

## Highlights

### **Extracting the iron concentration in silicon solar cells using photovoltaic parameters and machine learning**

Oleg Olikh, Oleksii Zavhorodnii

- The iron defect transformation effect on Si solar cells' performance was studied using SCAPS simulation
- Short-circuit current changes are most suitable for estimating iron impurity concentration.
- Open-circuit voltage changes are a non-monotonic function of iron concentration at low doping levels.
- Monochromatic illumination is more effective than AM1.5 for accurate iron concentration estimation.

# Extracting the iron concentration in silicon solar cells using photovoltaic parameters and machine learning

Oleg Olikh\*, Oleksii Zavhorodnii

Taras Shevchenko National University of Kyiv, 64/13, Volodymyrska Street, Kyiv, 01601, Ukraine

## ARTICLE INFO

### Keywords:

silicon  
iron-boron pairs  
solar cells  
SCAPS simulation  
defect influence  
estimation of iron contamination


## ABSTRACT

Transitioning to renewable energy sources is paramount for humanity's sustainable development, and silicon solar cells are at the forefront of solar energy conversion. Iron in these structures is a primary one of the most detrimental metallic impurities. This study examines the impact of iron defect variability on silicon solar cell performance across various scenarios. We have simulated solar cells using SCAPS software across a range of temperatures (290 – 340) K, base thicknesses (180 – 380)  $\mu\text{m}$ , doping levels ( $10^{15}$  –  $10^{17}$ )  $\text{cm}^{-3}$ , with iron concentrations varying from  $10^{10}$  to  $10^{14}$   $\text{cm}^{-3}$  under AM1.5 and monochromatic (940 nm) illumination. Analyzed across all cases were the effects of iron-boron pair dissociation on short-circuit current, open-circuit voltage, fill factor, and efficiency. The experimental measurements validated the simulation results, demonstrating good agreement for all photovoltaic parameters. This study investigates the potential of using photovoltaic parameter changes induced by iron-related defect restructuring to estimate iron concentration. It is shown that changes in short-circuit current obtained under monochromatic illumination are the most reliable, while the fill factor is the least effective. The study examined the correlation between changes in photovoltaic parameters caused by pair dissociation while establishing the expedience of applying principal component analysis in impurity concentration evaluation with the help of multiple parameters.

## 1. Introduction

The urgent challenges of global warming and fossil fuel depletion have spurred the search for renewable energy sources, accelerating the rapid advancement of photovoltaic (PV) energy and the widespread deployment of solar panels [1, 2]. However, PV energy faces several significant challenges: the need to optimize solar cell manufacturing processes, improve the efficiency and stability of PV devices (either by discovering new materials or by enhancing existing structures), develop energy management systems designed to regulate the energy flow in real-time and establish methods for monitoring the condition of photoelectric converters, specifically through defect characterization. Simultaneously, the application of machine learning (ML) methods, which can identify patterns and correlations that are not readily apparent through traditional analysis [3], opens new avenues for addressing the issues above in photovoltaics. For instance, ML methods optimize individual production stages — such as crystal growth [4] and plasma-enhanced vapor deposition [5] — and streamline entire solar cell production lines, which may involve up to ten processing steps and forty-seven process parameter inputs [6]. Moreover, ML techniques enhance PV devices by identifying key photovoltaic materials, analyzing references to various structural configurations in the literature [7] or optimizing perovskite solar cells [8], designing renewable energy-based demand-side management systems [9], and extracting additional information from internal quantum efficiency measurements [10]. Numerous reviews provide further insights into ML applications in PV systems [11, 12, 13, 14, 15].

We will now explore the application of artificial intelligence in greater depth for defect analysis. Most relevant studies focus on image analysis of solar cells, including electroluminescence, photoluminescence, and infrared thermography [16, 17, 18, 19, 20, 21, 22, 23]. These methods facilitate detecting and classifying defects such as cracks, finger failures, hot spots, scratches, and horizontal dislocations and are predominantly implemented using convolutional neural networks. Another widely adopted approach involves applying ML models to current-voltage ( $I$ - $V$ ) curves, enabling the identification of permanent and temporal faults in PV arrays [24, 25, 26, 27]. A key advantage of this defect characterization method is its reliance on  $I$ - $V$  measurements, a standard procedure for PV device assessment. Additionally, ML techniques specifically designed for analyzing point defects warrant particular attention. For instance,

 olegolikh@knu.ua (O. Olikh); nevermor464@gmail.com (O. Zavhorodnii)  
ORCID(s): 0000-0003-0633-5429 (O. Olikh); 0000-0001-8080-7661 (O. Zavhorodnii)

researchers have developed methods for determining the electronic structure of intrinsic defects [28, 29], detecting radiation-induced defects via Raman spectroscopy [30, 3], and extracting recombination-active center parameters from temperature- and injection-dependent lifetime measurements [31, 32, 33]. Although such defects represent one of the main limitations of PV devices [34, 35], studies in this area remain scarce. Moreover, existing research primarily focuses on characterizing defects in PV materials rather than solar cells and relies on data obtained through specialized equipment.

