Materials Advances

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Mater. Adv.*, 2025, 6, 598

Bayesian optimization and prediction of the durability of triple-halide perovskite thin films under light and heat stressors†

Deniz N. Cakan, [a] Eric Oberholtz,[a] Ken Kaushal,[ab] Sean P. Dunfield [a] and David P. Fenning *[ab]

Received 24th July 2024,
Accepted 20th November 2024

DOI: 10.1039/d4ma00747f

rsc.li/materials-advances

Perovskite thin films are leading candidates for the wide bandgap active layer in tandem solar cells; however, they are plagued by instability at elevated temperature, and under illumination. The large compositional design space of perovskites offers hope to engineer stability but is difficult to search effectively due to the generally low reproducibility in processing perovskite thin films. Here, we employ Bayesian optimization (BO) in conjunction with precision automation of perovskite thin film processing to improve the stability of perovskite films with a bandgap of ≈1.65 eV. The BO framework results in a 2.5× increase in the learning rate compared to traditional grid search. Additionally, we present a regression model that provides the first robust prediction of perovskite stability under light and heat based on readily-measured photoluminescence properties. This regression model achieves reasonable predictive power (coefficient of variation of the root mean square error = 27%) for the task of correlating fast and simple optical metrology with resource-intensive standardized light and heat tests that run for hundreds of hours. Feature importance ranking confirms Br content and photoluminescence emission stability under intense light as key indicators of durability.

1 Introduction

Halide perovskites have shown extraordinary promise in photovoltaic cells, achieving efficiencies of up to 26.1% in standalone devices and 34.6% in tandem devices with silicon.[1,2] These breakthroughs are primarily attributed to the high optoelectronic quality of solution-processed perovskite films[3] and advances in interface passivation, which mitigates issues such as high defect densities/ion diffusion.[4]

Despite these advancements, achieving long-term stability under both illumination and thermal stress continues to pose significant challenges. 1-sun illumination testing is essential for evaluating and demonstrating the durability of perovskite materials, as perovskites are known to undergo halide segregation under non-equilibrium conditions induced by light or heat.[5–8] Additionally, hybrid perovskites suffer from chemical degradation pathways. For example, the commonly used A-site cation methylammonium[9] and the passivation molecule phenethylammonium, used in forming 2D perovskites,[10,11] have both been shown to undergo deprotonation reactions that exhibit Arrhenius behavior.

Historical data show that few perovskite solar cells maintain operational stability beyond 5000 h under these conditions.[12,13] The urgency to improve perovskite durability has led the community to leverage accelerated aging tests such as ISOS-L-2 testing recommendations[5] (1-sun illumination and 65 °C/85 °C). Generally, the elevated temperature within the ISOS-L-2 testing provides some acceleration factor of perovskite degradation with respect to 25 °C testing, *e.g.* 40×,[14] but the acceleration will be dependent on the specific architecture of the cell and module and the bill of materials as well as process conditions.

Testing at much higher temperatures to achieve faster acceleration rates is generally not feasible as this would risk activating unrealistic degradation pathways, especially with the organic cations typically used in these perovskites.[15] Consequently, there is a practical limit to how much the aging factor can be increased, which in turn limits the learning rate of ISOS-L-2 or similar tests. The prolonged and resource-intensive feedback loop in the durability assessment of today's perovskite materials and perovskite solar cells motivates the integration of active learning and machine learning (ML) techniques to more efficiently sample and correlate early-term (*e.g.*, ≤1 h post-fabrication) tests with the outcomes of standardized tests such as ISOS-L-2.

[a] *Aiiso Yufeng Li Family Department of Chemical and Nano Engineering, University of California San Diego, La Jolla, CA 92093, USA. E-mail: dfenning@ucsd.edu*
[b] *Materials Science & Engineering Program, University of California San Diego, La Jolla, CA 92093, USA*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4ma00747f

Machine learning regression techniques, such as random forest regression (RFR) and Gaussian process regression (GPR), have demonstrated some limited ability to predict the durability of perovskite materials.[16,17] For example, Hartono *et al.* showed that an RFR prediction of the time-to-degradation of methylammonium lead iodide ($MAPbI_3$) films passivated with different capping layers could help predict a degradation onset over 700 min of ISOS-L-3 testing that includes humidity, light, and heat.[18]

Unfortunately, the time-dependence in the measurement of samples and their post-fabrication handling are often underestimated as variables in the process. Perovskite thin films are known to exhibit variation in optoelectronic performance following exposure to light,[19] humidity,[20] and oxygen,[21] as defect concentrations can shift or new phases can form.[4,22] This process-induced and path-dependent variance, often unrecorded, could diminish the capacity of data-driven models to discern underlying patterns in datasets with relatively small sample sizes. Additionally, perovskites are notably sensitive to process variables, such as the anti-solvent application timing[23] and rate.[24] The pioneering work toward machine learning of perovskite durability[16–18] has thus likely been limited in part by the high variance inherent in operator-driven fabrication and measurement processes, as noted by the authors themselves.[18] We can conclude that inconsistency from fabrication to characterization reduces experimental power.

