Identification of the dominant recombination process for perovskite solar cells based on machine learning

Le Corre, Vincent M.; Sherkar, Tejas S.; Koopmans, Marten; Koster, L. Jan Anton

Published in:
Cell Reports Physical Science

DOI:
10.1016/j.xcrp.2021.100346

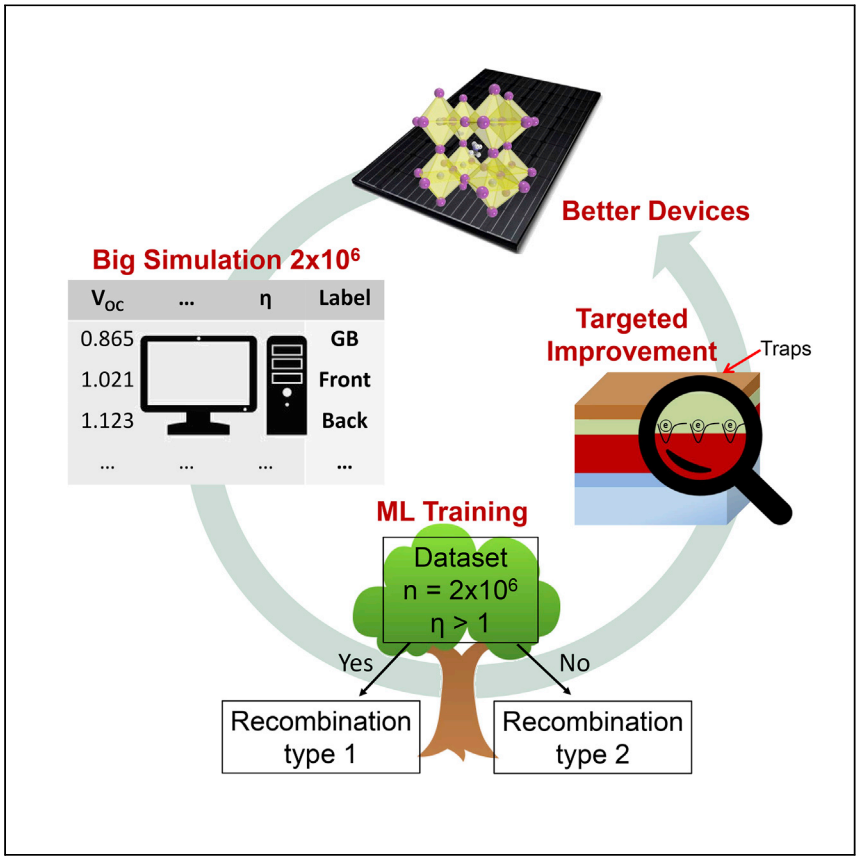
Publication date:
2021

Cell Reports
Physical Science

CellPress
OPEN ACCESS

Article

Identification of the dominant recombination process for perovskite solar cells based on machine learning



Le Corre et al. demonstrate the application of machine learning methods to identify the dominant recombination process in perovskite solar cells with 82% accuracy. The machine learning algorithms are trained and tested using large-scale drift-diffusion simulations, and their applicability on real solar cells is also demonstrated on devices previously reported.

Vincent M. Le Corre, Tejas S. Sherkar, Marten Koopmans, L. Jan Anton Koster

v.m.le.corre@rug.nl (V.M.L.C.)
l.j.a.koster@rug.nl (L.J.A.K.)

HIGHLIGHTS

In-depth analysis of the ideality factor of perovskite solar cells

Accurate prediction of the dominant recombination using machine learning methods

Le Corre et al., Cell Reports Physical Science 2, 100346
February 24, 2021
https://doi.org/10.1016/j.xcrp.2021.100346

Cell Reports
Physical Science

CellPress
OPEN ACCESS

Article

Identification of the dominant recombination process for perovskite solar cells based on machine learning

Vincent M. Le Corre,[1,*] Tejas S. Sherkar,[1] Marten Koopmans,[1] and L. Jan Anton Koster[1,2,*]

SUMMARY

Over the past decade, perovskite solar cells have become one of the major research interests of the photovoltaic community, and they are now on the brink of catching up with the classical inorganic solar cells, with efficiency now reaching up to 25%. However, significant improvements are still achievable by reducing recombination losses. The aim of this work is to develop a fast and easy-to-use tool to pinpoint the main losses in perovskite solar cells. We use large-scale drift-diffusion simulations to get a better understanding of the light intensity dependence of the open-circuit voltage and how it correlates to the dominant recombination process. We introduce an automated identification tool using machine learning methods to pinpoint the dominant loss using the light intensity-dependent performances as an input. The machine learning was trained using >2 million simulations and gives an accuracy of the prediction up to 82%.

INTRODUCTION

Big data science and machine learning (ML) have drawn a lot of attention, not only from industries such as Google, Facebook, and Amazon but also in the scientific community. There has been a tremendous increase in the number of studies using these techniques and in very different fields, from medicine to chemistry and physics.[1–9] ML techniques have proved to be very effective at predicting the properties of materials[3,9,10] and at speeding up the material discovery process by suggesting new promising structures.[4–8]

In an era in which high-throughput experimentation is made possible thanks to the use of autonomous robots[4,11] and in which a large amount of data can be processed easily with the new data science and artificial intelligence (AI) tools openly accessible online, science should turn to a more systematic/standard experimentation, data collection, and analysis methodology. This would allow scientists to spend more time exploring data or new ideas rather than going through tedious and sometimes poorly reproducible lab work. If this line of thinking is applied to photovoltaic research, the future workflow of material and device development for solar cells may look like Figure 1. In recent years, there have been only a few attempts to develop tools corresponding to one or more of the steps described in Figure 1. These studies showed that AI methods helped those in experimental planning to predict the next promising material to use/synthesize or experiment to undertake.[3,5–7] The use of robots for high-throughput experiments[7,12,13] and automated data analysis[3,7,11] provides several advantages in terms of speed of material development and size/reproducibility of the experimental output, demonstrating the power of this approach.

[1]Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG, Groningen, the Netherlands

[2]Lead contact

*Correspondence: v.m.le.corre@rug.nl (V.M.L.C.), l.j.a.koster@rug.nl (L.J.A.K.)

https://doi.org/10.1016/j.xcrp.2021.100346

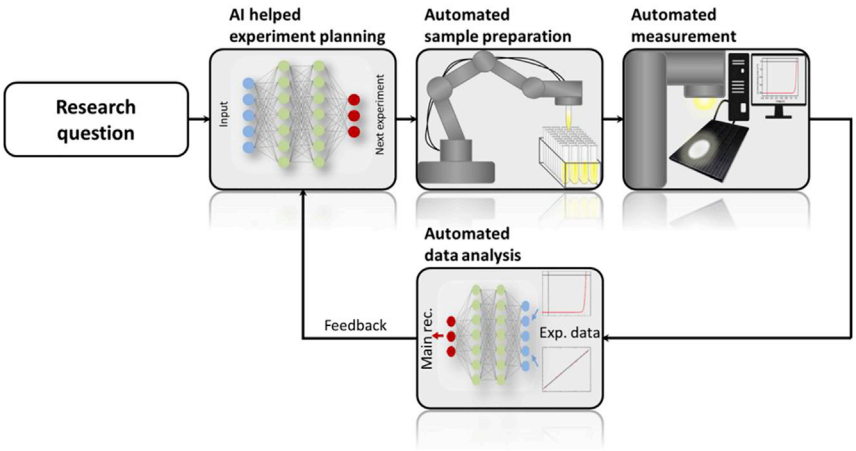


Figure 1. Workflow for automated research
Potential closed-loop workflow for automated and high-throughput experimentation in photovoltaic research.

