

Dear Editor,

Dear Editor and Reviewers, We sincerely thank you for taking the time to review our manuscript "Extracting the iron concentration in silicon solar cells using photovoltaic parameters and machine learning" (Ms. Ref. No.: SEJ-D-25-01089). Your insightful comments and constructive suggestions have greatly helped us improve the quality of our work. We particularly appreciate your careful reading and thoughtful feedback, which have led to significant improvements in both the technical content and presentation clarity of our manuscript. We have carefully addressed all the comments and made corresponding revisions to the manuscript. Below we provide our detailed point-by-point responses to each comment. We hope the revised manuscript better meets your expectations and standards for publication in Solar Energy.

## Response to Reviewer #2

**Comment 1.** *Why temperature range (290-340) K? what was the dissociation efficiency of FeB pairs at this temperature? How much is translated to cell degradation? Since at this temperature range, other factors also go side on that impact device characteristic.*

### Reply:

The selected temperature range (290–340) K was chosen for several reasons:

- **practical relevance:** the selected temperature range reflects the realistic operating conditions of most silicon solar cells, ensuring our results are applicable to real-world scenarios;
- **FeB pair dynamics:** within this temperature range, FeB pairs in boron-doped silicon exhibit significant thermal dissociation kinetics, enabling the study of their reversible transformation into interstitial iron ( $\text{Fe}_i$ ) and substitutional boron ( $\text{B}_s$ ).

The dissociation of FeB pairs is a thermally activated process. At the lower end of this range (290 K), FeB pairs are relatively stable, while at higher temperatures (above 320 K), a significant fraction of pairs dissociate, releasing  $\text{Fe}_i$ . [? ]

In the range of 290–340 K, especially under strong illumination, as in our experiments, the dissociation efficiency approaches 100%, and most FeB pairs decay [? ]. The dissociation of FeB pairs releases  $\text{Fe}_i$ , which is a highly recombination-active center. This leads to a pronounced decrease in minority carrier lifetime and, consequently, measurable degradation of photovoltaic parameters such as short-circuit current ( $I_{SC}$ ), open-circuit voltage ( $V_{OC}$ ), fill factor (FF), and efficiency ( $\eta$ ). For instance, simulations show relative changes (e.g.,  $(\epsilon I_{SC})$  can exceed (5–10)% at higher iron concentrations ( $10^{13} - 10^{14} \text{ cm}^{-3}$ ), directly correlating with iron-induced recombination [DOI: 10.1002/pip.594].

Our methodology does not induce any irreversible physical or chemical changes in the solar cell structure. The temperature range used (290–340) K is well below the threshold for processes that could cause permanent defect formation, dopant diffusion, or degradation. The intense illumination used to dissociate FeB pairs is also within the limits commonly applied in standard characterization procedures and does not lead to light-induced degradation or other irreversible effects under our experimental protocol.

**Comment 2.** *Complete details about the dissociation of FeB pairs using halogen lamp illumination? What was the time duration? How much was the decay efficiency.*

### Reply:

The illumination time was 30 seconds for each dissociation cycle. This duration was chosen based on literature evidence and our own preliminary tests, which confirm that at the applied intensity and temperature, this period is sufficient for near-complete dissociation of FeB pairs.

At the specified temperature and illumination conditions, the dissociation efficiency of FeB pairs exceeds 90%, as supported by both our experimental results and previous studies [? ]. After halogen lamp exposure, the majority of FeB pairs are converted to interstitial iron ( $\text{Fe}_i$ ), which is confirmed by the observed changes in photovoltaic parameters.

**Comment 3.** *Most of the models significantly fail for the AM1.5 illumination? What could be the reason for that?*

**Reply:**

The reason most models significantly fail under AM1.5 illumination is the increased complexity of physical interactions caused by its broad-spectrum nature. Unlike monochromatic illumination, which provides a targeted, direct signal related to iron impurities, AM1.5 introduces a wide range of wavelengths that trigger diverse effects—such as variable carrier generation, multiple recombination pathways, and competing influences from surface and bulk defects. This results in a more intricate and less direct relationship between photovoltaic parameter changes and iron concentration, making accurate predictions challenging. Even with AM1.5-specific training data, the inherent complexity overwhelms many models, leading to higher error rates compared to the simpler, more predictable conditions of monochromatic illumination.

**Comment 4.** *Authors are suggested to provide more explanation behind failure of the particular model, infact, for the results obtained.*

**Reply:**

The notably poor performance of the SVR model in our results is likely due to its limited ability to capture the complex, highly non-linear relationships between the input features (such as relative changes in PV parameters, temperature, doping, and thickness) and the target iron concentration, especially when compared to more flexible models like XGB or DNN. SVR with standard kernels can struggle when the data distribution is multi-dimensional and the dependencies are not well approximated by simple kernel functions. Additionally, SVR is sensitive to feature scaling and hyperparameter selection, which may further limit its effectiveness in this application. This explains why SVR underperforms relative to ensemble and deep learning models in our study.