On the other hand, there is significant opportunity for process automation in reducing variance in perovskites[25] – a practice that is becoming increasingly prevalent[26–32] due to its enhancement of experimental throughput and process precision, while benefiting from the low-temperature processing of perovskite thin films, in particular.

Here, we combine the use of perovskite automated spin coat assembly line[33] (PASCAL) and ML to investigate composition engineering for durable wide bandgap perovskite thin films. PASCAL automates the production and characterization of perovskite thin films, significantly reducing the variance in the optoelectronic properties of resulting films (with a coefficient of variation as low as 0.08% for photoluminescence peak energy).[33] This reduction in variance facilitates a systematic and effective exploration of perovskite composition space.

Specifically, we focused the search on the double-cation, triple-halide space ($FA_{0.78}Cs_{0.22}Pb(I_{0.8-x-y}Br_xCl_y)_3$ x = 0–0.20 and y = 0–0.10), chosen for its existing proofs of relative durability[34] and suitable band gaps ($\geq$1.6–1.7 eV) for two-terminal tandem applications with a silicon bottom cell. We investigate alloying the I-rich perovskite with Br and Cl to widen the bandgap, where Cl is also known to enhance film morphology,[35,36] durability,[37] and luminescence.[38] We deliberately avoid methylammonium (MA) precursors because of the relative instability at elevated temperatures of films containing remnant MA after annealing.[12,39]

We first benchmark the ability of readily available Bayesian optimization (BO) models[40] to explore and optimize over the compositional space relative to traditional grid-search techniques. The results indicate that BO can expedite the learning of durable compositions by 2.5× *via* more efficient sampling than traditional grid-search methods. Next, we used a random forest regression (RFR) model[41], a model which predicts outcomes *via* a series of decision tree regressors, to link film level optoelectronic characterization in PASCAL with lengthy 1-sun and 85 °C degradation tests (specifically, film level ISOS-L-2).[5,42] The produced RFR model achieves a modal R^2 correlation over 500 different samplings of the dataset of 67.8% and a coefficient of variation of the root mean square error (CVRMSE) of 27.5%. This result evidenced a mild but robust correlation and predictive power of perovskite composition and PL spectroscopy with the film color change measured in the durability test conducted under 1-sun and 85 °C conditions. This study highlights the successful application of simple ML models to predict the degradation test performance of perovskite thin films across diverse compositions, demonstrating the potential to merge high-precision experimental automation with ML to enhance the prediction of functional properties that are resource-intensive to determine, specifically the durability of halide perovskite thin films.

2 Experimental methods

2.1 Sample fabrication

Glass/(2-(3,6-dimethoxy-9*H*-carbazol-9-yl)ethyl)phosphonic acid (MeO-2PACz)/perovskite sub-cells were fabricated using a custom-built high-throughput experimentation platform, the perovskite automated spin coat assembly line (PASCAL).[33] Four 1.4 M endpoint solutions (Table S1, ESI†) for the double-cation triple-halide space ($FA_{0.78}Cs_{0.22}Pb(I_{0.8-x-y}Br_xCl_y)_3$ (x = 0–0.20, y = 0–0.10 and 1% excess PbI_2) were manually prepared and mixed using PASCAL to prepare interpolated compositions.[43] Thin films were deposited using PASCAL *via* spin coating inside a glovebox with a constant N_2 purge. Precisely, 40 μL of the absorber solution was dispensed 5 s before spinning, followed by spinning at 5000 rpm with 2000 rpm acceleration for 50 s. With 30 s remaining in the step, 200 μL of methyl acetate was dispensed at 2000 μL s^{-1} from 2 mm above the center of the spinning substrate. Once the spin was complete, the sample was immediately transferred to a hotplate at 100 °C for 30 min using PASCAL. Finally, the sample was moved to an aluminum-floored storage tray to cool to room temperature. Further details regarding PASCAL as a platform, the cleaning procedure of the substrates, and details of MeO-2PACz deposition are provided elsewhere.[33] Prototype cells from the initial grid search were fabricated using a glass/ITO/MeO-2PACz/perovskite/C60/BCP/Ag architecture. C60 (30 nm), BCP (5 nm), and Ag (80 nm) were deposited using a glovebox-attached physical vapor deposition system.

2.2 Scanning electron microscopy

SEM imaging (FEI) of select samples was performed using an in-lens detector with an accelerating current of 0.1 nA and a voltage of 5 kV.

2.3 X-Ray-diffraction

XRD (Rigaku) measurements of select samples were performed using an 8.04 keV Cu K_α focused beam light source with a K_β filter.

