remove the unwanted noise in the sample images, they are pre-processed using median and bilateral filters. As the PL images vary strongly in contrast, it is not suitable to apply intensity-based algorithms. Instead contrast-invariant methods are required. We used such an algorithm as presented in [7] to detect structures from PL images. In case of thin- line-like structures, it is difficult to distinguish them into recombination-active grain boundaries and dislocations only from PL images. A similar approach as mentioned in [8] is followed to detect the grain boundaries from Infrared images (IR) of the wafer samples. In a subsequent step, these detected grain boundaries are overlayed on the PL image of the same wafer sample to separate dislocations from recombination-active grain boundaries in PL images. Initially, all the PL structures that are detected as marked in magenta color as shown in the FIGURE 1(b). In the next step, grain boundaries detected from IR image are marked in orange color and aligned according to [9, pp. 58-59] on the same PL image. Superposed structures are then isolated into recombination-active grain boundaries and the remaining structures into dislocations.

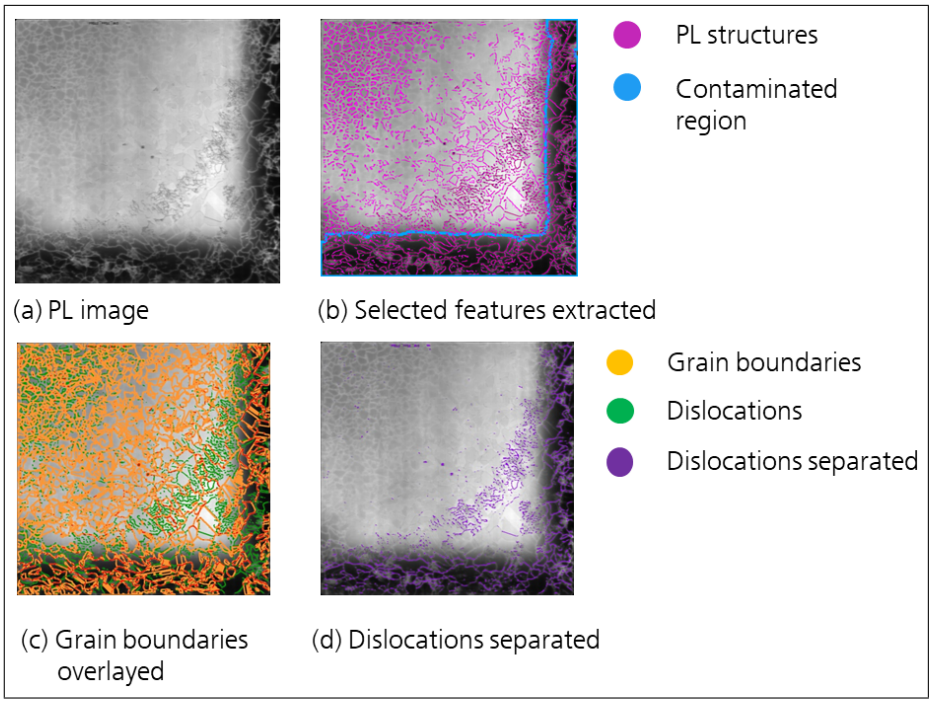


**FIGURE 1.** (a) PL image of the wafer from corner brick. (b) Detected PL structures are represented in magenta color and contaminated region in blue color. (c) Grain boundaries extracted from IR image and overlayed on the image. (d) Recombination-active grain boundaries and dislocation clusters separated.

As the wafer samples examined are from different manufacturers, features extracted from them are carefully chosen such that they are universal and independent of manufacturing process or material type.

In total 24 features are extracted from the PL image of each wafer sample similar to the feature extraction mentioned in [9]. These features can be broadly categorized into four different classes according to their characteristics. (a) features based on recombination active structures (RAS) on PL image, which include area fraction of RAS from

regular and contaminated regions, and histogram of RAS densities with four bins. (b) features based on intensity such as PL image intensity are binned into histogram with five bins and mean PL intensity of the wafer. (c) features related to grain boundaries such as thinned dislocations and (d) features based on wafer position in the brick such as top or bottom part of the brick and area fraction of contaminated region. Furthermore, the algorithm is optimized to achieve better performance speed and in-line capability.

In addition to those features extracted from the PL images, doping concentration of the wafers and six new features from GB data are extracted. These features include area fraction of GB's from IR image, fraction of recombination active GB's (extracted from PL image) over total GB's (extracted from IR image), recombination active GB's over recombination active defects and the area fraction of recombination active GB's. Further, percentage of GB and dislocations after separation are included.

