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. The location of revisions is pointed by red color and highlighted in yellow in "MarkedManuscript.pdf". 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 #1

Comment 1. *The introduction lacks the role of iron valency (ferrous or ferric) in silicon solar cells.*

Reply: Reviewer is correct that ferrous (Fe^{2+}) and ferric (Fe^{3+}) states deserve consideration, as they are the most common and stable charge states of iron. These ionic forms typically occur in compounds where iron forms chemical bonds (either ionic or covalent) with other elements, as well as in cases where iron is present as an impurity in solid materials. In silicon, iron can also exist in a trivalent (ferric) state when it substitutes for a silicon atom at a lattice site. However, under normal conditions, the concentration of substitutional iron is extremely low, less than 1% of the total iron impurity atoms [1]. Increasing the concentration of substitutional iron requires special sample processing, such as high-temperature annealing or irradiation. Moreover, substitutional iron acts as a weak recombination center, so its influence on the properties of silicon solar cells can be neglected. The ferrous form is virtually absent in silicon. The majority of iron impurity atoms in silicon occupy positions, where they can exist in either a neutral (Fe_i^0) or positively charged (Fe_i^+) state, depending on the position of the Fermi level [2]. In *n*-type silicon,Fe_i is more likely to exist in a neutral state, whereas in *p*-type silicon, it is more likely to be positively charged. Even when interstitial iron forms complex point defects, such as iron–boron pairs (Fe_iB_s), no valence bonds are formed. In interstitial configuration, iron acts as an active recombination center.

Consequently, in silicon solar cells, the role of iron valence is negligible, in contrast to other types of solar cells, such as perovskite solar cells [3]. Specifically, in $MAPbI_3$-based devices, both Fe^{3+} and Fe^{2+} point defects are observed, with Fe^{3+} being electronically inactive in terms of recombination.

We have added the relevant information to the Introduction (page 2, third paragraph from the top).

Comment 2. *The work didn't deal with the redox reaction of iron in silicon solar cells.*

Reply: Initially, it is important to note that redox reactions do not play a central role in silicon solar cells, unlike dye-sensitized solar cells or photoelectrochemical cells. However, changes in the charge states of individual atoms — whether intrinsic or impurity-related — occur during the photoelectric conversion process. For example, during the recombination of charge carriers at iron-related defects (such as interstitial iron atoms or iron-boron pairs), electron capture occurs in the initial stage:

$$\begin{aligned} Fe_i^+ + e^- &\rightarrow Fe_i^0\,, \\ (Fe_iB_s)^0 + e^- &\rightarrow (Fe_iB_s)^-\,, \end{aligned}$$

which typically corresponds to a reduction reaction. In the subsequent stage, hole capture alters the defect's charge state again

$$\begin{aligned} Fe_i^0 + h^+ &\rightarrow Fe_i^+\,, \\ (Fe_iB_s)^- + h^+ &\rightarrow (Fe_iB_s)^0\,, \end{aligned}$$

formally representing an oxidation reaction. When FeB pairs dissociate due to intense illumination, electron injection, or heating up to 200 °C, the iron atom becomes neutral by capturing an electron (reduction) during the first step of this two-stage process. After dissociation and the cessation of the external stimulus, the interstitial iron atom loses an electron (oxidation), becoming singly positively charged, which allows it to re-form a pair with a negatively charged dopant (boron). However, transitions between different charge states of iron-related defects in silicon are governed by electron and hole capture and emission processes, involving the exchange of carriers between the defect levels and the

conduction or valence band. These processes are described within Shockley–Read–Hall theory and do not constitute redox reactions in the classical chemical sense.

Consequently, the original manuscript did not emphasize redox reactions of iron in silicon solar cells, and the Reviewer's observation is valid. In the revised manuscript, we have included a note on the distinct nature of charge-state transitions in iron-related defects, highlighting how they differ from classical redox reactions (page 2, fourth paragraph from the top).

Comment 3. *In the title, "Extracting the iron concentration" should be changed to "Determination of iron concentration.*

Reply: We concur that the phrase "Determination of iron concentration" better emphasizes the quantitative aspect of the work and aligns well with the standard terminology used in similar studies. Accordingly, we have revised the manuscript title to: "Determination of iron concentration in silicon solar cells using photovoltaic parameters and machine learning."

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: We hope that the proposed approach will be applied to evaluate the concentration of iron in already installed solar modules and during the certification of photovoltaic converters. On the one hand, the selected temperature range of 290–340 K reflects the typical operating conditions of most silicon solar cells. On the other hand, the IEC 61215-2:2021 [4] standard requires that temperature coefficients of short-circuit current, open-circuit voltage, and maximum power be determined based on current–voltage measurements conducted within a temperature range of approximately +15 °C to +75 °C (288–348 K). In other words, the standard solar cell testing results can serve as input data for machine learning models trained within selected temperature range.