2.4 Photoluminescence

All PL measurements were conducted at least 3 min after the annealed sample was allowed to cool to room temperature on an aluminum tray (Fig. 1a). A 0.9 W 635 nm laser (ThorLabs) was used as the excitation source within PASCAL for a single-point initial PL spectra with approximately 0.1-sun's power density at the sample. The spectra were collected using a 200 μm slit grating spectrometer (StellarNet) with an 8 nm spectral resolution and were analyzed by fitting a single Gaussian to determine the PL peak intensity, peak emission energy, and full-width half-max. PL stability measurements were then taken using a 0.9 W 405 nm laser (ThorLabs) focused to approximately 4-sun's power density at the sample as an excitation source for a 60 s duration, and each spectrum over time was analyzed similarly to the initial PL spectral analysis. The initial and final measurements were used to calculate the relative PL intensity change and the PL peak shift. Photoluminescence quantum yield (PLQY) measurements were carried out on a home-built setup, and the details can be found elsewhere.[33]

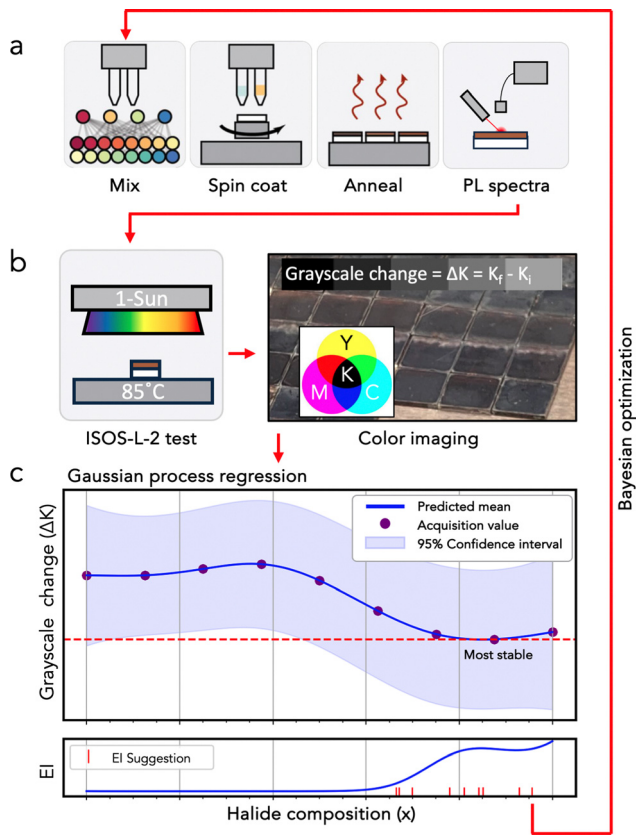

Fig. 1 (a) Schematic of automated sample fabrication *via* PASCAL. (b) ISOS-L-2 (1-sun and 85 °C) testing and images in the CMYK color scale. (c) Example schematic of a Gaussian process regression (above) fit on the change in grayscale (Δ*K*) in combination with the expected improvement (EI) acquisition function (below) together enabling Bayesian optimization (BO). Compositions suggested by EI are then fabricated by PASCAL to create an active learning loop.

2.5 Durability testing

After PL characterization within PASCAL, the sub-cells (glass/MeO-2PACz/perovskite) were then immediately manually transferred to a custom 1-sun illumination and 85 °C testing platform where an inert N_2 environment is maintained, mirroring the ISOS-L-2 recommendations (Fig. 1b). The testing platform is equipped with white light LEDs (ViparSpectra) calibrated to 1-Sun equivalent photon flux for the perovskite bandgap (Fig. S1, ESI†). The sealed chamber is equipped with a custom hot plate (temperature controlled to ±1%) and a quartz window and maintains a constant N_2 gas purge. The sub-cells were subjected to these heat and light conditions for 800 h.

2.6 Colormetrics

Following the colormetric method of Hartono *et al.*,[18] optical images (GoPro) of the perovskite films were taken in 20 min intervals as they degraded under 1-sun illumination and 85 °C test conditions. The RGB pixel values of each image were converted to a CMYK color scale. The *K* grayscale component of this CMYK color scale was then used as a proxy of film phase stability, where the desired black phase (high *K* value) transforms into a yellow PbI_2 and/or yellow hexagonal $FAPbI_3$ polymorph (low *K* value). The change in grayscale (Δ*K*) was calculated between the final and initial images. Use of this imaging technique was chosen for its compatibility with the batch characterization of films under light and heat.[18,44]

2.7 Active learning framework

An active learning framework based on BO[40] was tested as a means to search more efficiently for more durable compositions within the double-cation triple-halide perovskite composition space. The active learning approach used a Gaussian process regression (GPR) to make an initial fitting of a sample feature (composition) to a desired predicted feature (light and heat stability). An objective function, typically expected improvement[45–47] (EI), then operated on the GPR model predicted mean and uncertainty to suggest the composition at which the optimal durability under light and heat may be acquired.[48]

Halide perovskites at ≈1.65 eV are visible as dark brown films when in their photo-active phase and yellow when degraded to their photo-inactive phase. This visible color change was used as a proxy of the perovskite phase by analyzing the change in grayscale brightness (change in the *K* value from the CMYK color scale) of still images under the start and end of 1-sun illumination and 85 °C test conditions.