To pursue this reasoning for the study of perovskite solar cells (PSCs), we need appropriate characterization tools that can accurately pinpoint whether the bulk or an interface limits efficiency while being fast/easy and can also automatically analyze data. Here, we build a characterization tool to pinpoint bulk or interfaces recombination as the dominant loss based on light intensity-dependent current-voltage characteristics measurements.

Hybrid halide perovskites show desirable properties for use in thin-film solar cells, such as a large absorption coefficient, high charge carrier mobilities, and long diffusion lengths.[14–16] As a consequence, PSCs have attracted great interest, facilitating a remarkable increase in their efficiency over the past 10 years.[14,17] This increase has been driven mainly by novel fabrication techniques that ensure very compact perovskite absorber films[18–21] with large grain sizes,[22] minimizing non-radiative recombination losses through charge trapping at defects in the perovskite absorber bulk at the grain boundaries.

Efforts are being directed toward the tuning of electron and hole transport layers (ETLs, HTLs)[23–25] and interface engineering.[26–28] Nonetheless, in existing PSCs, Shockley-Read-Hall (SRH) trap-assisted recombination is still the dominant recombination loss, whether it happens in the bulk via grain boundaries or at the interfaces (HTL/perovskite and perovskite/ETL).[29–37] There exist many reports on the recombination in PSCs; however, most of these experimental studies are performed on perovskite thin films.[38–44]

Photoluminescence (PL) studies look at charge transfer rates at interfaces between perovskite thin films and transport layers, revealing the recombination kinetics.[24,43] Although these methods elucidate the fundamental optoelectronic properties and role of defect physics in recombination, they are often performed under non-operating solar cell conditions (i.e., non-solar fluences and/or not a full device configuration) and are often complex.[24,43] More recent reports also suggest that for high-efficiency perovskite, the main recombination center is consistently located at the interface between the perovskite and the TLs.[33,34] It is therefore important to identify the dominant recombination losses, so that efforts can be directed toward improving the quality of the perovskite material or the interface in question. While the PL measurements have proven to be very effective at pinning out the dominant

loss, they require several sets of devices with different structures, which make them very time-consuming.

Here, we present a more in-depth analysis on the light intensity measurement of PSCs and especially on the parameters influencing the slope ($\eta kT/q$ on a semi-log plot) of the open-circuit voltage (V_{OC}) versus light intensity curve, with η the so-called ideality factor, T the temperature, k the Boltzmann constant, and q the elementary charge. The ideality factor is typically used to conclude on the dominant recombination process very simply by saying that an ideality factor close to 1 means that band-to-band recombination dominates[45] and an ideality factor close to 2 corresponds to bulk SRH trap assisted recombination,[46] while when it is in between these two values, it is difficult to conclude.

However, here, we show that this interpretation is not complete, as other parameters influence the ideality factor and that the analysis of ideality factor needs to be coupled with other key figures such as the open-circuit voltage (V_{OC}), the short-circuit current (J_{SC}),the fill factor (FF), and doping. To remedy this issue, we trained—using large-scale drift-diffusion (DD) simulations[31,32,47]—a ML algorithm based on decision trees to classify the light intensity measurement output to give the most likely dominant recombination. Our trained model was able to determine, within a matter of seconds, whether the dominant recombination process was band-to-band, interface, or grain boundaries trap recombination with >80% accuracy, and this while only requiring the simple measurement of light intensity-dependent current-voltage characteristics (JVs). This represents, to the best of our knowledge, the first use of ML as identification tools of the dominant recombination process in PSCs over such a wide range of material properties. In a broader perspective, the ML approach could also be used for automated data analysis, as in the workflow described in Figure 1, as the experimental data used are simple, can be easily measured, and the ML algorithm provides a fast and easy way to analyze the data and provide feedback to choose the next relevant experiment.

RESULTS AND DISCUSSION

Relationship between ideality factor and recombination processes

The amount of charge recombination is directly related to the open-circuit voltage of a solar cell, whose light intensity dependence can reveal information about the dominant recombination mechanism under operating conditions. In this section, we describe the different possible scenarios that are relevant for PSCs: (1) band-to-band recombination, (2) SRH trap-assisted recombination, and (3) SRH trap-assisted recombination, with one pinned charge carrier density. Since Auger recombination is negligible under non-concentrated illumination,[48,49] it is excluded from this analysis.

The V_{OC} of a solar cell is the difference between the electron and hole quasi-fermi levels, and is approximated analytically by[45,50]

$$qV_{OC} = E_{gap} - kT\ln\left(\frac{N_{cv}^2}{np}\right) \qquad \text{(Equation 1)}$$

where E_{gap} is the band gap of the photoactive material, N_{cv} is the density of states in the conduction and valence bands, while n and p are electron and hole concentrations, respectively. The ideality factor is given by how the np-product depends on the light intensity. This yields to different η for different scenarios.

The total recombination rate in the solar cell is given by

$$R = R_b + R_{SRH}^{Bulk} + R_{SRH}^{Front\ int.} + R_{SRH}^{Back\ int.} \quad \text{(Equation 2)}$$

where R_b is the band-to-band recombination rate and R_{SRH} is the trap-assisted recombination rate described by the SRH statistics in the bulk (R_{SRH}^{Bulk}) of the perovskite or at the interface with the front ($R_{SRH}^{Front\ int.}$) and back ($R_{SRH}^{Back\ int.}$) TLs. At V_{OC}, the total recombination rate (R) is equal to the charge generation rate (G) in the device. We used our DD simulations (available open source on GitHub[47]) to study the different cases with dominant band-to-band recombination or SRH trap-assisted recombination with and without one pinned charge carrier density one at a time for a simple device structure (see the parameters used in the Supplemental experimental procedures). For more datails on the DD simulation please see Supplemental experimental procedures and Figure S4.

Dominant band-to-band recombination

When band-to-band recombination is dominant ($R_b \gg R_{SRH}^{Bulk} + R_{SRH}^{Front\ int.} + R_{SRH}^{Back\ int.}$) in the device, at V_{OC} the total recombination rate is

$$R \approx R_b = \gamma np = G. \quad \text{(Equation 3)}$$

So,

$$n = p = (G/\gamma)^{1/2}, \quad \text{(Equation 4)}$$

where γ is the band-to-band recombination constant. Following Equation 1, the V_{OC} is now approximated by

$$qV_{OC} = E_{gap} - kT\ln\left(\gamma N_{cv}^2 \times \frac{1}{G}\right). \quad \text{(Equation 5)}$$

Since the charge generation rate (G) is proportional to the light intensity (I), the semi-log plot of V_{OC} versus I gives a slope of kT/q or a ideality factor of $\eta = 1$. Therefore, $\eta = 1$ could indicate dominant band-to-band recombination in the device.[45] Note that for PSCs, γ has been found to be low[51–54] in the range of 10^{-9}–10^{-11} $cm^3\ s^{-1}$. This could be explained by lattice distortion leading to a spatial separation of electrons and holes decreasing the probability of charge carriers to recombine. As a consequence, a hypothetical PSC in which band-to-band recombination dominates must have a very high-quality perovskite and interfaces with very low defects density, which would lead to high V_{OC} as we show later in the article.