Additional information justifying the choice of temperature range has been added in the revised manuscript (page 4, second paragraph).

Complete dissociation of the FeB pairs was achieved during experiments conducted within the specified temperature range (further clarified below in response to the following Reviewer's comment).

The changes in photovoltaic parameter values resulting from FeB pair dissociation depend on both the concentration of iron impurities and the solar cell parameters (such as base thickness and dopant concentration), as well as the measurement conditions (illumination intensity and spectral composition, temperature). This issue is discussed in more detail in [5], which presents the results of silicon solar cell modeling used to create the training dataset. For instance, under AM1.5 illumination, the relative changes in short-circuit current range from −40.88% (for $N_{\mathrm{B}} = 10^{17}$ cm^{-3}, $T = 290$ K, $d_p = 380$ μm) to 6.11% (for $N_{\mathrm{B}} = 10^{15}$ cm^{-3}, $T = 315$ K, $d_p = 380$ μm). Similarly, the relative changes in open-circuit voltage vary from −12.95% (for $N_{\mathrm{B}} = 10^{15}$ cm^{-3}, $T = 340$ K, $d_p = 180$ μm) to 1.92% (for $N_{\mathrm{B}} = 1.778{\times}10^{15}$ cm^{-3}, $T = 340$ K, $d_p = 230$ μm). Notably, the most pronounced changes in solar cell characteristics do not necessarily occur at the highest iron concentrations. For example, the maximum negative change in open-circuit voltage (−12.95%) is observed at an iron concentration of 3.162×10^{12} cm^{-3}, rather than at the highest tested level.

This suggests that FeB pair dissociation, depending on the specific parameters, can not only degrade solar cell performance but also improve certain characteristics [5, 6]. However, it is important to emphasize that light-induced dissociation of FeB pairs does not lead to irreversible changes in photovoltaic parameters. Once the illumination ceases, the FeB pairs re-form, and the solar cell characteristics return to their original values [7, 8, 9, 10, 7, 8].

Information regarding the ranges of changes in photovoltaic parameters and their reversibility has been added to the revised manuscript (page 4, fourth paragraph from the top; page 3, first paragraph in Section 2.1).

We fully agree with the Reviewer that, within the chosen temperature range, factors beyond iron concentration and the selected descriptors may also influence solar cell characteristics. This is especially relevant for real solar cells, which, in addition to iron, may contain other recombination-active impurities. To better isolate iron-related effects from the influence of other factors, we used relative changes in PV parameters as model inputs. Since models trained

in this way provided sufficiently accurate predictions of iron concentration from both simulated and experimentally measured current–voltage characteristics, we consider this approach well justified.

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: It is known [11] that strong optical illumination (>100 W/m^2) results in almost complete (>99%) dissociation of the FeB pairs. The dissociation rate R_d is influenced by the pair concentration N_{FeB} and overall carrier generation rate G, which is proportional to the illumination intensity [7, 8, 9, 10]:

$$R_d = K \left(\frac{G}{N_{\mathrm{FeB}}} \right)^2 , \tag{1}$$

where K is the constant of the material. Additionally, it is well established that R_d depends on temperature [12, 13], the spectral composition of illumination [14], and the presence of recombination channels other than those associated with iron-related defects [7, 8].

For experimental validation of the proposed models, samples with an iron concentration of 2×10^{11} to 4×10^{13} cm^{-3} were used. Dissociation of FeB pairs was induced by intense illumination from a halogen lamp (7000 W/m^2). The illumination duration required to achieve near-complete dissociation depended on the iron concentration and reached up to 400 s for the solar cell with the highest N_{FeB} values. When selecting the illumination intervals, we considered the results of a previous study [14], which showed that in samples with a similar defect composition and exposed to a comparable light source, complete dissociation at $N_{\mathrm{Fe}} = 9 \times 10^{12}$ cm^{-3} occurred within 20 s.

Details regarding the selection of the illumination duration can be found on page 4 (first paragraph).