This study proposes an ML-based approach to extract impurity concentrations from  $I$ - $V$  curves. Specifically, we focus on quantifying iron in boron-doped crystalline silicon solar cells. This constraint is not overly restrictive, given that (i) Si-based solar cells dominate the current PV market [36, 37], with most being manufactured from boron-doped crystalline silicon (c-Si:B); and (ii) iron is one of the most prevalent, ubiquitous, and efficiency-limiting metallic impurities [38, 39]. It is well established that in  $p$ -type material, iron tends to bind with dopant atoms such as boron, forming iron–boron pairs under equilibrium conditions or existing as interstitial species only in the presence of sufficiently high free electron densities [40, 41]. The deliberate transition between these states can be readily induced through intense illumination, electron injection, or heating up to 200 °C and is commonly employed in various methods for assessing iron concentration [42, 43, 44, 45, 46, 47, 48, 49]. The approach proposed in this study leverages changes in photovoltaic parameters (PVPs) (short-circuit current, open-circuit voltage, efficiency, and fill factor) resulting from FeB dissociation as input features for ML algorithms. The specified PVPs can be easily extracted from  $I$ - $V$  characteristics, making this method advantageous compared to existing approaches:

(1) Unlike glow discharge mass spectrometry or secondary ion mass spectrometry [50], it is non-destructive. (2) It does not require specialized equipment or specially prepared samples, in contrast to Fourier-transform infrared spectroscopy, electron paramagnetic resonance, deep-level transient spectroscopy (DLTS), Laplace DLTS, carrier lifetime measurements, photoluminescence, or photoconductance [51, 52, 53, 43, 44, 47, 45]. (3) It is relatively simple and fast compared to other methods that also rely on measuring PV parameters, such as the kinetics of short-circuit current [46] and open-circuit voltage [48], but require lengthy experimental procedures or multiple illumination levels.

In our previous work [54], we employed a deep learning approach to estimate iron concentration based on the ideality factor, which was also derived from  $I$ - $V$  characteristics. However, the method proposed in this study imposes fewer constraints on the accuracy of  $I$ - $V$  measurements across the entire voltage range and the model used to describe the  $I$ - $V$  characteristics of actual structures. Thus, our approach allows for the simultaneous determination of iron impurity concentration along with key electrical parameters. These inline characterization techniques are crucial for ensuring efficient production lines and optimizing processes to produce reliable solar cells. Notably, luminescence imaging is increasingly used for solar cell characterization. However, ML methods have been proposed to extract  $I$ - $V$  characteristic data from such images [55, 56]. Integrating these methods with our approach into a unified pipeline would enable the extraction of iron concentration data from luminescence measurements.

## 2. Methodology

### 2.1. Data Collection

### 2.2. Data Pre-Processing

### 2.3. Machine Learning Algorithms

### 2.4. Model evaluation

[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 50, 55, 56, 57, 58, 59, 60, 61, 62]

## 3. Research Methodology

### 3.1. Simulation Details

## 4. Results and Discussion

## 5. Conclusion

## Acknowledgments

O.O. would like to acknowledge the financial support by National Research Foundation of Ukraine (Project No. 2023.03/0252 “Development of principles for the creation and machine-oriented characterization of porous silicon nanostructures with optimal heat transport properties”)

## Supplementary data

Supplementary data to this article can be found online at <http://surl.li/qneich>

## Data availability

Data will be made available on request.

