# Inversion of the Shockley-Read-Hall Equation Using Random Forests

Yoann Buratti[1], Josef Dick[1], Quoc Le Gia[1], Yan Zhu[1] and Ziv Hameiri[1]

[1] *University of New South Wales, Sydney, NSW, 2052, Australia*
*E-mail: yoann.buratti@student.unsw.edu.au*

Bulk defects in silicon (Si) wafers are key contributors to solar cell efficiency loss.[1] The identification and characterization of these defects are critical steps in the process of improving the efficiency and reliability of solar cells. In this study, we present an application of machine learning (ML) for extraction of defect parameters from temperature- and injection-dependent lifetime spectroscopy (TIDLS). A common technique to fit TIDLS measurements data is the defect parameter solution surface (DPSS) method developed by Rein.[2] Based on the Shockley-Read-Hall (SRH) equation,[3,4] the DPSS method finds the best fit among the combinations of defect parameters: the defect energy level ($E_t$) and the electron ($\sigma_n$) and hole ($\sigma_p$) capture cross-sections. Usually, the DPSS approach results in two possible defect parameter combinations, with one in the upper half of the bandgap and one in the lower half of the bandgap.

In this study, using the SRH equation, approximately half a million TIDLS curves were simulated representing a wide range of defect parameters combinations. Random forest (RF) regressors[5] are trained on 90% of the dataset ("training dataset") and evaluated on the remaining 10% of the dataset ("testing dataset"), using the coefficient of determination $R^2$ on the testing dataset as the scoring method. Since two defect parameters combinations are often obtained with the DPSS approach, the ML models will be trained independently in each bandgap half for each defect parameter, resulting in a total of six trained RF models.

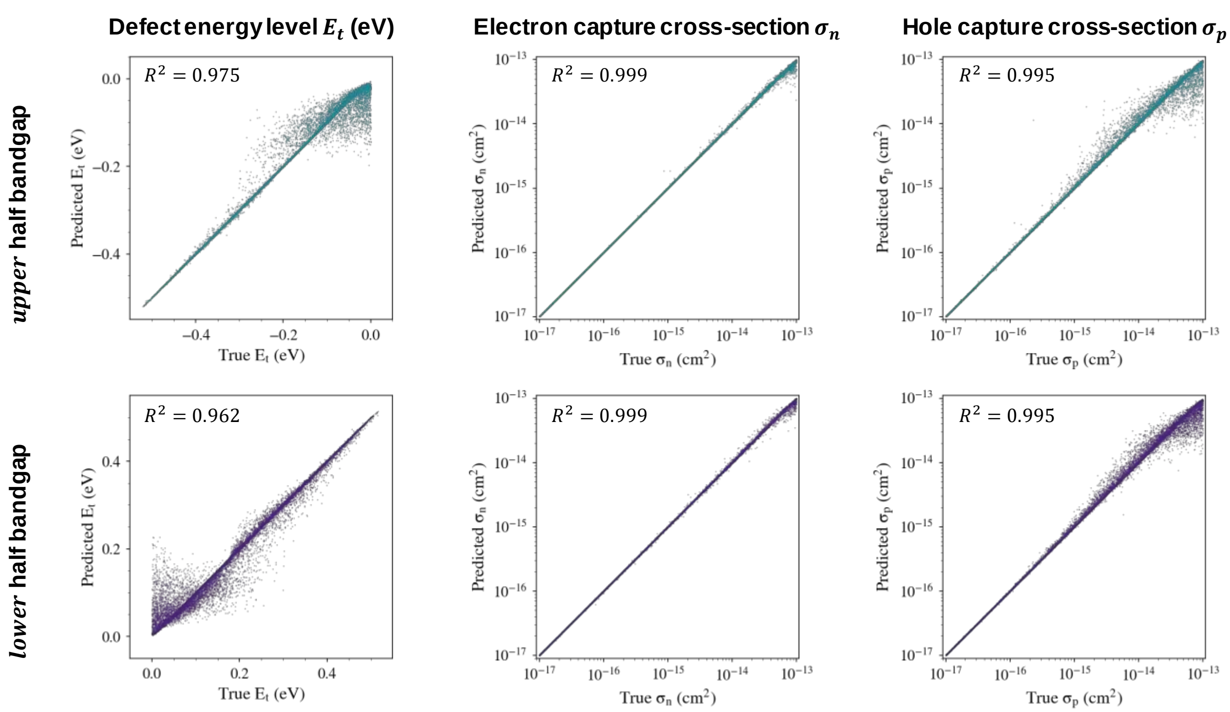


**Figure 1. Random forest predicted defect parameters against their true values**

The results of the six trainings are presented in Figure 1, where the predicted value from the RF models are plotted against their true value. The data shown is from the testing dataset and have previously not be seen by the RF models. The six graphs represent the combination of the three defect parameters ($E_t$, $\sigma_n$, $\sigma_p$) and the two bandgap halves (*upper*, *lower*). All testing $R^2$ score are above 96%, proving the successful prediction of defect parameters with a ML approach.

Limitations of the models, such as less accurate prediction towards the mid-gap ($E_t$ ~ 0 eV) are also observed in the ML results. Indeed, from the SRH equation, lifetime curves of defect with $E_t$ near the mid-gap have a weaker dependence on $E_t$. Therefore, TIDLS data of defects with different $E_t$ appear almost identical in this energy range, making it harder for ML to accurately separate them. However, it should be noted that this limitation is also true for conventional DPSS method. Another example of learned limitation is the overall better prediction of $\sigma_n$ over its counterpart $\sigma_p$. This observation can be explained by the fact that the simulation is done on a p-type wafer and, in certain conditions on the capture cross-section ratio and defect energy level, the TIDLS curves become less sensitive to $\sigma_p$ and therefore are harder to separate. It has been checked that, for a n-type wafer simulation, the reverse is true and that $\sigma_p$ is overall better predicted than $\sigma_n$.

This work pioneers the use of ML for lifetime spectroscopy, bringing the newest prowess of ML to material quality inspection. It opens a new era in the area of defect characterization, as it has the potential to overcome the limitations of current methods, such as dealing with the temperature dependence of various electrical parameters and provide new insights in the theory of lifetime spectroscopy.

**Acknowledgments**

This work was supported by the Australian Government through Australian Renewable Energy Agency [ARENA; project 2017/RND001]. 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. Schmidt, J. *et al.* Impurity-related limitations of next-generation industrial silicon solar cells. in *2012 IEEE 38th Photovoltaic Specialists Conference (PVSC) PART 2* 1–5 (2012). doi:10.1109/PVSC-Vol2.2012.6656779
2. Rein, S. *Lifetime Spectroscopy: A Method of Defect Characterization in Silicon for Photovoltaic Applications*. (Springer Science & Business Media, 2005).
3. Shockley, W. & Read, W. T. Statistics of the Recombinations of Holes and Electrons. *Physical Review* **87**, 835–842 (1952).
4. Hall, R. N. Electron-Hole Recombination in Germanium. *Physical Review* **87**, 387–387 (1952).
5. Breiman, L. Random Forests. *Machine Learning* **45**, 5–32 (2001).