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

Reply: Indeed, models trained on data obtained under AM1.5 illumination perform worse in almost all cases. We believe this can be attributed to the following reasons. It is well known that a regression model may achieve high accuracy by having the target variable be an unambiguous function of the descriptor. Moreover, the absolute value of the function's derivative should be as significant as possible. As shown earlier [5], the condition of uniqueness in the relationship between N_{Fe} and the changes in photovoltaic parameters under AM1.5 illumination is often not fulfilled in contrast to the case of monochromatic illumination. In addition, the ranges of changes in $\varepsilon I_{\mathrm{SC}}$, $\varepsilon V_{\mathrm{OC}}$, and $\varepsilon\eta$ under monochromatic illumination exceed those observed under AM1.5 illumination. For illustration, Fig. 1 shows some examples from [5]. The physical origin of the differences in the relationship between iron concentration and parameter changes under AM1.5 and monochromatic illumination lies in the fact that, in the former case, the generation of non-equilibrium charge carriers occurs throughout the entire volume of the solar cell, rather than being limited to the base region. As a result, the processes involved in photovoltaic conversion are more complex. In our view, these differences explain why the performance of ML models trained on AM1.5 illumination data is generally poorer. One way to address the ambiguity in the relationship between features and the target variable is to incorporate additional features into the model. As shown in the study (e.g., see Fig. 7 in manuscript), for AM1.5 illumination, increasing the number of descriptors often improves prediction accuracy, supporting our hypothesis.

This answer is incorporated in the text on page 17, second paragraph from the bottom.

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

Reply: We thank the Reviewer for the valuable comment. We fully agree that discussing potential reasons for the discrepancies in model performance is essential. In fact, our response to the previous comment already addresses part of this issue. Upon further examination of other cases, we note the following:

Increasing the number of descriptors while keeping the sample size constant leads to higher sparsity in the data space due to increased dimensionality, which in turn makes it more difficult for the model to generalize. In such cases, the model may also learn spurious correlations that do not generalize to test data — a risk that is particularly pronounced

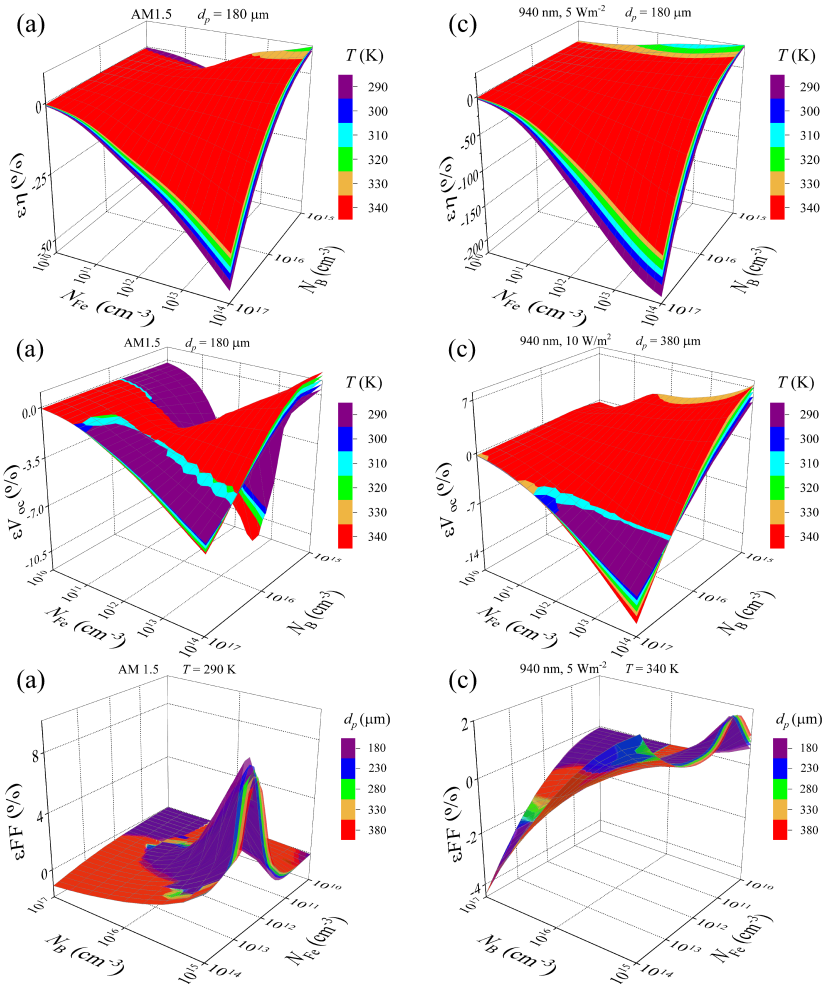

Figure 1: Relative changes in solar cell efficiency (first row), open-circuit voltage (second row), fill factor (third row) caused by a complete dissociation of Fe_iB_s. The dependencies are shown for two cases: AM1.5 illumination (left column) and 940 nm monochromatic light (right column). The figures are taken from reference [5].

in deep neural networks and deep tree-based ensemble methods. Moreover, a larger number of features results in more model parameters, increasing the risk of overfitting. These factors can collectively lead to a reduction in prediction accuracy. In our opinion, for $A^{AM1.5}$ models, both the positive effects achieved by increasing the number of descriptors (enhancing overall information variance and reducing ambiguity between the target variable and the input features) were dominant. However, for A^{940} models where this ambiguity was already low, the negative effects described above outweighed the benefits, which explains why prediction accuracy did not improve when using an expanded feature set.