## References

- [1] S. Karlilar Pata, M. Balcilar, Decarbonizing energy: Evaluating fossil fuel displacement by renewables in oecd countries, *Environ. Sci. Pollut. Res.* 31 (2024) 31304–31313.
- [2] J. L. Holechek, H. M. E. Geli, M. N. Sawalhah, R. Valdez, A global assessment: Can renewable energy replace fossil fuels by 2050?, *Sustainability* 14 (2022) 4792.
- [3] S. Park, J. Lee, S. Khan, A. Wahab, M. Kim, Machine learning-based heavy metal ion detection using surface-enhanced raman spectroscopy, *Sensors* 22 (2022) 596.
- [4] X. Qi, W. Ma, Y. Dang, W. Su, L. Liu, Optimization of the melt/crystal interface shape and oxygen concentration during the czochralski silicon crystal growth process using an artificial neural network and a genetic algorithm, *J. Cryst. Growth* 548 (2020) 125828.
- [5] L. Rachdi, M. Hofmann, Use of optical emission spectroscopy to predict silicon nitride layer properties, *Vacuum* 191 (2021) 110322.
- [6] Y. Buratti, C. Eijkens, Z. Hameiri, Optimization of solar cell production lines using neural networks and genetic algorithms, *ACS Appl. Energy Mater.* 3 (2020) 10317–10322.
- [7] L. Zhang, M. He, Unsupervised machine learning for solar cell materials from the literature, *J. Appl. Phys.* 131 (2022) 064902.
- [8] T. Liu, S. Wang, Y. Shi, L. Wu, R. Zhu, Y. Wang, J. Zhou, W. C. H. Choy, Machine-learning accelerating the development of perovskite photovoltaics, *Sol. RRL* 7 (2023) 2300650.
- [9] Z. Asghar, K. Hafeez, D. Sabir, B. Ijaz, S. S. H. Bukhari, J. Ro, Reclaim: Renewable energy based demand-side management using machine learning models, *IEEE Access* 11 (2023) 3846–3857.
- [10] Z. Abdullah-Vetter, B. Wright, T.-C. Wu, A. Shakiba, Z. Hameiri, Automatic quantitative analysis of internal quantum efficiency measurements of gaas solar cells using deep learning, *Adv. Sci.* 12 (2025) 2407048.
- [11] M. Di Sabatino, R. Hendawi, A. S. Garcia, Silicon solar cells: Trends, manufacturing challenges, and ai perspectives, *Crystals* 14 (2024) 167.
- [12] S. Datta, A. Baul, G. C. Sarker, P. K. Sadhu, D. R. Hodges, A comprehensive review of the application of machine learning in fabrication and implementation of photovoltaic systems, *IEEE Access* 11 (2023) 77750–77778.
- [13] R. Jaiswal, M. Martínez-Ramón, T. Busani, Recent advances in silicon solar cell research using data science-based learning, *IEEE J. Photovolt.* 13 (2023) 2–15.
- [14] Y. Buratti, G. M. Javier, Z. Abdullah-Vetter, P. Dwivedi, Z. Hameiri, Machine learning for advanced characterisation of silicon photovoltaics: A comprehensive review of techniques and applications, *Renewable Sustainable Energy Rev.* 202 (2024) 114617.
- [15] S. Bhatti, H. U. Manzoor, B. Michel, R. S. Bonilla, R. Abrams, A. Zoha, S. Hussain, R. Ghannam, Revolutionizing low-cost solar cells with machine learning: A systematic review of optimization techniques, *Advanced Energy and Sustainability Research* 4 (2023) 2300004.
- [16] H. Munawer Al-Otun, Classification of anomalies in electroluminescence images of solar pv modules using cnn-based deep learning, *Sol. Energy* 278 (2024) 112803.
- [17] L. Pratt, D. Govender, R. Klein, Defect detection and quantification in electroluminescence images of solar pv modules using u-net semantic segmentation, *Renew. Energy* 178 (2021) 1211–1222.
- [18] Z. Li, S. Zhang, C. Qu, Z. Zhang, F. Sun, Research on multi-defects classification detection method for solar cells based on deep learning, *PLOS ONE* 19 (2024) 1–16.
- [19] H.-H. Lin, H. K. Dandage, K.-M. Lin, Y.-T. Lin, Y.-J. Chen, Efficient cell segmentation from electroluminescent images of single-crystalline silicon photovoltaic modules and cell-based defect identification using deep learning with pseudo-colorization, *Sensors* 21 (2021) 4292.
- [20] W. Tang, Q. Yang, K. Xiong, W. Yan, Deep learning based automatic defect identification of photovoltaic module using electroluminescence images, *Sol. Energy* 201 (2020) 453–460.
- [21] C. Bu, T. Liu, R. Li, R. Shen, B. Zhao, Q. Tang, Electrical pulsed infrared thermography and supervised learning for pv cells defects detection, *Sol. Energy. Mat. Sol.* 237 (2022) 111561.
- [22] M. Turek, M. Meusel, Automated classification of electroluminescence images using artificial neural networks in correlation to solar cell performance parameters, *Sol. Energy. Mat. Sol.* 260 (2023) 112483.
- [23] C. Huang, Z. Zhang, L. Wang, Psopruner: Pso-based deep convolutional neural network pruning method for pv module defects classification, *IEEE J. Photovolt.* 12 (2022) 1550–1558.
- [24] Y. Liu, K. Ding, J. Zhang, Y. Li, Z. Yang, W. Zheng, X. Chen, Fault diagnosis approach for photovoltaic array based on the stacked auto-encoder and clustering with i-v curves, *Energy Convers. Manage.* 245 (2021) 114603.
- [25] Z. Chen, Y. Chen, L. Wu, S. Cheng, P. Lin, Deep residual network based fault detection and diagnosis of photovoltaic arrays using current-voltage curves and ambient conditions, *Energy Convers. Manage.* 198 (2019) 111793.
- [26] M. W. Hopwood, T. Gunda, H. Seigneur, J. Walters, Neural network-based classification of string-level iv curves from physically-induced failures of photovoltaic modules, *IEEE Access* 8 (2020) 161480–161487.
- [27] A. Mellit, S. Kalogirou, Artificial intelligence and internet of things to improve efficacy of diagnosis and remote sensing of solar photovoltaic systems: Challenges, recommendations and future directions, *Renewable Sustainable Energy Rev.* 143 (2021) 110889.
- [28] Y. Ma, H. Yu, Y. Zhong, S. Chen, X. Gong, H. Xiang, Transferable machine learning approach for predicting electronic structures of charged defects, *Appl. Phys. Lett.* 126 (2025) 044103.