Additions to Section 3.1 ("Train dataset")

The performance of models, such as SVR, degrades significantly under AM1.5 illumination compared to 940 nm illumination due to differences in the relative changes in short-circuit current and charge generation locations. Under 940 nm illumination, the light penetrates deeply, generating charge predominantly in the base where iron impurities strongly influence  $I_{SC}$ . The dissociation of iron-boron pairs thus produces a large  $eI_{SC}$ , creating a clear signal for models to detect iron concentration. In contrast, AM1.5 illumination, with its broad spectrum, generates charge across the cell—near the surface (emitter) for shorter wavelengths and deeper (base) for longer wavelengths—resulting in a smaller  $eI_{SC}$  due to the diluted impact of iron in the base. This reduced signal-to-noise ratio complicates the accurate prediction of iron concentration, particularly for models like SVR that struggle with weak or noisy relationships.

The distinct charge generation profiles further explain the observed results. Under 940 nm illumination, charge generation in the base aligns closely with iron-related recombination, establishing a direct link between iron concentration and photovoltaic parameter changes that even simpler models (e.g., Random Forest) can effectively capture. However, under AM1.5 illumination, charge generation spans the emitter, junction, and base, introducing additional recombination mechanisms (e.g., surface effects). This increased complexity challenges all models, with SVR performing poorly due to its limited ability to handle such nonlinearity and noise. More advanced models, such as XGB and DNN, outperform SVR by better resolving these subtle patterns, though their accuracy under AM1.5 remains lower than under 940 nm due to the inherent difficulties posed by the smaller  $eI_{SC}$  and distributed charge generation.

**Comment 5.** *What can be done to further increase the accuracy of these models?*

**Reply:**

The accuracy of the models could be further improved by:

- **expanding the training data set:** simulating a broader and denser range of device parameters (e.g., more values for doping, thickness, temperature, and iron concentration) would provide the models with more comprehensive training examples and improve generalization;
- **enhancing the physical realism of simulations:** incorporating even more detailed or advanced physical models in SCAPS, such as additional defect types, surface effects, or more accurate material parameters.

**Comment 6.** *There are some typing mistakes as well. for e.g. Page 14, line 44, t-tered...*

**Reply:**

We thank the Reviewer for noting the typographical error (“t-tered” instead of “t-altered”) and any others present in the manuscript. We carefully reviewed the entire text and corrected this and other typographical mistakes to improve the clarity and professionalism of the manuscript.

## **Response to Reviewer #4**

**Comment 1.** *Clearly describe the main objectives as to why and how using this ML approach can be better than traditional ways in the Introduction section.*

**Reply:**

In response, we have revised the Introduction to clearly state the main objectives and the advantages of the proposed machine learning approach over traditional methods for iron quantification in silicon solar cells.

**Objective:** The main objective of our study is to develop a non-destructive, efficient, and accurate method for quantifying iron concentration in silicon solar cells using machine learning models trained on photovoltaic parameter changes induced by FeB pair dissociation.

Why ML is advantageous:

- **non-destructive and equipment-free:** unlike traditional techniques (e.g., mass spectrometry, DLTS, photoluminescence), our approach does not require specialized equipment or destructive sample preparation;
- **based on standard measurements:** it relies solely on  $I-V$  curve measurements, which are routine and accessible in PV characterization;
- **faster and simpler:** the ML method enables rapid estimation of iron content without lengthy kinetic measurements or multiple illumination steps;
- **captures complex dependencies:** machine learning can model the subtle and non-linear relationships between iron concentration and PV parameter changes, which are difficult to describe analytically or with traditional regression methods;
- **inline process monitoring:** the approach is well-suited for inline quality control and process optimization in industrial solar cell manufacturing.

**Comment 2.** *Explain the Fe concentration measurement procedure and methods followed here for the experimental validation. In section 2.1, further clarification of experimental validation with details of number of cells per data samples, data collection methods and tools used will be helpful.*

**Reply:**

**Comment 3.** *Table 1 and PCA does not add much value.*

**Reply:**

We thank the Reviewer for this observation. In our study, we included PCA and Table 1 to transparently demonstrate that we systematically evaluated dimensionality reduction as a preprocessing step, which is a common practice in machine learning workflows.

However, as shown in our results, the initial set of features (relative changes in PV parameters, temperature, doping, and thickness) is already compact and physically meaningful, with limited redundancy. As a result, PCA did not provide a significant reduction in dimensionality or improvement in model accuracy. In some cases, especially under AM1.5 illumination, PCA even led to a slight decrease in prediction performance, likely due to the loss of direct physical interpretability of the features.