Another effect that may arise when increasing the number of features is multicollinearity. It occurs when new features are highly correlated with existing ones and can lead to instability in model estimates. However, in our case, it is unlikely that multicollinearity among the photovoltaic parameters (incidentally, their correlation is significantly higher in the 940 nm case than in the AM1.5 case — see Fig. S1) affects prediction accuracy. First, all algorithms used are sufficiently robust to multicollinearity. Second, a standard method for mitigating multicollinearity is PCA. Despite previous successes in applying principal component analysis to various photovoltaics-related machine learning tasks [15, 16, 17, 18, 19] — and the presence of favorable conditions for its use in our case, such as the high correlation between changes in photoelectric parameters caused by FeB pair dissociation — applying PCA as a data pre-processing step for evaluating iron impurity concentration did not improve the maximum prediction accuracy in most cases. One likely reason is that PCA preserves the overall variance in the data but does not account for the specific relationships between individual features and the target variable. For instance, some photovoltaic parameters may be

more informative for predicting low iron concentrations, while others may be more relevant for high concentrations. For instance, εFF) are shown [5] to be most pronounced at low iron concentrations, unlike other PVPs. When these components are merged through PCA transformation, such distinctions may be lost, leading to reduced accuracy. Another possible reason is that principal components — being linear combinations of original features — do not support localized decision splits (i.e., those based on individual feature thresholds), which are particularly effective in tree-based models.

When comparing the effectiveness of different algorithms, several points are worth noting. The poorest performance observed with SVR is likely due to its limited ability to model complex nonlinear relationships — such as those that apparently exist between iron concentration and variations in photoelectric parameters. In contrast, the superior performance of the XGB and DNN models is attributable to their ability to effectively capture nonlinear patterns, which in our case become more prominent with increasing descriptor dimensionality. Specifically, XGB performs best in relatively simpler scenarios, such as those involving 940 nm illumination (see Figs. 7i-7l, 9i-9l). This is because XGB requires less training data than DNN (12,000 examples are sufficient, though this is not a particularly large amount for a neural network), handles high-quality structured data well (simulated data are noise-free), and is more resilient to the influence of correlated features. At the same time, DNNs are more powerful tools for nonlinear modeling, capable of uncovering hidden structures in the data. This enables them to achieve better results on more complex tasks — such as the transition to $A^{AM1.5}$-models with high-dimensional features (see Figs. 7m-p, 9m-9p). As shown by the results in Subsection 3.4, successful prediction within the N_B–altered also requires accounting for higher-order relationships, which explains why DNN models achieved the best performance (Figs. 11e-11l). Regarding the RF and GB algorithms, as indicated in Tables 1 and 2, they tend to perform best only on less complex data (i.e., small sets of descriptors), where more powerful models are at higher risk of overfitting.

Reply has been incorporated into revised manuscript as Subsection 3.5.

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

Reply: When exploring potential strategies for improving model accuracy, it is essential to consider the efforts already undertaken towards this goal. Specifically, data normalization was applied, multiple descriptor sets were tested, and the impact of reducing feature correlation using PCA was evaluated. In addition, the main model parameters were optimized (see Tables S1–S5 and Figure S2) through bayesian optimization using the Optuna package, and cross-validation was employed.

Among the remaining justified options, the following approaches can be highlighted:

1. Expanding the training dataset, primarily by incorporating a greater number of boron concentration values into the database. This strategy aligns with a fundamental principle of machine learning—the need for as much data as possible.
2. Deriving new descriptors from existing ones that are more directly and simply related to iron concentration. These descriptors should not be limited to linear combinations, such as those used in PCA; instead, t heir development should be guided by underlying physical processes. Alternatively, generative models, such as variational autoencoders, can be employed to discover latent variables. Notably, this approach may also facilitate easier dataset expansion.
3. Incorporating additional features, especially those obtainable without extra measurements, is desirable because the requirements of additional experiments would substantially reduce the practicality of the proposed method. For example, the minority carrier diffusion length could serve as a valuable descriptor; however, to our knowledge, no methods exist for its determination based solely on single or dual I-V characteristic measurements. The most apparent step in this direction involves the use of relative changes in voltage (εV_{MP}) and current (εI_{MP}), corresponding to the maximum power point, as supplementary features. However, it should be noted that these quantities exhibit correlation with other photovoltaic parameters (see Fig. 2).