- [29] K. Choudhary, B. G. Sumpter, Can a deep-learning model make fast predictions of vacancy formation in diverse materials?, *AIP Adv.* 13 (2023) 095109.
- [30] J. Y. Chia, N. Thamrongsiripak, S. Thongphanit, N. Nuntawong, Machine learning-enhanced detection of minor radiation-induced defects in semiconductor materials using raman spectroscopy, *J. Appl. Phys.* 135 (2024) 025701.
- [31] S. Wang, B. Wright, Y. Zhu, Y. Buratti, Z. Hameiri, Extracting the parameters of two-energy-level defects in silicon wafers using machine learning models, *Sol. Energ. Mat. Sol.* 277 (2024) 113123.
- [32] Y. Buratti, J. Dick, Q. Le Gia, Z. Hameiri, Deep learning extraction of the temperature-dependent parameters of bulk defects, *ACS Appl. Mater. Interfaces* 14 (2022) 48647–48657.
- [33] Y. Buratti, Q. T. Le Gia, J. Dick, Y. Zhu, Z. Hameiri, Extracting bulk defect parameters in silicon wafers using machine learning models, *npj Computational Materials* 6 (2020) 142.
- [34] T. T. Le, Z. Zhou, A. Chen, Z. Yang, F. Rougieux, D. Macdonald, A. Liu, Reassessing iron–gallium recombination activity in silicon, *J. Appl. Phys.* 135 (2024) 133107.
- [35] M. Yamaguchi, K.-H. Lee, K. Araki, N. Kojima, Y. Ohshita, Analysis for efficiency potential of crystalline si solar cells, *J. Mater. Res.* 33 (2018) 2621–2626.
- [36] M. Bošnjaković, Advance of sustainable energy materials: Technology trends for silicon-based photovoltaic cells, *Sustainability* 16 (2024) 7962.
- [37] J. Zhang, Solar pv market research and industry competition report, *IOP Conf. Ser.: Earth Environ. Sci.* 632 (2021) 032047.
- [38] T. Buonassisi, A. A. Istratov, M. D. Pickett, M. Heuer, J. P. Kalejs, G. Hahn, M. A. Marcus, B. Lai, Z. Cai, S. M. Heald, Chemical natures and distributions of metal impurities in multicrystalline silicon materials, *Prog. Photovolt.: Res. Appl.* 14 (2006) 513–531.
- [39] M. Schubert, M. Padilla, B. Michl, L. Mundt, J. Giesecke, J. Hohl-Ebinger, J. Benick, W. Warta, M. Tajima, A. Ogura, Iron related solar cell instability: Imaging analysis and impact on cell performance, *Sol. Energy Mater. Sol. Cells* 138 (2015) 96–101.
- [40] L. Kimerling, J. Benton, Electronically controlled reactions of interstitial iron in silicon, *Physica B+C* 116 (1983) 297–300.
- [41] C. Möller, T. Bartel, F. Gibaja, K. Lauer, Iron-boron pairing kinetics in illuminated p-type and in boron/phosphorus co-doped n-type silicon, *J. Appl. Phys.* 116 (2014) 024503.
- [42] G. Zoth, W. Bergholz, A fast, preparation-free method to detect iron in silicon, *J. Appl. Phys.* 67 (1990) 6764–6771.
- [43] S. Rein, S. W. Glunz, Electronic properties of interstitial iron and iron-boron pairs determined by means of advanced lifetime spectroscopy, *J. Appl. Phys.* 98 (2005) 113711.
- [44] J. Schmidt, D. Macdonald, Recombination activity of iron-gallium and iron-indium pairs in silicon, *J. Appl. Phys.* 97 (2005) 113712.
- [45] M. Goodarzi, R. A. Sinton, H. Jin, P. Zheng, W. Chen, Q. Wang, D. Macdonald, Accuracy of interstitial iron measurements on p-type multicrystalline silicon blocks by quasi-steady-state photoconductance, *IEEE J. Photovolt.* 7 (2017) 1216–1223.