## Learning Model and Feature Evaluation

We want to predict a solar cell quality parameter $y_i$ based on the extracted wafer features $\vec{x_i}$ of a sample $i$. To train the features and to develop prediction models, we used elastic-net algorithm as in [6]. A multi-linear model such as elastic-net algorithm is a better choice in this case as the feature dependency is more intuitive when compared to any non-linear regression models. Therefore, we determine model parameters $\vec{\beta} = (\beta_1, \cdots, \beta_P)$ where P is the number of features. Error between the predicted value $\hat{y}$ and $y$ for $n$ training samples is optimized according to

$$\vec{\beta} = argmin_{\vec{\beta}} \left\{ \frac{1}{2n} \sum_{i=1}^{n} \left( y_i - \beta_0 - \sum_{j=1}^{P} x_{i,j}\beta_j \right)^2 + \lambda \sum_{j=1}^{P} \left|\beta_j\right| \right\}. \quad (1)$$

The second term in equation 1 indicates a penalty term for models with large coefficients. With increasing regularization value $\lambda$ the model complexity reduces. By analyzing the $\ell_1$ norm of the coefficients the total number of active features stays low. An optimum value of $\lambda$ is determined within a model evaluation step for feature selection.

## EXPERIMENTAL

The data set of about 7300 wafers from 10 manufacturers are investigated within an incoming control and processed to PERC solar cells. For our prediction model, the data set is divided into a training set with about 2900 wafers from 32 bricks and 7 boxes, and a test data set with about 4400 wafers from 42 distinct bricks and 18 boxes. A separate validation set is not required as the training is done with cross validation. Four prediction models are trained to predict $\eta$, $V_{oc}$, $J_{sc}$ and FF, one each and their results are shown in FIGURE 2.

From the previously computed model trained on various bricks from all the manufacturers, error's are evaluated according to manufacturer and material type. Consistency of the estimators in prediction models can be explained from error distribution of $V_{oc}$ and $J_{sc}$, shown in FIGURE 3. Here M-1,M-2,M-3...M-10 are the ten different manufacturers and N is the no of wafer samples. And material from manufacturers M1-M9 is classical mc-silicon, whereas M-10 is HPM-silicon material. Further HPM-silicon wafers from manufacturer's M-3 and M-9 are termed as M-3b and M-9b. Despite of wide range of wafer samples used for training, prediction models converge well for the data from different manufacturers.

## Results on Solar Cell Quality Prediction

The prediction results are displayed in TABLE 1. The best correlation of 0.90 is obtained for $V_{oc}$, due to its strong dependence on material parameters. For $J_{sc}$, the correlation coefficient decreases to 0.88 due to the increasing impact of process-related effects, such as texturization. For FF, the correlation coefficient is the lowest with 0.48 as this parameter mainly depends on process-related effects and thus cannot really be predicted in the as-cut stage. As efficiency combines all three parameters, the correlation coefficient of the $\eta$ prediction is with 0.83 lower than that of $V_{oc}$ but still close to it, which shows that an $\eta$ prediction is possible with a sophisticated prediction model as presented here.