Separately, a strategy for improving the quality of the training dataset, which involves increasing the realism of the simulation model, can be distinguished. For example, instead of employing SCAPS, three-dimensional simulators such as Synopsys Sentaurus TCAD, Silvaco ATLAS TCAD, or Solcore can be used to generate labeled datasets that account for the presence of additional impurities commonly found in Cz-Si, such as oxygen and carbon. Furthermore, incorporating experimental uncertainty into the training data by introducing noise can enhance the model's robustness and predictive accuracy.

This answer is incorporated in the revised text in Subsection 3.7.

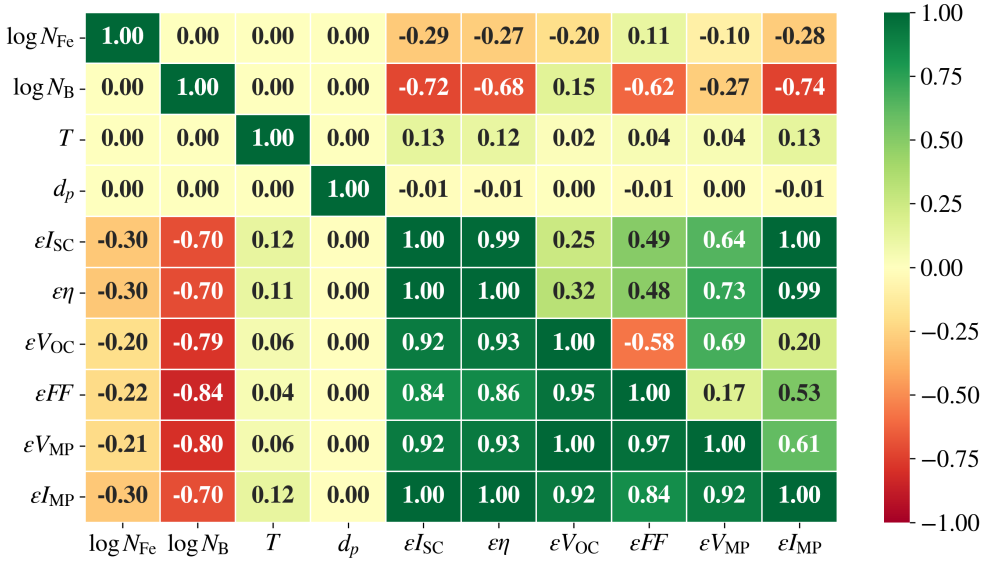

Figure 2: Correlation plot of features in hypothetical expanded training set. Data above and below the main diagonal correspond to AM1.5 and 940 nm illumination, respectively.

Comment 6. *There are some typing mistakes as well. for e.g. Page 14, line 44, t-ltered…*
Reply: We thank the Reviewer for pointing out the typographical errors and apologize for their presence. We have carefully reviewed the entire manuscript and corrected typographical issues to improve its clarity and professionalism.

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: The authors fully agree with the reviewer's assessment concerning the importance and, indeed, the necessity of including a synthesis in the Introduction that outlines the primary objective of this work and highlights the advantages of the proposed approach.

The primary objective of this study is to introduce a novel methodology for quantifying iron impurities in silicon solar cells and to examine the specific aspects of its implementation. The proposed approach utilizes variations in photovoltaic parameters resulting from FeB pair dissociation as input features for machine learning models. Due to the ability of machine learning algorithms to capture latent correlations and the simplicity of applying trained models, the method is characterized by notable ease of use. The required measurements for generating input data, specifically the acquisition of two current–voltage characteristics or potentially only two short-circuit current values, are fast, non-destructive, do not require prior sample preparation, and can be performed on fully fabricated solar cells without specialized equipment. These features clearly distinguish the proposed method from traditional techniques. An additional advantage is that the necessary equipment is also used for solar cell characterization, which enables straightforward integration of the proposed approach into routine procedures employed in the quality control or certification of photovoltaic devices.

The text has been revised according to the Reviewer's advice — see page 2-3, last paragraph in Section 1.

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: We used a set of n^+-p-p^+-Si samples for the experimental validation. The structure was fabricated from a 380 μm thick p-type boron-doped Czochralski silicon (100) wafer with a doping level of $N_B = 1.36 \times 10^{15}$ cm^{-3}. The n^+–layer, with a sheet resistance of $20 - 30$ $\Omega/\square$ and a thickness of 0.7 μm, and the p^+–layer, with a sheet resistance of $10 - 20$ $\Omega/\square$ and a thickness of 0.6 μm, were formed by phosphorus and boron diffusion, respectively.

The iron concentration was determined using a methodology described in [20, 21], which is based on fitting the kinetics of the short-circuit current following FeB pairs dissociation, measured under low-intensity monochromatic illumination. The kinetics of I_{SC} were measured at 340 K, which allowed tracking the complete recovery of FeB pairs within 1 hour only (the characteristic recovery time being approximately 600 s at 340 K, compared to 13,000 s at 300 K [8, 9]). The determined iron concentration ranged from 2×10^{11} to 4×10^{13} cm^{-3} across the various solar cells.