The change in grayscale was defined as a figure of merit (Δ*K*) to track phase stability under 1-sun illumination and 85 °C test conditions. A Gaussian process regression was fit over the triple-halide composition space to the figure of merit Δ*K* to produce a model of durability as a function of composition (Fig. 1c). The model utilized a single objective Gaussian process from BoTorch.[40] To implement active learning, expected improvement (EI) was used as the acquisition function and

the batch integrated negative posterior variance function was utilized to facilitate exploration (bottom subplot of Fig. 1c).[48–50] See the data availability statement for further details.

2.8 Random forest regression

Using the scikit-learn library in Python,[41,51] a random forest regression[52] was fit on the composition, PL intensity, PL peak energy, PL intensity change, and PL peak shift under illumination (details in Section 2.4) to predict ΔK. The model was 5-fold cross-validated using a 80 : 20 train-test split. The samples chosen to train and test the model were taken at random. To evaluate model robustness, the model was re-fitted 500 times using 5-fold cross-validation, generating a distribution of performance to inspect the variance in model results given the limited sample size. A SHAP table, which provides importance values to each feature in the RFR model, was then constructed for a model with modal R^2 and again for the highest R^2 model, where R^2 is predicted ΔK degradation *vs.* actual ΔK. See the data availability statement for further details.

2.9 Results and discussion

2.9.1 Comparing BO to grid search. To prepare a "ground truth" dataset using traditional search methods for comparison against the active learning approach, a fine grid mesh of 80 unique compositions was generated and the films were tested under 1-sun illumination and 85 °C conditions (Fig. 2a). A Gaussian process regression (GPR) was used to regress the change in grayscale (ΔK) on the composition, and the result is shown in Fig. 2b.

On inspection of the fitted ΔK values produced by the GPR model from the grid search dataset (Fig. 2b), three regions of relatively higher durability were observed at ≈5% Br, ≈11% Br, ≈17% Br, and ≤2% Cl (red circles in Fig. 2b). Compositions at ≈5% Cl and ≤5% Br were found to degrade relatively quickly under illumination and thermal stress, similar to results shown in the literature.[53] While the variation of durability across the full compositional space was relatively smooth, a high ΔK contrast was co-located with higher (≈5%) Cl loadings, likely due to the limited solubility of Cl in the I-rich Br-poor perovskite.[34,54] For the BO experiment, the search space was re-initialized by selecting 15 unique compositions considering a coarser grid search (black points in Fig. 2c). An expected improvement (EI) acquisition function was used with the GPR model of the space to suggest the next batch of 15 compositions to sample (red crosses in Fig. 2c). First to conduct BO, an EI acquisition function with a hyperparameter at 100% exploitation (BO-exploit) was used to ensure rapid honing in on a local optimum based on the information already gained. This batch of compositions suggested by EI was prepared using PASCAL, films were tested under 1-sun and 85 °C conditions, and the new measured ΔK values were then used to update the active learning model. The resultant GPR is shown in Fig. 2d. In testing points suggested in BO-exploit, the model identified the same 3 regions (≈5%, 11%, and 17% Br, at ≤2% Cl) of relatively heightened durability as those found by the model produced *via* grid search. In comparing the two GPR models