Dominant SRH recombination without pinning of one charge carrier density

If SRH recombination is dominant ($R_{SRH} >> R_b$) in a device, at V_{OC}, the total recombination rate is

$$R \approx R_{SRH} = \frac{C_n C_p \Sigma_T}{C_n(n+n_1) + C_p(p+p_1)} np = G, \quad \text{(Equation 6)}$$

where Σ_T is the electron trap density, C_n and C_p are the capture coefficients for electrons and holes, respectively, and $n_1 = p_1 = N_{cv}\exp(-E_{trap}/kT)$. Typically, $n >> n_1$ and $p >> p_1$ when traps act as recombination centers,[32,55] and therefore these constants can be neglected.

Now, if $n \approx p$ and assuming $C_n = C_p$ we get

$$R \approx R_{SRH} = C_p \Sigma_T p = G, \quad \text{(Equation 7)}$$

$$p = n = G/C_p\Sigma_T. \quad \text{(Equation 8)}$$

Following Equation 1, the V_{OC} is now approximated by

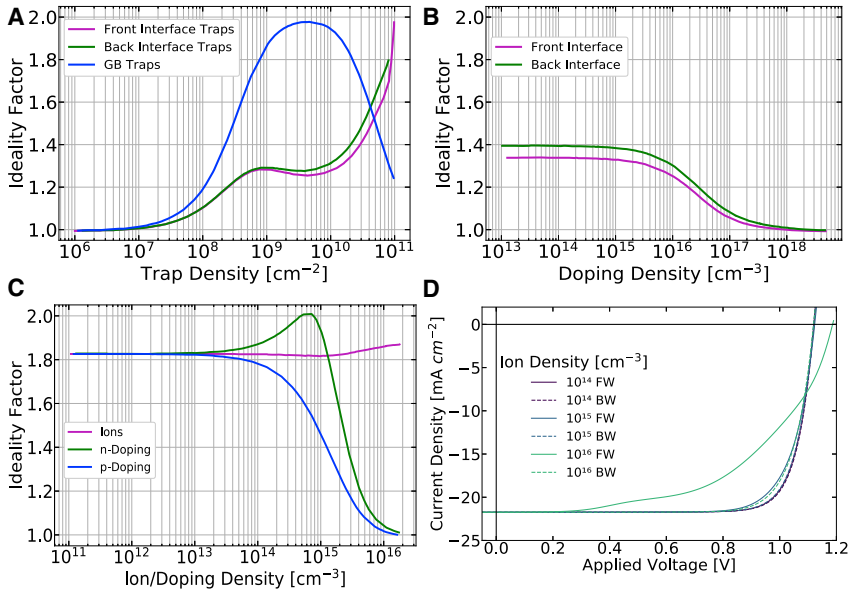


Figure 2. Simulated evolution of the ideality factor depending on physical parameters

(A) Varying trap density.

(B) Varying TL doping.

(C) Varying ion and doping density in the perovskite layer.

(D) Effect of the ions density on the degree of hysteresis.

Note that for (A), (B), and (C), the traps at the grain boundaries and front or back interfaces are considered separately.

The parameters used in the simulation can be found in Table S1.

$$qV_{OC} = E_{gap} - kT\ln\left(C_p\Sigma_T N_{cv}^2 \times \frac{1}{G^2}\right). \quad \text{(Equation 9)}$$

In this case, the light intensity dependence of the V_{OC} shows a slope of $2kT/q$ or a ideality factor of $\eta = 2$. Therefore, $\eta = 2$ represents dominant SRH recombination in the device (see Figure 2A).

Now, considering the special case of perovskite, it is unlikely to have evenly distributed trap states throughout the material bulk. In fact, studies have shown that defects tend to migrate out of the bulk and toward the grain boundaries (GBs) and interfaces, leaving low trap densities within the bulk.[56,57] For this reason, we only consider traps located at the GBs in this article.

In addition, as much as the previous derivation works well for traps evenly spread through the perovskite layer, one must be more careful when considering spatially localized traps such as traps located at GBs. In fact, upon high trap density at the GB, a depletion region is created because of the charge of the traps; hence, it somewhat pins the opposite charge carrier density to compensate for the space-charge region leading to $\eta < 2$ (see next section).

Dominant SRH recombination with one pinned charge carrier density

In this scenario, trap-assisted recombination is still dominant; however, one of the charge carriers is pinned to a finite value (i.e., it remains roughly constant upon changing light intensity). This can happen in PSCs for several reasons:

- High trap density at the GB, creating a depletion region (see above) (Figure 2A)

- Traps at the perovskite-TL interface with a doped TL where the charge carrier density is pinned to the doping level of the respective charge transport layer (Figure 2B)
- Traps at the perovskite-TL interface thin undoped TL and an ohmic contact with the respective electrode. The ohmic contact pins the charge carrier density as the latter induces many charges in the charge transport layer[58] (Figure 2A)
- Doping of the perovskite bulk

If we take the example of a perovskite-ETL interface, $n >> p$ and $n \approx N_D^+$, where N_D^+ is the doping level in ETL. Therefore, when SRH recombination is dominant at the perovskite-ETL interface, following Equation 6, at V_{OC}, the total recombination rate is

$$R \approx R_{SRH} = C_p \Sigma_T p = G, \quad \text{(Equation 10)}$$

$$p = G/C_p\Sigma_T \text{ and } n \approx N_D^+ \quad \text{(Equation 11)}$$

Now, following Equation 1, the V_{OC}, when SRH recombination dominates at the perovskite-ETL interface, is approximated by

$$qV_{OC} = E_{gap} - kT\ln\left(\frac{C_p\Sigma_T N_{cv}^2}{N_D^+} \times \frac{1}{G}\right) \quad \text{(Equation 12)}$$

Similarly, when SRH recombination dominates at the HTL-perovskite interface, the V_{OC} is approximated by

$$qV_{OC} = E_{gap} - kT\ln\left(\frac{C_n\Sigma_T N_{cv}^2}{N_A^-} \times \frac{1}{G}\right) \quad \text{(Equation 13)}$$

where N_A^- is the doping level in the HTL. The light intensity dependence of the V_{OC} now shows a slope of kT/q or an ideality factor of $\eta = 1$. Therefore, just as in the case of dominant band-to-band recombination, $\eta = 1$, which would make the prediction of the dominant recombination process based only on the ideality factor very inaccurate.

Figure 2A shows the evolution of the ideality factor depending on the trap density either at the grain boundary or at one of the interfaces (see Figure S1 for more details). We can note that as expected from the derivation for very low trap densities $\eta \rightarrow 1$, as band-to-band recombination dominates. However, we can see that the evolution of the ideality factor respective to the trap density is not monotonously increasing and depends on where the traps are located within the device. Thus, the ideality factor can definitely not be used as a quantitative measure of the number of traps in the systems. We can also see the different scenarios leading to a reduced ideality factor because of the pinning of one charge carrier, whether it is related to a large number of traps at the GB (Figure S1A, blue line), to a pinning because of thin TL and interfacial traps (Figure S1A, magenta and green line), or because of the doping of the TL with interfacial traps (Figure S1B) (see Figures S1 and S2).

Doping the perovskite bulk, whether it comes from the addition of a compound or because of the oxidation of the metal,[59–61] also leads to a different ideality factor, as seen in Figure 2C. In fact, upon sufficient doping, one of the charge carrier densities is pinned, and again, $\eta \rightarrow 1$. Henceforth in the article, we do not consider doped perovskite.