To create the experimental test dataset, we measured the I-V characteristics under monochromatic illumination, both before and immediately after FeB pair dissociation. Photovoltaic parameters were extracted from the I-V curves using a standard procedure [22], followed by calculations of their relative changes. For each solar cell, measurements were performed at 1-3 temperatures within the range of 305-340 K.

The similar experimental setup was used both for N_{Fe} determination and experimental test dataset creation. The current-voltage characteristics and Isc kinetics were measured using a Keithley 2450 source meter. The light–emitting diode SN–HPIR940nm–1W with light wavelength 940 nm and intensity of about 5 W/m^2 was used as low–intensity monochromatic light source. The LED radiation intensity was stabilized by a W1209 thermostat and a power supply regulated by a circuit incorporating positive feedback and digital control. The solar cell temperature was driven by a thermoelectric cooler controlled by an STS-21 sensor and maintained constant by a PID algorithm embedded in the software that serves the experimental setup.

A total of 20 solar cells were used for experimental validation, and 30 I-V curves were analyzed. However, at low temperatures and low iron concentrations, the expected relative changes in the fill factor were below 0.5%, making it difficult to determine the εFF with sufficient accuracy. As a result, εFF could be reliably determined in only 8 cases (corresponding to 5 samples). Consequently, the number of samples in the experimental datasets used to test models trained on different descriptor sets (described in detail in Subsection 2.2) ranged from 30 to 8.

The detailed information can be found in revised manuscript (last five paragraphs in Subsection 2.1).

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

Reply: In general, principal component analysis (PCA) is a widely used and effective tool in machine learning, particularly for addressing problems in photovoltaics. For example, it has been successfully applied to fault detection and classification based on I–V curves [16, 15], performance prediction of organic [17] and perovskite [18] solar cells, as well as measurement of the internal quantum efficiency in GaAs solar cells [19].

Despite previous successes in applying principal component analysis to various photovoltaics-related machine learning tasks — and the presence of favorable conditions for its use in our case, such as the high correlation between changes in photoelectric parameters caused by FeB pair dissociation — applying PCA as a data pre-processing step for evaluating iron impurity concentration did not improve the maximum prediction accuracy. One likely reason is that PCA preserves the overall variance in the data but does not account for the specific relationships between individual features and the target variable. For instance, some photovoltaic parameters may be more informative for predicting low iron concentrations, while others may be more relevant for high concentrations. For instance, εFF) are shown [5] to be most pronounced at low iron concentrations, unlike other PVPs. When these components are merged through PCA transformation, such distinctions may be lost, leading to reduced accuracy. Another possible reason is that principal components — being linear combinations of original features — do not support localized decision splits (i.e., those based on individual feature thresholds), which are particularly effective in tree-based models.

At the same time, we consider it important to document that this commonly used technique was systematically evaluated and found to be less effective than using the original features for iron concentration determination. Conversely, we fully agree that Table 1 does not offer substantial additional value and have therefore excluded it from the revised manuscript.

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: We thank the Reviewer for the valuable comment. We fully agree that discussing potential reasons for the discrepancies in model performance is essential.

We note that this comment is similar in content to Comments 3 and 4 from Reviewer #2. We trust that Reviewer #4 will be understanding that, in response to this remark, we do not repeat the content of previous replies, but instead note that the relevant information has been incorporated into the revised manuscript in Subsection 3.5.

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

Reply: First of all, it is important to note that the selection of feature combinations was guided by the degree of correlation between variations in individual photovoltaic parameters and the iron concentration (see Fig. S1 in the Supplementary Material). Specifically, the 4-dimensional feature set included the measurement conditions (T), solar cell characteristics (d_p and N_B), and the PVP most strongly correlated with $\log N_\mathrm{Fe}$, i.e. $\varepsilon I_\mathrm{SC}$. In extending the set to five dimensions, we added the parameter most strongly correlated with N_Fe among the remaining unused ones — namely, $\varepsilon\eta$ — to the existing features. Subsequent steps followed a similar procedure.