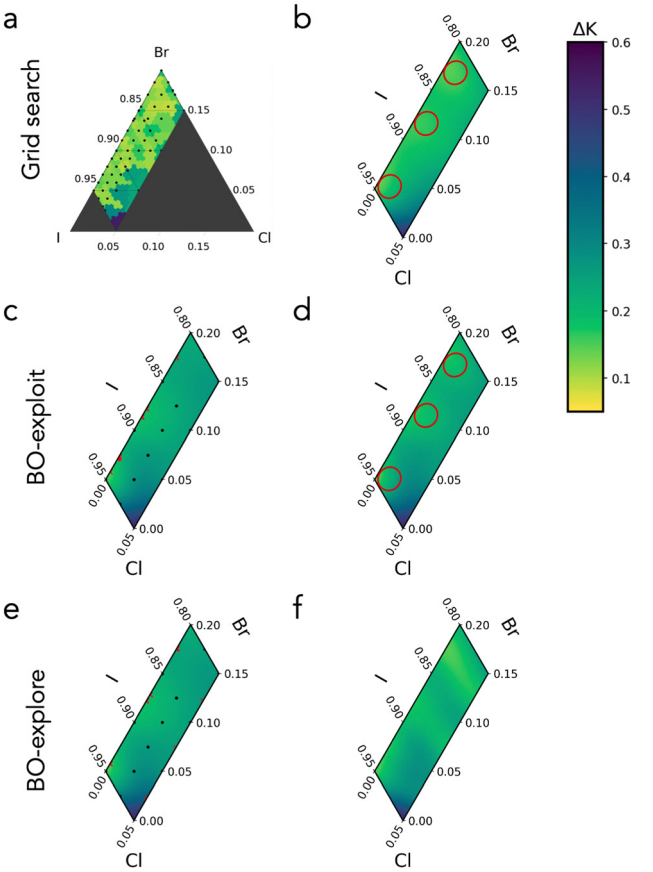

Fig. 2 (a) Grid search points in black with measured ΔK values depicted by a color bar. (b) GPR model produced by fitting grid search points. The color bar depicts predicted ΔK values. Red circles indicate regions of relatively high performance. (c) Initialization round composition points in black with BO-EI 100% exploit hyperparameter suggested points denoted by red crosses. The color bar depicts predicted ΔK values. (d) Updated model after testing the batch of points suggested by BO-EI with the 100% exploit hyperparameter. The color bar depicts predicted ΔK values. Red circles indicate regions of relatively high performance. (e) Initialization round points in black with BO-EI with 50% exploit:50% explore hyperparameter suggested points denoted by red crosses. The color bar depicts predicted ΔK values. (f) Updated model after testing the batch of points suggested by BO-EI with 50% exploit:50% explore hyperparameters. The color bar depicts predicted ΔK values.

produced from grid search and BO-exploit (Fig. S2, ESI†), there is minimal difference observed in these areas of relatively higher durability (≤5% absolute error relative to the full changes in grayscale). Thus, we conclude that these regions of relatively enhanced durability can be found within a fraction (60% less) of the grid search sample budget.

From these 3 regions (red circles in Fig. 2b and composition in Table 1), the most durable compositions from the GPR model were fabricated as glass/MeO-2PACz/perovskite subcells, followed by the basic characterization of the crystallographic phase *via* XRD, film morphology *via* SEM, and quasi-Fermi level splitting *via* photoluminescence quantum yield. Details of the basic characterization techniques are in the Experimental methods section.

XRD revealed the perovskite to be in the cubic phase for all three compositions with a slightly smaller lattice with

increasing Br, revealed by a shift in all peaks to a higher scattering angle (Fig. S3a, ESI†). SEM images of all 3 compositions revealed apparent grain domains on the order of 500 nm, with no obvious distinctions between compositions (Fig. S3a–c, ESI†).

In Cl-free films, unpassivated PLQY revealed a decrease from about 0.08% to 0.03% with increasing Br fractions, equivalent to an increase in the implied voltage deficit from ≈110 mV to ≈160 mV to Fig. S4, ESI†). This implied voltage deficit was substantially decreased in the Cl-containing film with even greater Br ($Br_{0.174}$, Fig. S5, ESI†). This decrease in the implied voltage deficit is perhaps due to the improved passivation of non-radiative defects by Cl in the film.

Prototype cells from the initial grid compositions were fabricated using a p–i–n architecture (glass/ITO/MeO-2PACz/perovskite/C60/BCP/Ag architecture). Fig. S6a (ESI†) shows the device results, and an image of the devices under test in the automated JV testing apparatus is shown in Fig. S6b (ESI†). The measured champion efficiency was an 18.5% PCE (Fig. S6c, ESI†), while the median PCE of all compositions was around 16.2% with an interquartile range of about 2% absolute. The full distribution of the PCE is shown in Fig. S6d (ESI†). The ≥18% PCE achieved is comparable to the state-of-the-art for wider bandgap perovskite devices when no targeted interface passivation is used.[55] Pearson's correlations between half-cell PL metrics and finalized devices show limited correlations between the PL metrics and device performance (Fig. S7, ESI†), indicating that the performance of the pristine devices is not limited by the optoelectronic quality of the absorber.

2.9.2 Exploitation *vs.* exploration: was information about the broader composition space lost when using the active learning focused on exploitation? Within Bayesian optimization, the exploitation:exploration hyperparameter in the acquisition function tunes the balance between honing in on a local optimum based on the information already gained (exploitation) *vs.* searching out new information toward a global optimum (exploration). Typically, experiments have a budget constraint that dictates the number of samples that can reasonably be generated and tested. Here, fifteen samples per batch were allocated as a budget for iterative model development.

The effect of the exploitation:exploration hyperparameter on the model performance was investigated by increasing exploration from 0% to 50%. In doing so, the model recommended mid-points between previously tested compositions. Interestingly, in testing these more exploratory compositions (red crosses in Fig. 2e) under 1-sun illumination and 85 °C test conditions, several compositions with a halide fraction of 5% Cl showed relatively heightened durability. Thus, this model (BO with exploration) uniquely predicts that some Cl may be alloyed while maintaining film durability (Fig. 2f), which is in relative agreement with the GPR model from the 80 sample grid (Fig. 2b). Limited Cl solubility has also been experimentally confirmed by Xu *et al.* in triple-cation (Cs, FA, and MA) and triple-halide (I, Br, and Cl) compositions.[56] Furthermore, the exploration parameter effectively prevented maxima trapping at low Cl loadings that occurred when the exploitation parameter was set to 100%. Thus, we conclude that maintaining some exploration, even in relatively well understood composition spaces like the triple-halide perovskite, enables a more robust compositional search that avoids getting trapped at local minima.