Another important feature of PSCs is the presence of moving ions—we also simulated devices with different ion densities (see Figures 2C and 2D), and realize that

contrary to doping, they do not affect the ideality factor significantly for stabilized JVs. In addition, we show that if a device shows little to no hysteresis for high voltage scan rate, then there is also no significant effect of ions on the stabilized performance and the ideality factor. This is well in line with the work by Tessler and Vaynzof,[62] who showed that when a JV characteristic does not show signs of hysteresis, the device model can still reproduce the same material and device parameters without including ions. Hence, we do not consider ions in the following simulations, as they would also increase significantly the computational time.

To summarize, we showed that the analysis of the ideality factor is not as straightforward as expected and that other factors need to be taken into consideration. Thus, despite giving valuable information, the ideality factor cannot be used on its own to conclude anything about the dominant recombination process but must be analyzed in correlation with other parameters.

Dataset for ML

ML algorithms have proven to be very efficient classifiers, and here, recognizing the dominant recombination process is nothing more than a classification problem. We are looking at three or four potential classes: (1) band-to-band recombination, (2) SRH recombination in the bulk, and (3) interfacial SRH at the front (3a) or back (3b) interface. To accurately train a ML algorithm, we need a dataset on which the ML can be trained and tested. To build a usable dataset, we need to define enough descriptive features and, of course, have properly labeled data. In other words, we need a large amount of data on PSCs with light intensity-dependent performance and an already-known dominant recombination process for each of the data points. This represents one of the great challenges in applying ML strategies to PSC research, as there is no readily available dataset or database reporting performance values of PSCs with the corresponding limiting process. Despite the huge amount of data published every year on PSCs, there is no real standard on what to characterize when a newly made device is reported, aside from the typical JV curve under 1 sun AM 1.5; even more problematic, most of the time, we focus on reporting what has been improved compared to a reference, which is indeed important, but it would also be insightful to investigate what is still limiting the device.

To overcome the issue of not having an experimental dataset to work with, we decided to use DD simulations[31,32,47] to generate the needed data. DD simulations have already proven their ability to capture the physics of PSCs by reproducing a wide variety of devices with different perovskite composition and structure.[31,32,63–65] Hence, we are confident that it can give us meaningful and reliable data to work with. To obtain a representative dataset, we randomly picked parameters within a reasonable range (see Table S2) based on commonly reported experimental values in the literature for the thickness and mobilities of all layers. For the recombination parameters that are more difficult to access with experiments, we used previously published simulation work using either DD[31,32,63–65] or first-principles[51–54] simulation to set meaningful boundaries. We also included parasitic, lumped series and shunt resistances to better represent realistic PSCs, as this effect can strongly influence the performance of real solar cells. However, it should be noted that the light intensities used here were purposely chosen close to 1 sun to limit the influence of the parasitic resistances. A high series resistance will mostly affect the performance at higher light intensities, while a low shunt resistance is detrimental at low light intensities. Due to computational constraints, we have used a lumped rather than a distributed series resistance.[66] This implies that our results apply to small, lab-scale devices. We also chose to fix the band gap of the

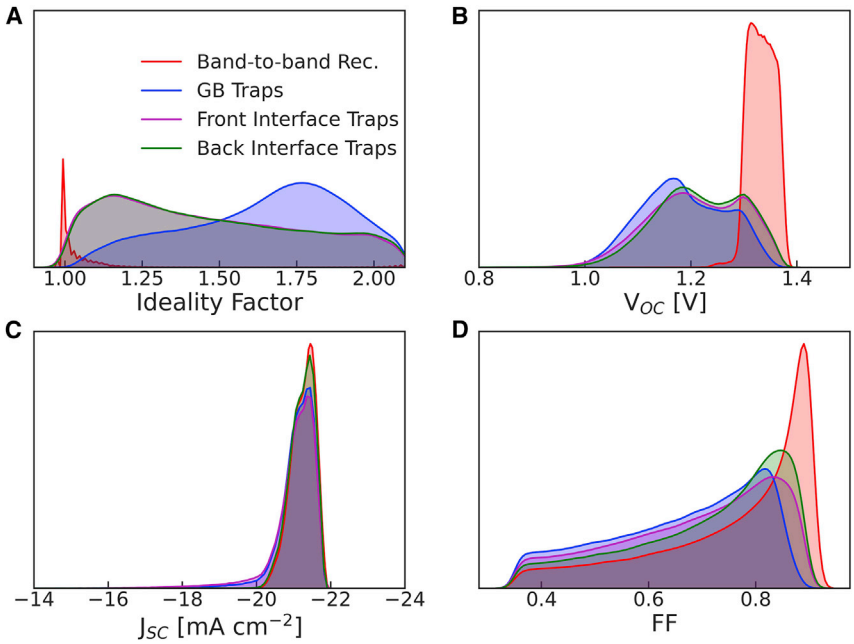


Figure 3. Density distribution of the main features for the different recombination types simulated

(A) The ideality factor.

(B) The open-circuit voltage.

(C) The short-circuit current.

(D) The fill factor. All given under 1 sun illumination.

Parameters used in the simulation can be found in the Supplemental experimental procedures.

perovskite to 1.6 eV (which corresponds to the band gap of MAPbI3), as we, for now, focus on a proof of concept; but obviously the analysis presented below would also hold for a different band gap by extending the simulation space for the said band gap.

Figure 3 presents the summary of the main performance parameters under 1 sun illumination and the ideality factor for >2 million simulated solar cells. The ideality factor was fitted from simulated values from 5 different light intensities, thus representing a total number of simulations of >10 million. One of the main advantages of using a dataset based on simulations is that each point can properly be labeled to their dominant recombination process (under 1 sun at maximum power point condition), as the contribution of each loss can be very easily disentangled from one another and the recombination fraction of each process can be calculated.

Without going further, we can already draw some conclusions from the density plot presented in Figure 3, the ideality factor not being sufficient to accurately distinguish the different recombination processes. In fact, as expected from the derivation made in the previous section $\eta \approx 1$ is not always a sign of dominant band-to-band recombination. However, if we consider ideality factor and V_{OC}, then we realize that $\eta \approx 1$ and $V_{OC} > 1.2V$ is very likely to be band-to-band recombination. As mentioned in the previous section, band-to-band limited devices exhibit impressively high V_{OC} with a loss lower than 0.4 V compared to the band gap; this is due to the very low values of the band-to-band recombination constant in perovskite. However, to be band-to-band recombination limited, the perovskite layer needs to be of very high quality with very low trap densities. This suggests that by adding