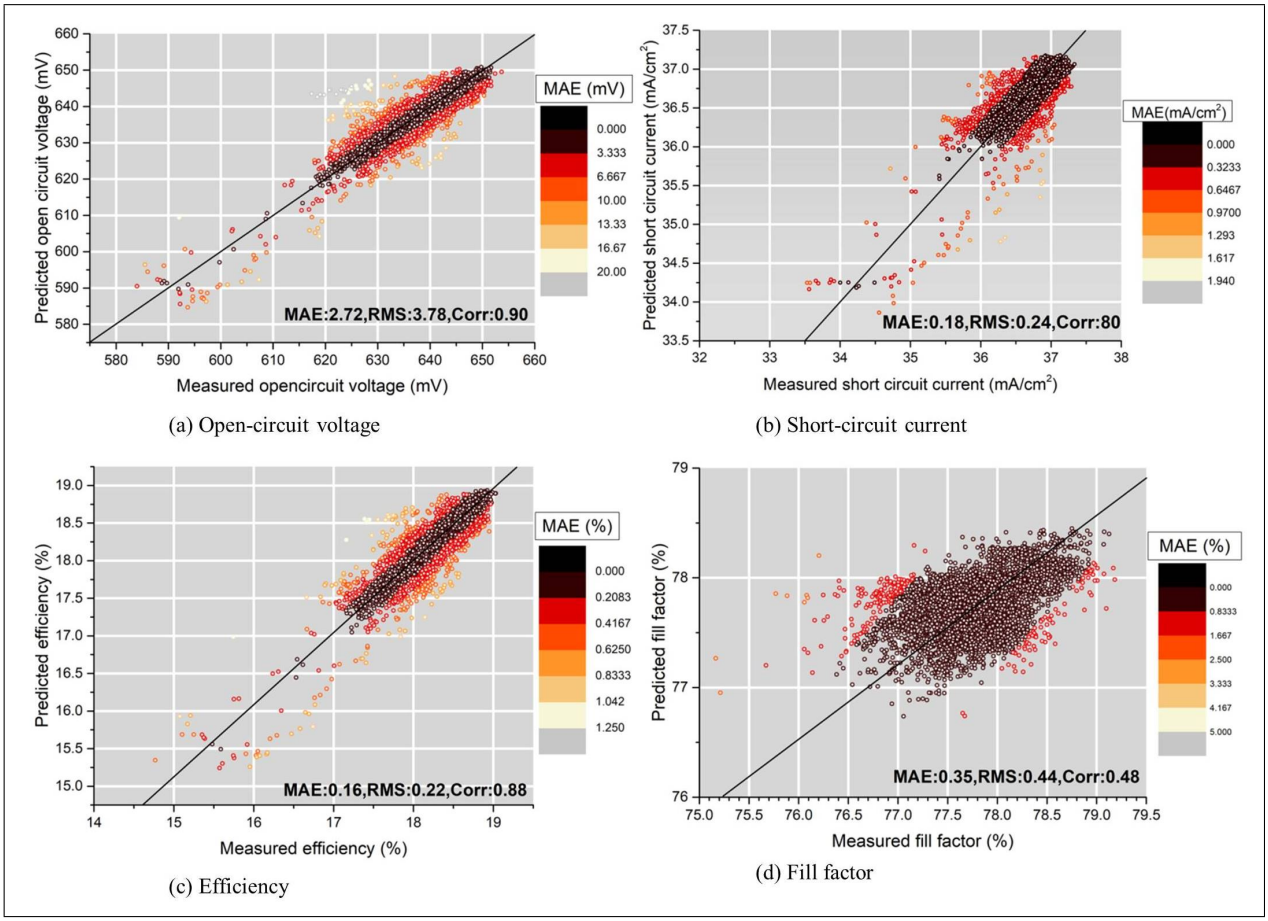


**FIGURE 2.** Measured and predicted solar cell parameters.
* Mean absolute error (MAE), root mean squared error (RMS) and correlation (Corr).