Table 1 Compositions of predicted most durable films, $FA_{0.78}Cs_{0.22}Pb(I_xBr_yCl_z)_3$

I_x	Br_y	Cl_z
0.95	0.05	0
0.884	0.116	0
0.823	0.174	0.003

Overall, the GPR produced in each of these BO campaigns and grid search of the compositional space provided insight into the possible presence of islands of durability in the perovskite composition space. Interestingly, all results indicated 3 regions of relatively increased durability at ≈5% Br, ≈11% Br, and ≈17% Br at predominately Cl-poor fractions. The most durable composition was identified as $FA_{0.78}Cs_{0.22}Pb(I_{0.95}Br_{0.05})_3$. Conversely, all investigations of the double-cation, triple-halide space revealed that Cl largely hindered durability in I-rich and Br-poor alloys, with the least durable composition of $FA_{0.78}Cs_{0.22}Pb(I_{0.95}Cl_{0.05})_3$. This may be due to the large mismatch in the ionic radius between I and Cl, which leads to microscopic phase heterogeneity due to limited solubility and subsequent degradation under illumination.[34,57] Additionally, the poor durability in films with Cl may be exacerbated due to the lack of volatile MACl-containing precursors where the MA deprotonates, enabling MA^0 and HCl to leave the film.[58]

2.9.3 Predicting perovskite degradation after light and heat accelerated stress tests. Toward increasing the learning rate of durability testing, a random-forest-regression (RFR) model was constructed in an attempt to correlate perovskite composition, fast optical characterization, and ISOS-L-2 film level durability. An RFR is chosen for its superior performance in higher dimensionality over GPRs. GPRs also assume that the response surface is smooth, which may not always be the case in perovskites, such as at phase boundaries like those arising due to the limited solubility of Cl in I-rich perovskites.[56]

However, RFRs built on a small sample size (≤200) typically suffer from high variance in model performance due to the nature of initializing train and test splits. To reduce this variance, cross-validation was utilized, which uses pre-allocation of randomly sampled parts of the dataset for training and a separate subset for testing. However, even with a typical 5-fold cross-validation, at low N, RFR is still quite susceptible to high variance. Thus, to better understand the distribution of possible model outcomes, a 5-fold cross-validation RFR model was built from the 80 unique compositions and redrawn 500 times to produce a distribution.

The RFR models were constructed using normalized features representing composition data, PL spectra data, and ΔK data of the grid search samples (Fig. 3a, with raw feature values in Fig. S8a–e and the half-cell reverse scan PCE in Fig. S8f, ESI†). All produced models are available online at github.com/fenning-research-group/BODP4T.

A modal model achieved a predicted ΔK *vs.* observed R^2 of 67.8%, an example of which is plotted in Fig. 3a. As can be

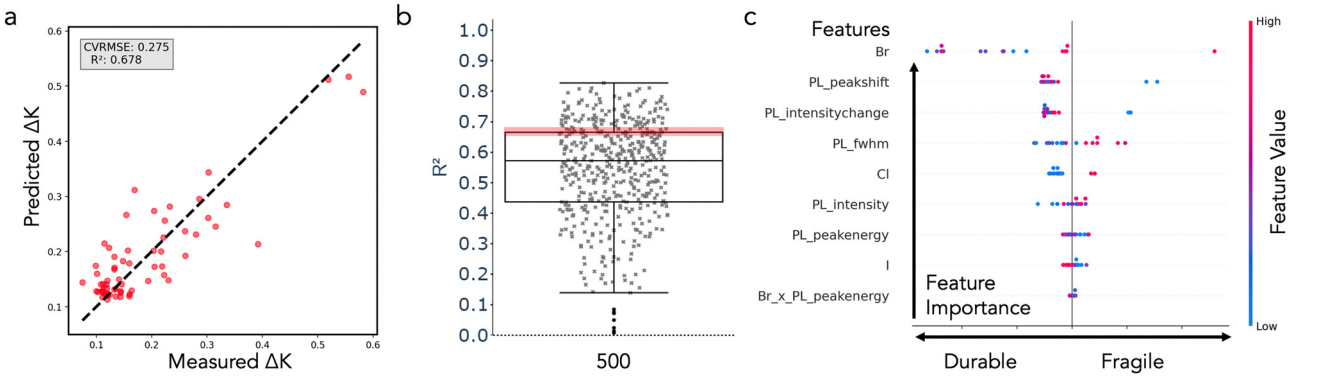

Fig. 3 (a) Parity plot of an RFR model yielding a modal R^2 goodness-of-fit metric. (b) Box plot of the 500 iterations of the 5-fold cross-validated model. Mode shaded in red. (c) SHAP table for the mode R^2 model produced in this study.