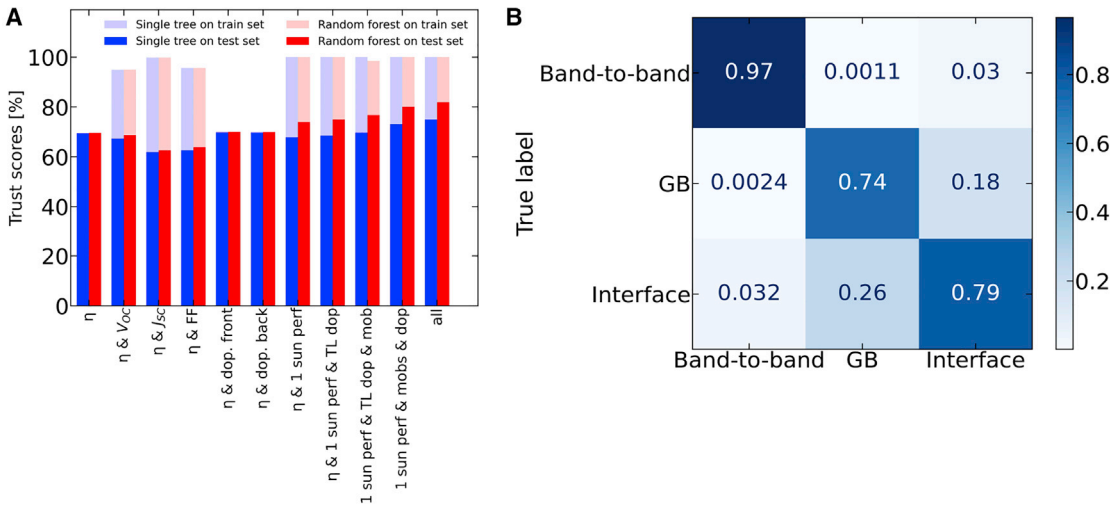


Figure 4. Machine learning performance

(A) Evolution of the trust score on training and test dataset depending on the features used to train single tree or random forest (RF). The full dataset consists of $>10^6$ simulated devices, 25% of which was used as a test set. The best performance (82%) is obtained for RF using 5 different light intensity performances, doping, and mobilities of all of the layers as features.

(B) Confusion matrix on the prediction for the best-performing RF trained showing how often a prediction is correct (diagonal), depending on the dominant recombination process.

other features on top of the ideality factor, we could improve the accuracy of the prediction of the dominant recombination process. Hence, moving to the ML technique could help us predict the dominant recombination process by adding/tuning features that can preferably be easily measured.

ML tree-based methods to identify the dominant recombination process for PSCs

As mentioned earlier, ML techniques are very effective for classification problems. Here, we chose to use two tree-based methods: single decision tree (ST) and random forest (RF) because they present the advantage of being more transparent than a typical neural network and that you can get more insight into what governs the decision by "drawing" the trained tree. All of the ML training was made using the scikit-learn[67] toolbox. We began by defining three classes for the dominant recombination process: (1) band-to-band, (2) GB traps, and (3) interfacial traps and trained our ML to distinguish between the different scenarios. To avoid any bias in our learning, the dataset was balanced between the three classes and 25% of the dataset was used as a test set. The common approach used in the literature is to use only the ideality factor as a feature, so we trained both ST and RF algorithms to use only η, giving an overall accuracy of ≈70% and ≈70% on the training and testing set, respectively (see Figure 4). The trust score—i.e., how much can you trust the prediction made by the ML—for each class can also be calculated (Figure S3), giving 85% for band-to-band recombination dominant, 63% for the GB traps, and 57% for interfacial traps. This means, for example, that if one inputs new numbers into the ML and the output is interfacial recombination, then there is a 57% chance that this is indeed the dominant recombination process.

We note that the ideality factor alone gives a decent prediction as to whether the system is band-to-band or trap limited, but it is not enough to distinguish whether the traps are located at the interface or within the bulk at the GBs. This is to be expected, as it is difficult to distinguish between the two trap-assisted recombination

processes. However, Figure 3A shows that there is still a difference in the density distribution for the solar cell 1 sun performance, suggesting that adding those features as descriptors should improve the prediction accuracy. In Figure 4A, we see that adding more features such as the 1 sun performance or the mobility and whether the TLs are doped help to improve the prediction of the ML algorithm. Note that in the simulation, the TL is considered doped if it satisfies the doping criteria introduced by Le Corre et al.[63]

The best result, with an overall accuracy >82%, was obtained with a RF classifier when using the performance of all of the simulated light intensities, the mobility of all of the layers, and whether the TLs are doped. This improvement is due to a better differentiation between interfacial and GB traps recombination.

The normalized confusion matrix on the prediction, shown in Figure 4B, gives more insight into the accuracy of the prediction by showing the fraction of the predictions that are correct on the diagonal. The off-diagonal elements give the fraction of wrongly assigned predictions. In our best case, it shows that when the ML output is either band-to-band, GB, or interfacial recombination, the results can be trusted in 97%, 74%, and 79% of the cases, respectively. However, the accuracy of the prediction is not perfect, which is mainly due to the fact that some interfacial dominant cases are mistaken as GB dominant and vice versa, and also the presence of mixed cases, in which both GB and interface make a comparable contribution to the overall recombination.

To validate our analysis, we chose experimental values from the literature that have the appropriate band gap (i.e., the right perovskite composition), in which the dominant loss is known and in which enough of the features are available. The first set of two devices is made of vacuum-processed MAPbI3 in a nip or pin structure,[32] and the main recombination process was determined by fitting DD simulations to light intensity-dependent current-voltage characteristics. The second set of data is made of triple-cation perovskite $CsPbI_{0.05}[(FAPbI_3)_{0.89}\ (MAPbI_3)_{0.11}]_{0.95}$ in a pin structure with changing interlayer or TLs, and the dominant recombination process was accessed by performing PL measurement and also DD simulation for devices 3 and 4.[33,63,68] For more details on the devices, see Table S3.

We used the ideality factor, 1 sun illumination performance, and doping of the TLs (Table 1) as features for the RF. We found good agreement between the prediction and the dominant recombination previously reported for these devices. Hence, despite the accuracy of the prediction not being perfect, it still predicts quite well the dominant process for experimental devices, proving that it could be used to analyze real experimental data for PSCs. To that aim, all of the data and the Python code used in this study are available on GitHub,[69,70] as well as the open-source DD simulation package SIMsalabim, used to simulate the PSCs.[31,32,47]

This result brings the hope that in the future, we could use an even larger dataset with experimental data to train ML methods and predict the dominant losses. This would be especially useful with the aim of more automated and high-throughput production of data, in which the analysis needs to be performed automatically and accurately or used as an online tool, where the trained models would be available and could be used as a characterization tool for new studies. We could even envision adding all kinds of other experimental results as features to improve the accuracy of the ML even more. In addition, this line of thinking could also be easily adapted for other kinds of photovoltaic technology, such as organic or quantum dots solar cells.

Table 1. Parameters used as features for the random forest algorithm to predict the dominant recombination of experimental perovskite solar cells

Device no.	Ref.	V_{OC} (V)	J_{SC} (mA cm^{-2})	*FF*	η	Front TL doped	Back TL doped	Experimental dominant	ML prediction: prediction recombination
1	[32]	1.08	−20.0	0.73	2	yes	yes	mixed GB/interface	GB
2	[32]	1.12	−20.2	0.81	1.55	yes	yes	mixed GB/interface	GB/interface
3	[33,63]	1.09	−21.65	0.779	1.42	no	no	interface	interface
4	[33]	1.17	−21.7	0.786	1.42	no	no	interface	interface

Find more information on the device structures in the Supplemental experimental procedures. FF, fill factor; GB, grain boundary; ML, machine learning; TL, transport layer.

Despite being widely reported in the literature, the analysis of the dominant recombination based on the ideality factor turns out to be less reliable than expected for PSCs, as it is also strongly affected by other factors. In fact, by performing a large-scale simulation, we have shown that the ideality factor does not only depend on the dominant recombination process but also on carrier pinning, either because of the contacts or doping. Hence, to draw any meaningful conclusion on the main loss process, the analysis needs to be complemented by other features.