The presence of iron impurities in silicon increases the recombination rate of charge carriers. However, each photovoltaic parameter captures a different aspect of how iron-related recombination influences solar cell performance. Specifically, the short-circuit current reflects the photogenerated current (I_ph), which primarily depends on the diffusion length and recombination in the quasi-neutral region. In contrast, the open-circuit voltage is influenced not only by I_ph but also by the saturation current (I_0) and the ideality factor (n), both of which are determined by recombination processes in the space-charge region [23]: $V_\mathrm{OC} = nkT\left[\ln\left(I_\mathrm{ph}/I_0\right)+1\right]$. The fill factor, in turn, is less sensitive to variations in n [24]:

$$FF \approx \frac{\upsilon_\mathrm{OC} - \ln\left(0.72+\upsilon_\mathrm{OC}\right)}{\upsilon_\mathrm{OC}+1}, \tag{2}$$

where υ_OC being the normalized open-circuit voltage $\upsilon_\mathrm{OC} = qV_\mathrm{OC}/nkT$. Ultimately, the efficiency integrates all of these effects: $\eta \propto I_\mathrm{SC}\cdot V_\mathrm{OC}\cdot FF$.

Incorporating multiple descriptors associated with variations in photovoltaic parameters provides a more comprehensive representation of the effects of iron impurities. This expanded descriptor set increases the informational diversity available for model training, which can enhance predictive accuracy. However, adding more features also raises the risk of overfitting and increases data sparsity, which may hinder the model's ability to generalize. Therefore, simply increasing the number of features does not guarantee improved performance. Instead, optimal predictive accuracy is typically achieved with specific, well-chosen combinations of descriptors.

The rationale for selecting specific feature combinations was added in the third paragraph of Subsection 2.2, while the physical basis for the observed improvement in prediction accuracy with an increased number of descriptors was included in the second paragraph of Subsection 3.5. A more detailed discussion of the adverse effects associated with expanding the feature set is provided in the response to Comment 4 from Reviewer #2 and in the third and fourth paragraphs of Subsection 3.5 of the revised manuscript.

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

Reply: We want to thank the Referee for the insightful question. To tackle this problem, it is first necessary to consider each parameter separately. The ability of the models to make predictions beyond the range of the training dataset depends on the relationship between the target variable and the descriptors. As previously shown [5], the relationship between the relative changes in the photoelectric parameters and iron concentration can vary significantly depending on the temperature and the concentration of the doping impurity in the base material. This is because a number of parameters, including those of silicon such as bandgap, carrier mobilities, intrinsic recombination coefficients, and Fermi level position, as well as iron-related defects such as carrier capture cross-sections, depend on temperature and

carrier concentration. These relationships, which were taken into account when creating the training set, are sufficiently complex, particularly with respect to N_B, as demonstrated by the results obtained for the N_B–altered dataset. As previously mentioned, even the task of interpolation within the range of N_B used during training was only satisfactorily addressed by the DNN models due to their high ability to identify hidden patterns. Predicting beyond the training range of N_B is even more challenging. Therefore, in our view, the capabilities of the developed models are quite limited.

The temperature-related relationships are simpler, as demonstrated by the reasonably good results obtained using tree-based models for the T–altered dataset (see Fig. 9, Table 2). However, for the training set that included T values beyond the original range, the performance of these models degraded significantly (see Fig. 10). Once again, the best average results were achieved by the deep neural network models (see Fig. 9), and only with their use can relatively accurate predictions beyond the training range of temperature be expected.

The situation is more straightforward concerning the base width, as its influence on the dependence of N_{Fe} on variations in PVPs value is much weaker and often approximately linear [5]. Therefore, in our opinion, the developed models can be expected to remain applicable for predictions beyond the range selected in this study. Naturally, the extension of the base thickness range is only possible within certain limits, until the ratio between d_p and the diffusion length of minority carriers changes significantly.

In summary, we note that if predictions are required for cases outside the range of parameter values used during training, it is advisable to rely exclusively on DNN models. Among the considered parameters, base thickness allows for the largest permissible deviations, whereas for temperature, deviations are unlikely to exceed 10 K. The use of the developed models with doping impurity concentrations different from those in the training set is not recommended. This requirement effectively serves as an additional condition that the training set must satisfy to ensure satisfactory prediction accuracy. It is worth noting that in our study, the parameter ranges were selected to closely correspond to those of real solar cells and their operating or testing conditions.

A detailed discussion of the applicability of the models beyond the selected parameter ranges is provided in Subsection 3.7 (last four paragraphs).

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: From a practical standpoint, the developed models are applicable for iron quantification in monocrystalline silicon solar cells with an n^+-p-p^+ structure, featuring a boron-doped base and a wide range of base thicknesses ($180 - 380$ μm) and doping concentrations ($10^{15} - 10^{17}$ cm^{-3}). The proposed ML-based approach provides a non-destructive, fast, and simple method for determining iron concentrations that does not require specialized equipment. The methodology used to develop models can be adapted for other types of solar cells and/or for quantifying other impurities.

Thanks to the Reviewer's comment, we supplemented Conclusion section (second to last paragraph).

References

[1] E. Wright, J. Coutinho, S. Öberg, V. J. B. Torres, Mössbauer parameters of fe-related defects in group-iv semiconductors: First principles calculations, J. Appl. Phys. 119 (2016) 181509.