clearly seen, a modal model provides some measure of explanatory power. The best model had an R^2 of 80.8% (Fig. S10a, ESI†) with a 22.3% CVRMSE, where the RMSE is normalized by the mean value of the actual (ΔK). The 22.3% CVRMSE indicates reasonable power in the prediction of yellowing during degradation. A box plot of the R^2 goodness-of-fit metric for all 500 iterations of the model prediction of ΔK is shown in Fig. 3b. The data exhibit a single mode, with a characteristic R^2 of 67.8% and the exemplary modal model of Fig. 3a had a coefficient of variation of the RMSE of 27.5%. Of all 500 models, the modal coefficient of variation of the RMSE was 26.6% (Fig. S9, ESI†).

2.9.4 Features that predict perovskite film durability. To examine the importance of the chemical and material property features underlying the model prediction, a SHAP table[59] was constructed for the RFR model with R^2 modal performance (Fig. 3c). In this SHAP table, Br content was assigned as having the greatest impact on the model's fit and was calculated to be negatively correlated with the predicted ΔK. However, since there is no clear division of the Br fraction (feature value) with the model ΔK prediction (durable *vs.* fragile), a direct correlation is not suggested. This assignment agrees with the 3 distinct Br regions of relatively heightened durability found in the prior GPR models. The high impact of Br on the model is explained by the well-characterized sensitivity of Br in perovskites. For example, too high a Br content ($\geq$20%) has been shown to lead to halide segregation and degradation under illumination, bias, and thermal stress.[6–8]

Interestingly, the PL peak shift and PL intensity change under photo-exposure were the next highest importance features to the model's prediction of durability. Intuitively, a stable PL emission correlates with an optoelectronically invariant film under stressors, which should directly relate to durability. Similarly, the PL FWHM was calculated to negatively correlate with durability, which makes sense as perovskites not properly mixed at the A- and X- sites have been shown to phase separate into photo-inactive phases.[6,55,60] The Cl fraction was ranked just below the PL FWHM with a clear division in the SHAP table, indicating that the model predicts a more durable composition with a lower Cl fraction. This trend with the Cl fraction generally agrees with the prior GPR models.

Notably, the most predictive model from the 500 iterations, with an R^2 of 80.8% and a CVRMSE of 22.3% (Fig. S10a, ESI†), shared the same top 5 ranked features in its corresponding SHAP table (Fig. S11b, ESI†).

The SHAP table shows that the Br fraction and material properties as measured by PL (intensity, FWHM, and peak emission energy stability over high-intensity uniform light for 60 s) significantly influence model predictions. While few studies[56] have linked high-intensity PL to durability under light and heat, this work introduces the first robust and quantitative evaluation of using rapid optical characterization to predict 1-sun illumination and 85 °C film durability across multiple perovskite compositions, establishes a correlation, and provides a predictive model. While the SHAP table indicates the PL peak shift and intensity change as being impactful to the model's prediction, it does not explain a direct correlation. However, the table does show more instances of PL blueshift and increasing PL intensity under illumination as being predictive of a durable composition. Moreover, the SHAP table shows a clear relationship between the PL FWHM, Cl fraction, and durability within the model.

RFR models are inherently different from linear regression models in that they do not require regression on independent features. The decision tree-based structure of RFR allows it to capture complex, non-linear relationships and interactions between features, even without explicitly defined interaction terms.

To elucidate the direct correlations between individual input variables and ΔK, the full Pearson correlation matrix is provided in Fig. S11 (ESI†). There are negligible direct correlations between the PL peak shift and ΔK, indicating complex interactions with other features rather than a dominant standalone effect. For example, there is a mild positive correlation between the I and PL peak shift (r = 0.54) and a modest negative correlation between the Br and PL peak shift (r = −0.48). Additionally, there is a weak positive (negative) correlation between the I (Br) and PL intensity.

These correlations indicate that while there may be interactions between these variables, their predictive value within the RFR model arises from complex, non-linear relationships rather than direct, linear correlations. Additionally, the RFR

model's ability to handle features with some degree of collinearity supports the inclusion of all PL metrics.

3 Outlook and future work

Learning curves of the RFR trained on increasing fractions of the data show that the model performance has not yet saturated as the number of training examples was increased in full (Fig. S12, ESI†). While the training score is relatively high, there are positive trend lines of R^2 and the mean-squared error (MSE) with increasing data fraction. This trend with an increasing training set size suggests that the model could be further improved with an increase in the overall training data size. Similarly in the cross-validation scores, there is a monotonic improvement that appears yet to have reached an asymptote. Together these trend lines indicate that the model may see a modest improvement with a larger dataset. Additionally, future exploration of other regression models such as ridge regressions,[61] physics-informed Bayesian models,[16,62,63] or fully Bayesian neural networks[64] may further guide better prediction of perovskite film degradation.