We show using ML procedures that the prediction of the dominant recombination process can be improved by adding the performances under different light intensities, doping, and mobilities. A maximum prediction accuracy of 82% was obtained using a RF algorithm, which represents a 12% improvement compared to the use of the ideality factor only. The trained RF was also able to accurately predict the dominant recombination of experimental devices.

This study shows that performing a large-scale simulation and using ML tools could lead to a new type of automated data analysis that would be suitable for high-throughput experimentation.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, L. J.A. Koster (l.j.a.koster@rug.nl).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Original data have been deposited to Mendeley Data: https://doi.org/10.17632/xbzw29tjz4.1 (full dataset and Python code for the ML training). The codes to reproduce the main results of this article can be found at:

https://github.com/kostergroup/SIMsalabim (DD code)
https://github.com/kostergroup/Perovskite-Device-Doctor (Python code for the ML training)

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrp.2021.100346.

ACKNOWLEDGMENTS

This work was supported by a grant from STW/NWO (VIDI 13476). This is a publication by the FOM Focus Group "Next Generation Organic Photovoltaics," participating in the Dutch Institute for Fundamental Energy Research (DIFFER).

AUTHOR CONTRIBUTIONS

V.M.L.C. performed most of the DD simulations, wrote the ML codes, performed the data analysis, and wrote the manuscript. T.S.S. initiated the project and performed the preliminary simulations. M.K. wrote some of the codes used to perform the simulations. V.M.L.C. and L.J.A.K. conceived the project. L.J.A.K. supervised the project. All of the authors contributed to the revision of the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 1, 2020
Revised: November 25, 2020
Accepted: January 21, 2021
Published: February 11, 2021

REFERENCES

1. Haghighatlari, M., Vishwakarma, G., Altarawy, D., Subramanian, R., Kota, B.U., Sonpal, A., Setlur, S., and Hachmann, J. (2019). ChemML: A Machine Learning and Informatics Program Package for the Analysis, Mining, and Modeling of Chemical and Materials Data. https://chemrxiv.org/articles/preprint/ChemML_A_Machine_Learning_and_Informatics_Program_Package_for_the_Analysis_Mining_and_Modeling_of_Chemical_and_Materials_Data/8323271.

2. Hachmann, J., Afzal, M.A.F., Haghighatlari, M., and Pal, Y. (2018). Building and deploying a cyberinfrastructure for the data-driven design of chemical systems and the exploration of chemical space. Mol. Simul. *44*, 921–929.

3. Sanchez-Lengeling, B., Roch, L.M., Perea, J.D., Langner, S., Brabec, C.J., and Aspuru-Guzik, A. (2019). A Bayesian Approach to Predict Solubility Parameters. Adv. Theory Simul. *2*, 1800069.

4. Tabor, D.P., Roch, L.M., Saikin, S.K., Kreisbeck, C., Sheberla, D., Montoya, J.H., Dwaraknath, S., Aykol, M., Ortiz, C., Tribukait, H., et al. (2018). Accelerating the discovery of materials for clean energy in the era of smart automation. Nat. Rev. Mater. *3*, 5–20.

5. Nagasawa, S., Al-Naamani, E., and Saeki, A. (2018). Computer-Aided Screening of Conjugated Polymers for Organic Solar Cell: Classification by Random Forest. J. Phys. Chem. Lett. *9*, 2639–2646.

6. Cao, B., Adutwum, L.A., Oliynyk, A.O., Luber, E.J., Olsen, B.C., Mar, A., and Buriak, J.M. (2018). How To Optimize Materials and Devices via Design of Experiments and Machine Learning: Demonstration Using Organic Photovoltaics. ACS Nano *12*, 7434–7444.

7. Sun, S., Hartono, N.T., Ren, Z.D., Oviedo, F., Buscemi, A.M., Layurova, M., Chen, D.X., Ogunfunmi, T., Thapa, J., Ramasamy, S., et al. (2019). Accelerated Development of Perovskite-Inspired Materials via High-Throughput Synthesis and Machine-Learning Diagnosis. Joule *3*, 1437–1451.

8. Häse, F., Roch, L.M., and Aspuru-Guzik, A. (2019). Next-Generation Experimentation with Self-Driving Laboratories. Trends Chem. *1*, 282–291.

9. Zakutayev, A., Wunder, N., Schwarting, M., Perkins, J.D., White, R., Munch, K., Tumas, W., and Phillips, C. (2018). An open experimental database for exploring inorganic materials. Sci. Data *5*, 180053.

10. Wilbraham, L., Sprick, R.S., Jelfs, K.E., and Zwijnenburg, M.A. (2019). Mapping binary copolymer property space with neural networks. Chem. Sci. (Camb.) *10*, 4973–4984.

11. Xia, R., Brabec, C.J., Yip, H.-L., and Cao, Y. (2019). High-throughput optical screening for efficient semitransparent organic solar cells. Joule *3*, 2241–2254.

12. Xie, C., Tang, X., Berlinghof, M., Langner, S., Chen, S., Späth, A., Li, N., Fink, R.H., Unruh, T., and Brabec, C.J. (2018). Robot-Based High-Throughput Engineering of Alcoholic Polymer: Fullerene Nanoparticle Inks for an Eco-Friendly Processing of Organic Solar Cells. ACS Appl. Mater. Interfaces *10*, 23225–23234.

13. Chen, S., Hou, Y., Chen, H., Tang, X., Langner, S., Li, N., Stubhan, T., Levchuk, I., Gu, E., Osvet, A., and Brabec, C.J. (2018). Exploring the Stability of Novel Wide Bandgap Perovskites by a Robot Based High Throughput Approach. Adv. Energy Mater. *8*, 1701543.

14. Kojima, A., Teshima, K., Shirai, Y., and Miyasaka, T. (2009). Organometal halide perovskites as visible-light sensitizers for photovoltaic cells. J. Am. Chem. Soc. *131*, 6050–6051.

15. Stranks, S.D., Eperon, G.E., Grancini, G., Menelaou, C., Alcocer, M.J., Leijtens, T., Herz, L.M., Petrozza, A., and Snaith, H.J. (2013). Electron-hole diffusion lengths exceeding 1 micrometer in an organometal trihalide perovskite absorber. Science *342*, 341–344.

16. Edri, E., Kirmayer, S., Mukhopadhyay, S., Gartsman, K., Hodes, G., and Cahen, D. (2014). Elucidating the charge carrier separation and working mechanism of CH3NH3PbI(3-x)Cl(x) perovskite solar cells. Nat. Commun. *5*, 3461.

17. National Renewable Energy Laboratory (2019). Best research-cell efficiencies. https://www.nrel.gov/pv/assets/images/efficiency-chart.png.

18. Liu, M., Johnston, M.B., and Snaith, H.J. (2013). Efficient planar heterojunction perovskite solar cells by vapour deposition. Nature *501*, 395–398.

19. Ahn, N., Son, D.-Y., Jang, I.-H., Kang, S.M., Choi, M., and Park, N.-G. (2015). Highly Reproducible Perovskite Solar Cells with Average Efficiency of 18.3% and Best Efficiency of 19.7% Fabricated via Lewis Base Adduct of Lead(II) Iodide. J. Am. Chem. Soc. *137*, 8696–8699.