We agree that, in the context of our current dataset, Table 1 and the PCA analysis do not add substantial value to the main findings. Nevertheless, we believe it is useful for readers to see that this option was considered and to understand why it may not be necessary for similar studies with small, well-chosen feature sets.

**Comment 4.** *In section 3.1" As expected, increasing the number of descriptors enhances model performance (Fig. 4 and Fig.S7). The only exception occurs under AM1.5 illumination with PCA, where adding a fifth descriptor may degrade predictions rather than improve them." And similar discussions about the impact of changing number of features and dimensions is included with respect to different algorithms. It will be good to have some insight into the reasons for this. Also, the physical significance or meaning of it.*

**Reply:**

**Why increasing the number of features generally improves performance:** adding more physically meaningful descriptors-such as relative changes in  $I_{SC}$ ,  $V_{OC}$ ,  $\eta$ , and  $FF$ , along with structural parameters like temperature, doping, and base thickness-provides the model with more information about the underlying device physics. Each additional feature captures a different aspect of how iron-related recombination affects photovoltaic performance, helping the model to distinguish iron effects from other influences. This typically leads to improved prediction accuracy, as the model can learn more nuanced relationships.

**Why adding features can sometimes degrade performance (especially with PCA or under AM1.5):** adding more features does not always guarantee better results. In some cases-such as under AM1.5 illumination with PCA-adding additional descriptors can introduce redundancy or noise, especially if the new feature is highly correlated with existing ones or does not provide independent information about iron concentration. PCA, in particular, transforms the original features into linear combinations (principal components) that may not always preserve the most physically relevant information for the specific prediction task. This can occasionally lead to a decrease in model accuracy, as seen in our results.

**Physical significance:** the physical meaning of these observations is that the most informative features for iron quantification are those most directly and sensitively affected by FeB dissociation (such as  $\epsilon I_{SC}$  and  $\epsilon \eta$ ). Including additional PV parameters or structural descriptors is beneficial up to the point where they add unique information. Beyond that, extra features may simply add noise or complexity, especially if the model or preprocessing (like PCA) cannot effectively filter out irrelevant components.

**Comment 5.** *Include reasons for improvement in prediction with certain combinations of features in the discussion section.*

**Reply:**

We thank the Reviewer for this valuable suggestion. The improvement in prediction accuracy observed with certain combinations of features is closely related to the physical relevance and complementarity of the selected descriptors.

Each photovoltaic parameter-such as the relative changes in  $\epsilon I_{SC}$ ,  $\epsilon V_{OC}$ ,  $\epsilon \eta$ , and  $\epsilon FF$ -reflects a different aspect of how iron-related recombination affects device performance after FeB pair dissociation. For example:

- $\epsilon I_{SC}$  is highly sensitive to changes in minority carrier lifetime, which is directly impacted by iron contamination;
- $\epsilon \eta$  and  $\epsilon FF$  capture broader effects on overall device performance, including losses due to recombination and resistive effects;
- $\epsilon V_{OC}$  is influenced by both recombination and the quality of the p-n junction.

Structural parameters such as temperature ( $T$ ), base thickness ( $d_p$ ), and boron doping ( $N_B$ ) further modulate how iron impacts the device, since the recombination dynamics and the extent of FeB dissociation depend on these factors.

Combining multiple PV parameters and structural descriptors provides the model with a more complete picture of the device's physical state. This allows the machine learning algorithm to distinguish between changes caused specifically by iron and those due to other factors (e.g., temperature or doping).

Some features (like  $\epsilon I_{SC}$ ) may show similar changes for different iron concentrations under certain conditions, but when combined with additional parameters (like  $\epsilon \eta$  or  $\epsilon V_{OC}$ ), the model can resolve these ambiguities and make more accurate predictions.

The impact of iron on PV parameters is not strictly linear, especially when considering interactions with temperature and doping. Including a carefully chosen set of features enables the model to capture these complex relationships.

However, as discussed, adding features that are highly correlated or not physically informative can introduce noise or redundancy, sometimes reducing accuracy-especially after dimensionality reduction (PCA) or with less flexible models.

We propose adding this explanation to the Discussion section (Section 3.1, after the paragraph discussing Figs. 4 and S7), where the impact of feature number and combinations on model performance is already described. This will provide readers with both the physical intuition and practical implications behind the observed trends.