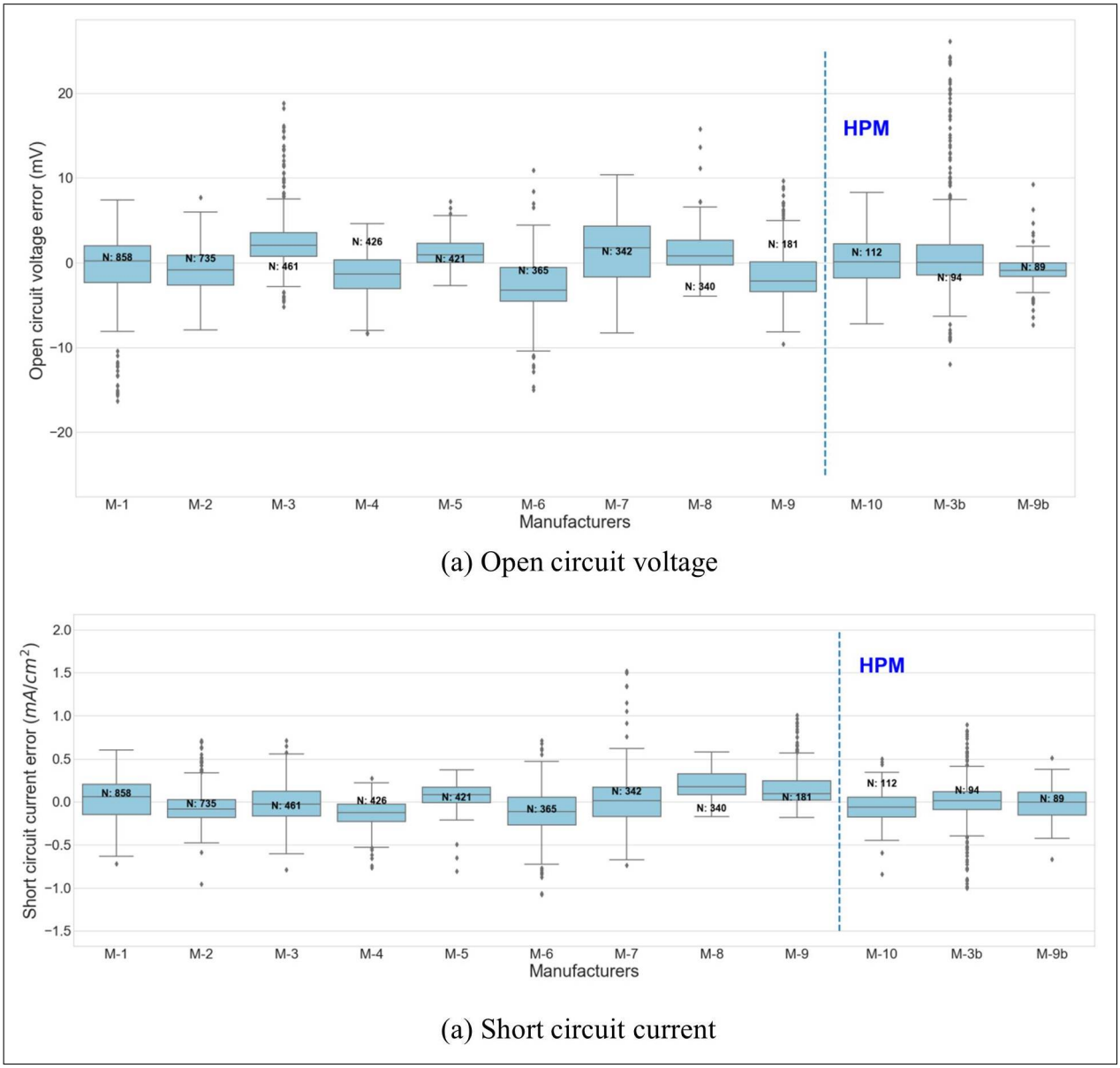


**FIGURE 3.** Error distribution of Voc and Jsc for 10 manufacturers.

**TABLE 1.** Prediction quality

| | $\eta_{\text{abs}}$ [%] | | | $V_{oc}$ [mV] | | | $J_{sc}$ [$mA/cm^2$] | | | $FF_{\text{abs}}$ [%] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mae | rms | corr | mae | rms | corr | mae | rms | corr | mae | rms | corr |
| Without GB data | 0.19 | 0.27 | 0.83 | 2.94 | 4.17 | 0.87 | 0.19 | 0.26 | 0.76 | 0.34 | 0.46 | 0.52 |
| With GB data | 0.16 | 0.22 | 0.88 | 2.72 | 3.78 | 0.90 | 0.18 | 0.24 | 0.80 | 0.34 | 0.44 | 0.52 |

* $\eta_{\text{abs}}$ and $FF_{\text{abs}}$ are the absolute values.

Additional GB-information is also evaluated which improved the classification of features as well by reducing the error values in the prediction models. The second row in the TABLE 1 displays the prediction quality of the models which include features extracted from GB data. $V_{\text{oc}}$ distribution of HPM and mc-Si materials are shown color-coded as a function of the area fraction of dislocations and grain boundaries in FIGURE 4. High $V_{\text{oc}}$ is obtained among the wafers with less defect structures. As the dislocations increase, the $V_{\text{oc}}$ is decreased in both material types. Crystal defects are more dominant in mc-silicon when compared to HPM-silicon material. Hence, $V_{\text{oc}}$ distribution of mc-silicon wafers along the dislocations axis is more widely spread. In contrast, there isn't a considerable change in the $V_{\text{oc}}$ values as the low density defect structures, i.e grain boundaries change. This behaviour is also observed from the FIGURE 5, where grain boundaries do not have any impact in predicting $V_{\text{oc}}$.