- [46] O. Olikh, V. Kostylov, V. Vlasuk, R. Korkishko, Y. Olikh, R. Chupryna, Features of FeB pair light-induced dissociation and repair in silicon n+-p-p+ structures under ultrasound loading, *J. Appl. Phys.* 130 (2021) 235703.
- [47] S. Herlufsen, D. Macdonald, K. Bothe, J. Schmidt, Imaging of the interstitial iron concentration in crystalline silicon by measuring the dissociation rate of iron–boron pairs, *Phys. Status Solidi RRL* 6 (2012) 1–3.
- [48] A. Herguth, Quantification of iron in boron-doped silicon solar cells from open circuit voltage measurements, *IEEE J. Photovolt.* 12 (2022) 937–947.
- [49] D. H. Macdonald, L. J. Geerligs, A. Azzizi, Iron detection in crystalline silicon by carrier lifetime measurements for arbitrary injection and doping, *J. Appl. Phys.* 95 (2004) 1021–1028.
- [50] M. Di Sabatino, Detection limits for glow discharge mass spectrometry (gdms) analyses of impurities in solar cell silicon, *Measurement* 50 (2014) 135–140.
- [51] D. K. Schroder, *Semiconductor Material and Device Characterization*, John Wiley & Sons, New Jersey, third edition, 2006.
- [52] R. C. Kurchin, J. R. Poindexter, V. Vähänissi, H. Savin, C. del Cañizo, T. Buonassisi, How much physics is in a current-voltage curve? inferring defect properties from photovoltaic device measurements, *IEEE J. Photovolt.* 10 (2020) 1532–1537.
- [53] A. R. Peaker, V. P. Markevich, J. Coutinho, Tutorial: Junction spectroscopy techniques and deep-level defects in semiconductors, *J. Appl. Phys.* 123 (2018) 161559.
- [54] O. Olikh, O. Lozitsky, O. Zavhorodnii, Estimation for iron contamination in si solar cell by ideality factor: Deep neural network approach, *Prog. Photovoltaics Res. Appl.* 30 (2022) 648–660.
- [55] P. Kunze, J. M. Greulich, A. Tummaliéh, W. Wirtz, H. Hoeffler, N. Woehrlé, S. Glunz, S. Rein, M. Demant, Contactless inline iv measurement of solar cells using an empirical model, *Sol. RRL* 7 (2023) 2200599.
- [56] M. Battaglia, E. Comi, T. Stadelmann, R. Hiestand, B. Ruhstaller, E. Knapp, Deep ensemble inverse model for image-based estimation of solar cell parameters, *APL Machine Learning* 1 (2023) 036108.
- [57] H. Minagawa, T. Tezuka, H. Tsuchida, Effective combinations of features in predicting the range of incident ions using machine learning, *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.* 553 (2024) 165383.
- [58] L. Breiman, Random forests, *Mach. Learn.* 45 (2001) 5–32.
- [59] A. Natekin, A. Knoll, Gradient boosting machines, a tutorial, *Front. Neurobot.* 7 (2013).
- [60] S. Akinpelu, S. Abolade, E. Okafor, D. Obada, A. Ukpong, S. Kumar R., J. Healy, A. Akande, Interpretable machine learning methods to predict the mechanical properties of abx3 perovskites, *Results Phys.* 65 (2024) 107978.
- [61] W. Cao, X. Liu, J. Ni, Parameter optimization of support vector regression using henry gas solubility optimization algorithm, *IEEE Access* 8 (2020) 88633–88642.
- [62] M. F. Hanif, M. U. Siddique, J. Si, M. S. Naveed, X. Liu, J. Mi, Enhancing solar forecasting accuracy with sequential deep artificial neural network and hybrid random forest and gradient boosting models across varied terrains, *Advanced Theory and Simulations* 7 (2024) 2301289.