[2] A. A. Istratov, H. Hieslmair, E. Weber, Iron and its complexes in silicon, Appl. Phys. A: Mater. Sci. Process. 69 (1999) 13–44.

[3] J. R. Poindexter, R. L. Z. Hoye, L. Nienhaus, R. C. Kurchin, A. E. Morishige, E. E. Looney, A. Osherov, J.-P. Correa-Baena, B. Lai, V. Bulović, V. Stevanović, M. G. Bawendi, T. Buonassisi, High tolerance to iron contamination in lead halide perovskite solar cells, ACS Nano 11 (2017) 7101–7109.

[4] IEC 61215-2:2021 - Terrestrial photovoltaic (PV) modules – Design qualification and type approval – Part 2: Test procedures, International Standard IEC 61215-2:2021, Geneva, Switzerland, 2021. https://webstore.iec.ch/publication/61174.

[5] O. Olikh, O. Zavhorodnii, Iron's impact on silicon solar cell execution: Comprehensive modeling across diverse scenarios, Materials Science and Engineering: B 317 (2025) 118192.

[6] J. Schmidt, Effect of dissociation of iron–boron pairs in crystalline silicon on solar cell properties, Progress in Photovoltaics: Research and Applications 13 (2005) 325–331.

[7] L. J. Geerligs, D. Macdonald, Dynamics of light-induced feb pair dissociation in crystalline silicon, Appl. Phys. Lett. 85 (2004) 5227–5229.

[8] 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.

[9] N. Khelifati, H. S. Laine, V. Vähänissi, H. Savin, F. Z. Bouamama, D. Bouhafs, Dissociation and formation kinetics of iron–boron pairs in silicon after phosphorus implantation gettering, Phys Status Solidi A 216 (2019) 1900253.

[10] 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.
[11] 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.
[12] J. Lagowski, P. Edelman, A. M. Kontkiewicz, O. Milic, W. Henley, M. Dexter, L. Jastrzebski, A. M. Hoff, Iron detection in the part per quadrillion range in silicon using surface photovoltage and photodissociation of iron-boron pairs, Appl. Phys. Lett. 63 (1993) 3043–3045.
[13] K. Lauer, C. Möller, D. Debbih, M. Auge, D. Schulze, Determination of activation energy of the iron acceptor pair association and dissociation reaction, in: Gettering and Defect Engineering in Semiconductor Technology XVI, volume 242 of *Solid State Phenomena*, Trans Tech Publications Ltd, 2016, pp. 230–235.
[14] O. Olikh, O. Datsenko, S. Kondratenko, Influence of illumination spectrum on dissociation kinetics of iron–boron pairs in silicon, Phys. Status Solidi A 221 (2024) 2400351.
[15] W. Gao, R.-J. Wai, A novel fault identification method for photovoltaic array via convolutional neural network and residual gated recurrent unit, IEEE Access 8 (2020) 159493–159510.
[16] S. Fadhel, C. Delpha, D. Diallo, I. Bahri, A. Migan, M. Trabelsi, M. Mimouni, PV shading fault detection and classification based on I-V curve using principal component analysis: Application to isolated PV system, Sol. Energy 179 (2019) 1–10.
[17] T. W. David, G. A. Soares, N. Bristow, D. Bagnis, J. Kettle, Predicting diurnal outdoor performance and degradation of organic photovoltaics via machine learning; relating degradation to outdoor stress conditions, Prog. Photovoltaics Res. Appl. 29 (2021) 1274–1284.
[18] Y. Liu, W. Yan, S. Han, H. Zhu, Y. Tu, L. Guan, X. Tan, How machine learning predicts and explains the performance of perovskite solar cells, Sol. RRL 6 (2022) 2101100.
[19] 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.
[20] O. Olikh, V. Kostylyov, V. Vlasiuk, R. Korkishko, R. Chupryna, Intensification of iron–boron complex association in silicon solar cells under acoustic wave action, J. Mater. Sci.: Mater. Electron. 33 (2022) 13133–13142.
[21] O. Olikh, V. Kostylyov, V. Vlasiuk, 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.
[22] B. Paviet-Salomon, J. Levrat, V. Fakhfouri, Y. Pelet, N. Rebeaud, M. Despeisse, C. Ballif, New guidelines for a more accurate extraction of solar cells and modules key data from their current–voltage curves, Prog. Photovoltaics Res. Appl. 25 (2017) 623–635.
[23] D. Yang (Ed.), Handbook of Photovoltaic Silicon, Springer Berlin, Heidelberg, first edition, 2019.
[24] M. A. Green, Accuracy of analytical expressions for solar cell fill factors, Sol. Cells 7 (1982) 337–340.