20. Correa-Baena, J.-P., Anaya, M., Lozano, G., Tress, W., Domanski, K., Saliba, M., Matsui, T., Jacobsson, T.J., Calvo, M.E., Abate, A., et al. (2016). Unbroken perovskite: Interplay of morphology, electro-optical properties, and ionic movement. Adv. Mater. *28*, 5031–5037.

21. Saliba, M., Matsui, T., Seo, J.-Y., Domanski, K., Correa-Baena, J.-P., Nazeeruddin, M.K., Zakeeruddin, S.M., Tress, W., Abate, A., Hagfeldt, A., and Grätzel, M. (2016). Cesium-containing triple cation perovskite solar cells: improved stability, reproducibility and high efficiency. Energy Environ. Sci. *9*, 1989–1997.

22. Nie, W., Tsai, H., Asadpour, R., Blancon, J.-C., Neukirch, A.J., Gupta, G., Crochet, J.J., Chhowalla, M., Tretiak, S., Alam, M.A., et al. (2015). Solar cells. High-efficiency solution-processed perovskite solar cells with millimeter-scale grains. Science *347*, 522–525.

23. Momblona, C., Gil-Escrig, L., Bandiello, E., Hutter, E.M., Sessolo, M., Lederer, K., Blochwitz-Nimoth, J., and Bolink, H.J. (2016). Efficient vacuum deposited pin and nip perovskite solar cells employing doped charge transport layers. Energy Environ. Sci. *9*, 3456–3463.

24. Shao, S., Abdu-Aguye, M., Qiu, L., Lai, L.-H., Liu, J., Adjokatse, S., Jahani, F., Kamminga, M.E., Gert, H., Palstra, T.T., et al. (2016). Elimination of the light soaking effect and performance enhancement in perovskite solar cells using a fullerene derivative. Energy Environ. Sci. *9*, 2444–2452.

25. Choi, J., Song, S., Hörantner, M.T., Snaith, H.J., and Park, T. (2016). Well-defined nanostructured, single-crystalline tio2 electron transport layer for efficient planar perovskite solar cells. ACS Nano *10*, 6029–6036.

26. Zhou, H., Chen, Q., Li, G., Luo, S., Song, T.B., Duan, H.-S., Hong, Z., You, J., Liu, Y., and Yang, Y. (2014). Photovoltaics. Interface engineering of highly efficient perovskite solar cells. Science *345*, 542–546.

27. Correa-Baena, J.-P., Tress, W., Domanski, K., Anaraki, E.H., Turren-Cruz, S.-H., Roose, B., Boix, P.P., Gratzel, M., Saliba, M., Abate, A., and Hagfeldt, A. (2017). Identifying and suppressing interfacial recombination to achieve high open-circuit voltage in perovskite solar cells. Energy Environ. Sci. *10*, 1207–1212.

28. Tan, H., Jain, A., Voznyy, O., Lan, X., García de Arquer, F.P., Fan, J.Z., Quintero-Bermudez, R., Yuan, M., Zhang, B., Zhao, Y., et al. (2017). Efficient and stable solution-processed planar perovskite solar cells via contact passivation. Science *355*, 722–726.

29. Johnston, M.B., and Herz, L.M. (2016). Hybrid perovskites for photovoltaics: Charge-carrier recombination, diffusion, and radiative efficiencies. Acc. Chem. Res. *49*, 146–154.

30. Tress, W., Marinova, N., Inganäs, O., Nazeeruddin, M.K., Zakeeruddin, S.M., and Graetzel, M. (2015). Predicting the Open-Circuit Voltage of CH3NH3PbI3 Perovskite Solar Cells Using Electroluminescence and Photovoltaic Quantum Efficiency Spectra: the Role of Radiative and Non-Radiative Recombination. Adv. Energy Mater. *5*, 1400812.

31. Sherkar, T.S., Momblona, C., Gil-Escrig, L., Bolink, H.J., and Koster, L.J.A. (2017a). Improving the performance of perovskite solar cells: Insights from a validated device model. Adv. Energy Mater. *7*, 1602432.

32. Sherkar, T.S., Momblona, C., Gil-Escrig, L., Ávila, J., Sessolo, M., Bolink, H.J., and Koster, L.J.A. (2017b). Recombination in Perovskite Solar Cells: Significance of Grain Boundaries, Interface Traps, and Defect Ions. ACS Energy Lett. *2*, 1214–1222.

33. Stolterfoht, M., Wolff, C.M., Márquez, J.A., Zhang, S., Hages, C.J., Rothhardt, D., Albrecht, S., Burn, P.L., Meredith, P., Unold, T., and Neher, D. (2018a). Visualization and suppression of interfacial recombination for high-efficiency large-area pin perovskite solar cells. Nat. Energy *3*, 847–854.

34. Stolterfoht, M., Caprioglio, P., Wolff, C.M., Márquez, J.A., Nordmann, J., Zhang, S., Rothhart, D., Hörmann, U., Redinger, A., Kegelmann, L., et al. (2018b). The perovskite/transport layer interfaces dominate non-radiative recombination in efficient perovskite solar cells. arXiv, 1810.01333. https://arxiv.org/abs/1810.01333.

35. Fang, R., Wu, S., Chen, W., Liu, Z., Zhang, S., Chen, R., Yue, Y., Deng, L., Cheng, Y.-B., Han, L., and Chen, W. (2018). [6,6]-Phenyl-C_{61}-Butyric Acid Methyl Ester/Cerium Oxide Bilayer Structure as Efficient and Stable Electron Transport Layer for Inverted Perovskite Solar Cells. ACS Nano *12*, 2403–2414.

36. Sidhik, S., Cerdan Pasarán, A., Esparza, D., López Luke, T., Carriles, R., and De la Rosa, E. (2018). Improving the Optoelectronic Properties of Mesoporous TiO_2 by Cobalt Doping for High-Performance Hysteresis-free Perovskite Solar Cells. ACS Appl. Mater. Interfaces *10*, 3571–3580.

37. Yang, G., Chen, C., Yao, F., Chen, Z., Zhang, Q., Zheng, X., Ma, J., Lei, H., Qin, P., Xiong, L., et al. (2018). Effective Carrier-Concentration Tuning of SnO_2 Quantum Dot Electron-Selective Layers for High-Performance Planar Perovskite Solar Cells. Adv. Mater. *30*, e1706023.

38. Wehrenfennig, C., Eperon, G.E., Johnston, M.B., Snaith, H.J., and Herz, L.M. (2014). High charge carrier mobilities and lifetimes in organolead trihalide perovskites. Adv. Mater. *26*, 1584–1589.

39. Duan, H.-S., Zhou, H., Chen, Q., Sun, P., Luo, S., Song, T.-B., Bob, B., and Yang, Y. (2015). The identification and characterization of defect states in hybrid organic-inorganic perovskite photovoltaics. Phys. Chem. Chem. Phys. *17*, 112–116.

40. Stranks, S.D., Burlakov, V.M., Leijtens, T., Ball, J.M., Goriely, A., and Snaith, H.J. (2014). Recombination kinetics in organic-inorganic perovskites: excitons, free charge, and subgap states. Phys. Rev. Appl. *2*, 034007.

41. Yamada, Y., Nakamura, T., Endo, M., Wakamiya, A., and Kanemitsu, Y. (2014). Photocarrier recombination dynamics in perovskite $CH_3NH_3PbI_3$ for solar cell applications. J. Am. Chem. Soc. *136*, 11610–11613.