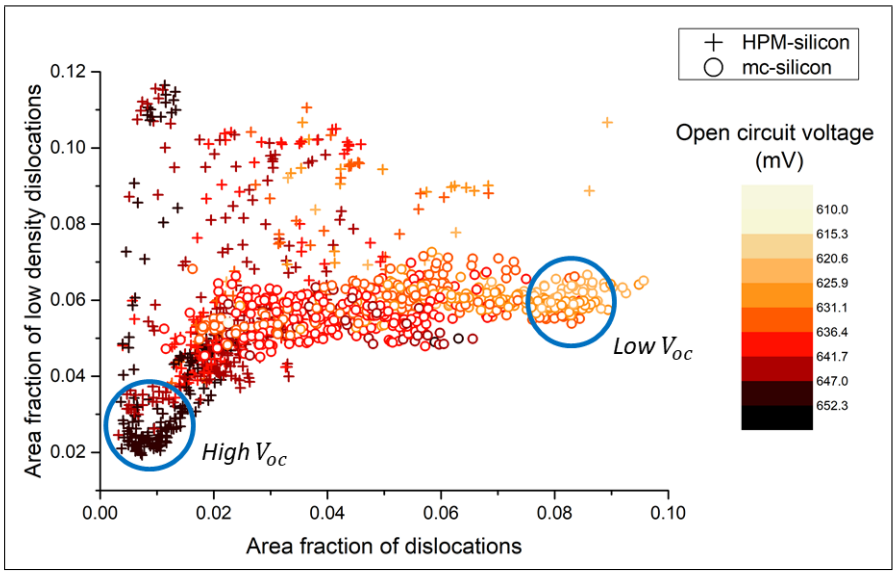


**FIGURE 4.** Distribution of open circuit voltage for HPM (crosses) and mc-silicon (circles) materials as a function of dislocations and low density dislocations which represent grain boundaries.

## Relevance of Features

Due to the regularization term, the presented prediction model activates the most relevant features only. Figure 5 gives an overview of active features for the prediction of $J_{\text{sc}}$, $V_{\text{oc}}$ and and $\eta$ with high feature weights of the model. It can be observed that PL-intensity values are vital for the prediction of all four $IV$-parameters, whereas the area fraction of contaminated regions does not contributes to the prediction models. This might be due to the region of low PL-intensity is a better feature to quantify these regions. Regions of low PL-intensity decrease $J_{\text{sc}}$ values due to high recombination effect in these regions. Regions of dense dislocations decrease the expected $V_{\text{oc}}$ value. Interestingly, the presence of grain boundaries which are reflected in feature vector by low density dislocations can increase $J_{\text{sc}}$. A possible reason may be that crystal defects positively affect the texturing process and thus $J_{\text{sc}}$ as discussed in [10]. In addition to these features, the impact of the model is also distributed among other features whose weights are comparatively less to discussed feature set. Among them, area fraction of recombination active GB over GB (extracted from NVCD) has negative impact on all the $IV$-parameters.