The observed improvement in prediction accuracy with certain feature combinations can be attributed to the complementary physical information provided by each descriptor. For instance,  $\epsilon I_{SC}$  is highly sensitive to iron-induced changes in carrier lifetime, while  $\epsilon\eta$ ,  $\epsilon V_{OC}$ , and  $\epsilon FF$  reflect additional aspects of device performance affected by iron contamination. Including temperature, doping, and thickness as features allows the model to account for structural and operational dependencies that modulate the impact of iron. Thus, combining these features enables the model to more accurately isolate and quantify the effects of iron, leading to better predictions. However, adding features that are redundant or weakly informative may introduce noise, which can sometimes reduce accuracy, particularly after PCA or under complex illumination conditions.

If the Reviewer is referring to all possible groupings or subsets of features, our study systematically evaluated several physically meaningful combinations, as described in Section 2.2 and shown in the results (e.g., Figs. 4 and S7). This approach allowed us to identify which combinations of descriptors yield the highest prediction accuracy, and to discuss the physical reasons for these outcomes.

Suggested manuscript addition (Section 3.1, after discussion of feature combinations):

In this study, we systematically evaluated various combinations of features, focusing on physically meaningful subsets relevant to iron-induced changes in photovoltaic parameters. For the machine learning models used, the order of input features does not affect the results; rather, the inclusion of complementary and independent features is key to improving prediction accuracy.

**Comment 6.** *Would this prediction be applicable beyond the range of  $N_B$ ,  $d_p$ ,  $T$  that was chosen in the study? why or why not?*

**Reply:**

The predictive models developed in our study were trained exclusively on simulated data within specific ranges of boron concentration ( $N_B$ ), base thickness ( $d_p$ ), and temperature ( $T$ ), as detailed in Section 2.1. As with most machine learning models, their accuracy and reliability are highest within the parameter space covered by the training data.

The predictions of our models are not guaranteed to be accurate outside the ranges of  $N_B$ ,  $d_p$ , and  $T$  used for training. This is because the models have not seen data from those regions and may not capture physical behaviors or interactions that occur at more "extreme" values.

Extrapolating to unseen parameter values can lead to unreliable or non-physical predictions, as the underlying device physics may change in ways not represented in the training data.

For example, at much higher or lower doping levels, base thicknesses, or temperatures, the impact of iron on photovoltaic parameters may differ due to changes in recombination mechanisms, carrier mobilities, or defect formation, which are not captured by the current model.

The applicability and reliability of the model's predictions beyond the studied parameter ranges ( $N_B$ ,  $d_p$ ,  $T$ ) will depend on which parameter is being extrapolated.

**Boron concentration:**

Extrapolating beyond the training range of boron doping is likely to result in significant errors. This is because  $N_B$  directly affects the formation and dissociation dynamics of FeB pairs, as well as the overall recombination activity in the device. The relationship between iron-related recombination and PV parameters is highly sensitive to  $N_B$ , so predictions outside the trained range may not be physically meaningful.

**Base thickness:**

In contrast, the base thickness has a much weaker influence on the photovoltaic response to iron contamination within the studied range. Our simulations and prior studies indicate that, for typical device designs, moderate changes

in  $d_p$  have only a minor effect on the relevant PV parameters and iron sensitivity. Therefore, extrapolation slightly beyond the studied  $d_p$  range may not degrade prediction accuracy as dramatically as for  $N_B$ .

#### **Temperature:**

The kinetics of FeB dissociation under carrier injection are influenced not only by temperature but also by recombination-enhanced defect reactions, further complicating the temperature dependence [ <https://doi.org/10.1155/2015/15457> ]. Therefore, extrapolating the model to temperatures outside the training range is likely to result in unreliable predictions, as the underlying physics governing iron-related recombination and FeB pair dynamics may change significantly.

**Comment 7.** *Discuss clearly the advantages and limitations of this model in conclusion. Eg practical significance, is it only for B doped cells only this thickness etc.*

#### **Reply:**

The proposed ML-based approach provides a non-destructive, rapid, and accurate method for quantifying iron concentration in boron-doped silicon solar cells using standard photovoltaic measurements. Its main advantages include ease of implementation, suitability for inline process monitoring, and the ability to capture complex, physically meaningful dependencies. However, the method is currently limited to the parameter ranges and device types represented in the training data, and its applicability is primarily for boron-doped c-Si cells with known structural parameters. Future work will focus on expanding the model's range and adapting the approach to other cell types and impurity systems.

Add this summary as a dedicated paragraph at the end of Conclusion section, right after main findings:

The proposed ML-based approach provides a non-destructive, rapid, and accurate method for quantifying iron concentration in boron-doped silicon solar cells using standard photovoltaic measurements. Its main advantages include ease of implementation, suitability for inline process monitoring, and the ability to capture complex, physically meaningful dependencies. However, the method is currently limited to the parameter ranges and device types represented in the training data, and its applicability is primarily for boron-doped c-Si cells with known structural parameters. Future work will focus on expanding the model's range and adapting the approach to other cell types and impurity systems.

## **References**