42. D'Innocenzo, V., Srimath Kandada, A.R., De Bastiani, M., Gandini, M., and Petrozza, A. (2014). Tuning the light emission properties by band gap engineering in hybrid lead halide perovskite. J. Am. Chem. Soc. *136*, 17730–17733.

43. Leijtens, T., Eperon, G.E., Barker, A.J., Grancini, G., Zhang, W., Ball, J.M., Kandada, A.R.S., Snaith, H.J., and Petrozza, A. (2016). Carrier trapping and recombination: the role of defect physics in enhancing the open circuit voltage of metal halide perovskite solar cells. Energy Environ. Sci. *9*, 3472–3481.

44. deQuilettes, D.W., Zhang, W., Burlakov, V.M., Graham, D.J., Leijtens, T., Osherov, A., Bulović, V., Snaith, H.J., Ginger, D.S., and Stranks, S.D. (2016). Photo-induced halide redistribution in organic-inorganic perovskite films. Nat. Commun. *7*, 11683.

45. Koster, L.J.A., Mihailetchi, V.D., Ramaker, R., and Blom, P.W.M. (2005). Light intensity dependence of open-circuit voltage of polymer: fullerene solar cells. Appl. Phys. Lett. *86*, 123509.

46. Mandoc, M.M., Kooistra, F.B., Hummelen, J.C., De Boer, B., and Blom, P.W.M. (2007). Effect of traps on the performance of bulk heterojunction organic solar cells. Appl. Phys. Lett. *91*, 263505.

47. Sherkar, T.S., Le Corre, V.M., Koopmans, M., Wobben, F., and Koster, L.J.A. (2020). SIMsalabim GitHub repository, version 3.76. https://github.com/kostergroup/SIMsalabim.

48. Tress, W. (2017). Perovskite solar cells on the way to their radiative efficiency limit–insights into a success story of high open-circuit voltage and low recombination. Adv. Energy Mater. *7*, 1602358.

49. Wang, Z., Lin, Q., Wenger, B., Christoforo, M.G., Lin, Y.-H., Klug, M.T., Johnston, M.B., Herz, L.M., and Snaith, H.J. (2018). High irradiance performance of metal halide perovskites for concentrator photovoltaics. Nat. Energy *3*, 855–861.

50. Schiff, E.A. (2003). Low-mobility solar cells: a device physics primer with application to amorphous silicon. Sol. Energ. Mat. Sol. *78*, 567–595.

51. Davies, C.L., Filip, M.R., Patel, J.B., Crothers, T.W., Verdi, C., Wright, A.D., Milot, R.L., Giustino, F., Johnston, M.B., and Herz, L.M. (2018). Bimolecular recombination in methylammonium lead triiodide perovskite is an inverse absorption process. Nat. Commun. *9*, 293.

52. Ambrosio, F., Wiktor, J., De Angelis, F., and Pasquarello, A. (2018). Origin of low electron-hole recombination rate in metal halide perovskites. Energy Environ. Sci. *11*, 101–105.

53. Wolf, C., Cho, H., Kim, Y.-H., and Lee, T.-W. (2017). Polaronic Charge Carrier-Lattice Interactions in Lead Halide Perovskites. ChemSusChem *10*, 3705–3711.

54. Zhang, X., Shen, J.-X., Wang, W., and Van de Walle, C.G. (2018). First-Principles Analysis of Radiative Recombination in Lead-Halide Perovskites. ACS Energy Lett. *3*, 2329–2334.

55. Simmons, J., and Taylor, G. (1971). Nonequilibrium steady-state statistics and associated effects for insulators and semiconductors containing an arbitrary distribution of traps. Phys. Rev. B *4*, 502.

56. Azpiroz, J.M., Mosconi, E., Bisquert, J., and De Angelis, F. (2015). Defect migration in methylammonium lead iodide and its role in perovskite solar cell operation. Energy Environ. Sci. *8*, 2118–2127.

57. Buin, A., Comin, R., Xu, J., Ip, A.H., and Sargent, E.H. (2015). Halide-Dependent Electronic Structure of Organolead Perovskite Materials. Chem. Mater. *27*, 4405–4412.

58. Wetzelaer, G.A.H., and Blom, P.W.M. (2014). Ohmic current in organic metal-insulator-metal diodes revisited. Phys. Rev. B Condens. Matter Mater. Phys. *89*, 241201.

59. Zhou, Y., Chen, J., Bakr, O.M., and Sun, H.-T. (2018). Metal-Doped Lead Halide Perovskites: Synthesis, Properties, and Optoelectronic Applications. Chem. Mater. *30*, 6589–6613.

60. Shao, S., Liu, J., Portale, G., Fang, H.-H., Blake, G.R., ten Brink, G.H., Koster, L.J.A., and Loi, M.A. (2018). Highly Reproducible Sn-Based Hybrid Perovskite Solar Cells with 9% Efficiency. Adv. Energy Mater. *8*, 1702019.

61. Kahmann, S., Shao, S., and Loi, M.A. (2019). Cooling, Scattering, and Recombination–The Role of the Material Quality for the Physics of Tin Halide Perovskites. Adv. Funct. Mater. *29*, 1902963.

62. Tessler, N., and Vaynzof, Y. (2020). Insights from device modeling of perovskite solar cells. ACS Energy Lett. *5*, 1260–1270.

63. Le Corre, V.M., Stolterfoht, M., Perdigón Toro, L., Feuerstein, M., Wolff, C., Gil-Escrig, L., Bolink, H.J., Neher, D., and Koster, L.J.A. (2019). Charge transport layers limiting the efficiency of perovskite solar cells: how to optimize conductivity, doping, and thickness. ACS Appl. Energy Mater. *2*, 6280–6287.

64. Neukom, M.T., Schiller, A., Züfle, S., Knapp, E., Ávila, J., Pérez-Del-Rey, D., Dreessen, C., Zanoni, K.P.S., Sessolo, M., Bolink, H.J., and Ruhstaller, B. (2019). Consistent Device Simulation Model Describing Perovskite Solar Cells in Steady-State, Transient, and Frequency Domain. ACS Appl. Mater. Interfaces *11*, 23320–23328.

65. Calado, P., Telford, A.M., Bryant, D., Li, X., Nelson, J., O'Regan, B.C., and Barnes, P.R. (2016). Evidence for ion migration in hybrid perovskite solar cells with minimal hysteresis. Nat. Commun. *7*, 13831.

66. Seeland, M., and Hoppe, H. (2015). Comparison of distributed vs. lumped series resistance modeling of thin-film solar cells and modules: Influence on the geometry-dependent efficiency. Physica Status Solidi A Appl. Mater. Sci. *212*, 1991–2000.

67. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: machine learning in Python. J. Mach. Learn. Res. *12*, 2825–2830.

68. Caprioglio, P., Stolterfoht, M., Wolff, C.M., Unold, T., Rech, B., Albrecht, S., and Neher, D. (2019). On the Relation between the Open-Circuit Voltage and Quasi-Fermi Level Splitting in Efficient Perovskite Solar Cells. Adv. Energy Mater. *9*, 1901631.

69. Le Corre, V.M. (2020a). Perovskite-Device-Doctor GitHub repository, version 1. https://github.com/kostergroup/Perovskite-Device-Doctor.

70. Le Corre, V.M. (2020b). Dataset for: Identification of the Dominant Recombination Process for Perovskite Solar Cells Based on Machine Learning, V2. https://doi.org/10.17632/xbzw29tjz4.2.