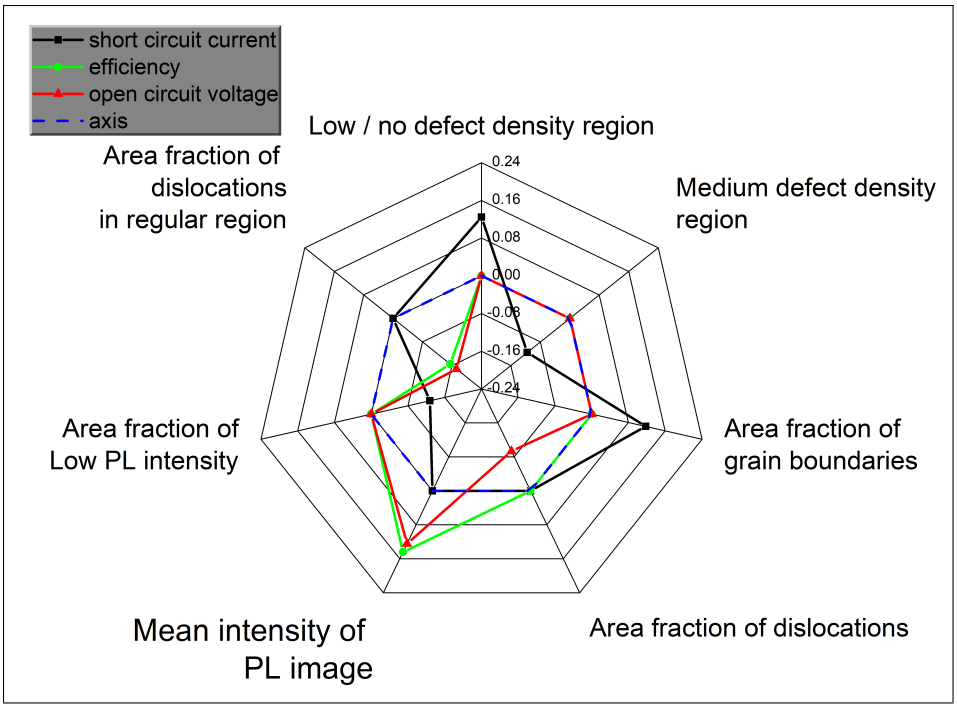


**FIGURE 5.** Contribution of features to the prediction models for weights greater 10% of a model.

## DISCUSSION

In PL images, intensity of the image is correlated to minority charge carrier lifetime of the material. And lifetime of the material is a vital parameter in defining the quality of the wafer and finished solar cells. This behaviour can also be observed from the Figure 5 as low PL intensity regions of the wafer which usually have less carrier lifetime have negative impact on $J_{sc}$. And the mean intensity or global intensity of the PL image has positive impact on $V_{oc}$ and $\eta$. Also crystal defects on the wafers decrease the quality of the solar cells due to high recombination effect. This explains the negative impact of dislocation features on $V_{oc}$ and $\eta$ prediction models as well. In case of fill factor is correlation is not high, as expected due to the influence from solar cell process. Feature relevance from the trained models also visualize the physical relevance of the material characteristics and their impact on the quality of the finished solar cells.

Along with the material quality of the wafer sample, solar cell process also has its influence in final solar cell quality. During the PERC process, some of the features which hinder the quality of the wafer tend to change and thus effect the prediction model. One such case is passivation of grain boundaries in HPM material. Almost 90% of the random angle grain boundaries remain inactive in certain cases [11].

## CONCLUSION AND OUTLOOK

We train regression models to predict $\eta$, $V_{oc}$, $J_{sc}$ and FF based on robust defect features in PL-images. Low prediction errors can be achieved not only for $V_{oc}$ but also for $\eta$ although this parameter is strongly influenced by process related effects. Intensity and dislocation information from PL images are crucial for these models. Adding new features which are computed from the GB data has improved the prediction accuracy of the models and allowed better interpretation of relevant features. The feature models for $V_{oc}$ and $J_{sc}$ prediction show opposite activation for regions with less dense dislocation clusters. The proposed method can be used to optimize crystallization and solar cell process, as the features are physically interpretable.

## REFERENCES


1. I. VDMA, ITRPV 1 – 62 (2017).
2. T. Luka, C. Hagendorf, and M. Turek, Photovoltaics International 1 – 6 (2016).
3. T. Trupke, R. Bardos, M. Schubert, and W. Warta, Applied Physics Letters **89**, p. 044107 (2006).
4. J. Haunschild, M. Glatthaar, M. Demant, J. Nievendick, M. Motzko, S. Rein, and E. R. Weber, Solar Energy Materials and Solar Cells **94**, 2007 – 2012 (2010).
5. T. Trupke, R. A. Bardos, M. D. Abbott, F. w. Chen, J. E. Cotter, and A. Lorenz, "Fast photoluminescence imaging of silicon wafers," in *2006 IEEE 4th World Conference on Photovoltaic Energy Conference*, Vol. 1 (2006), pp. 928–931.
6. M. Demant, S. Rein, J. Haunschild, T. Strauch, H. Höffler, J. Broisch, S. Wasmer, K. Sunder, O. Anspach, and T. Brox, Progress in Photovoltaics: Research and Applications **24**, 1533–46 (2015).
7. M. Felsberg and G. Sommer, Trans. Sig. Proc. **49**, 3136–3144December (2001).
8. T. Strauch, M. Demant, P. Krenckel, S. Riepe, and S. Rein, 7th International Conference on Silicon Photovoltaics 1 – 7 (2017).
9. M. Demant, *Quality Rating of Silicon Wafers-a Pattern Recognition Approach* (Fraunhofer Verlag, 2016).
10. J. Nievendick, M. Demant, J. Haunschild, A. Krieg, F. Souren, S. Rein, M. Zimmer, and J. Rentsch, "Appearance of rift structure created by acidic texturization and their impact on solar cell efficiency," in *Proceedings of the 35th IEEE PVSC* (2010), pp. 447–452.
11. K. Adamczyk, R. Snden, G. Stokkan, E. Looney, M. Jensen, B. Lai, M. Rinio, and M. D. Sabatino, Journal of Applied Physics **123**, p. 055705 (2018), https://doi.org/10.1063/1.5018797 .