While this investigation focuses on the advancement of the durability of glass/MeO-2PACz/perovskite sub-cells, the stability of the perovskite is a necessary condition for the stability of optoelectronic devices built around perovskite absorbers. The transferability of models built on sub-cell characterization to completed devices will be investigated in future work.

3.1 Conclusion

In this work, reproducible, automated perovskite thin film deposition is integrated with machine learning (ML) to enable the successful prediction of perovskite durability under light and elevated temperatures across the triple-halide composition space. The Bayesian optimization of durability within the triple-halide composition space effectively identified compositions with improved durability with ≈60% less samples tested than in a traditional grid search. This represents a roughly 2.5× increase in the rate of exploration within the high-dimensional halide space for the optimization of elevated light and heat durability. Overall, three islands of distinct Br loadings emerged within the triple-halide search space that showed enhanced durability, where $x_{Br} \approx 5\%$, 11%, and 17%. Generally, compositions had poor durability when the Cl-loading fraction was ≥2% of the halide fraction. Further investigation of compositions in or near these "durable" Br levels is warranted.

From the rich characterization and durability dataset, a random forest regression (RFR) model was built that predicted the durability of films fabricated under 1-sun illumination and 85 °C conditions with moderate predictive power, leveraging inputs on the composition and PL emission variation collected over 60 s from the automated characterization within PASCAL. This model provides a first step toward faster and less resource-intensive alternatives to accelerated testing using standardized 1-sun and 85 °C test conditions. SHAP feature importance ranking from the RFR revealed that changes to the PL spectra from the rapid characterization of newly-synthesized films are critical features in the durability prediction. Through these models, the study provides a framework toward accelerating the durability screening of novel perovskite compositions. Together, these results highlight the potential of combining precision automation with ML to improve the durability of halide perovskite thin films.

Author contributions

D. N. C. and D. P. F. conceived of the BO framework for the investigation. D. N. C. curated all data, conducted all analysis, built all visualizations, and wrote the original draft. E. O. assisted with PASCAL operation. K. K. and S. P. D. constructed the ISOS-L-2 degradation testing platform. D. P. F. led the research team, acquired funding, and supervised the research. All authors contributed to the review and editing of the final manuscript.

Data availability

The data and analysis code can be found at **https://github.com/fenning-research-group/BODP4T/tree/master/manuscript_data_availability**.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We would like to thank Rishi E. Kumar for his contributions to building PASCAL and early efforts with prototyping software for a Bayesian optimization task and Connor J. Dolan for his support with scanning electron microscopy. The compositional search and solar cell fabrication in this work was supported by the California Energy Commission, EPC-19-004. This material is based upon work supported in part by the U.S. Department of Energy's Solar Energy Technologies Office (SETO) Award Number DE-EE0010503.

Notes and references

1 *Best research-cell efficiency chart*, 2024, **https://www.nrel.gov/pv/cell-efficiency.html**.
2 L. LONGi Green Energy Technology Co., *34.6%! Record-breaker LONGi Once Again Sets a New World Efficiency for Silicon-perovskite Tandem Solar Cells*, 2024, **https://www.longi.com/en/news/2024-snec-silicon-perovskite-tandem-solar-cells-new-world-efficiency**, Accessed: 2024-06-20.
3 Y. Yuan, G. Yan, C. Dreessen, T. Rudolph, M. Hülsbeck, B. Klingebiel, J. Ye, U. Rau and T. Kirchartz, *Nat. Mater.*, 2024, **23**, 391–397.

4 S. P. Dunfield, L. Bliss, F. Zhang, J. M. Luther, K. Zhu, M. F. A. M. Hest, M. O. Reese and J. J. Berry, *Adv. Energy Mater.*, 2020, **10**, 1904054.
5 M. V. Khenkin, E. A. Katz, A. Abate, G. Bardizza, J. J. Berry, C. Brabec, F. Brunetti, V. Bulović, Q. Burlingame, A. Di Carlo, R. Cheacharoen, Y.-B. Cheng, A. Colsmann, S. Cros, K. Domanski, M. Dusza, C. J. Fell, S. R. Forrest, Y. Galagan, D. Di Girolamo, M. Grätzel, A. Hagfeldt, E. von Hauff, H. Hoppe, J. Kettle, H. Köbler, M. S. Leite, S. Liu, Y.-L. Loo, J. M. Luther, C.-Q. Ma, M. Madsen, M. Manceau, M. Matheron, M. McGehee, R. Meitzner, M. K. Nazeeruddin, A. F. Nogueira, Ç. Odabaşı, A. Osherov, N.-G. Park, M. O. Reese, F. De Rossi, M. Saliba, U. S. Schubert, H. J. Snaith, S. D. Stranks, W. Tress, P. A. Troshin, V. Turkovic, S. Veenstra, I. Visoly-Fisher, A. Walsh, T. Watson, H. Xie, R. Yldrm, S. M. Zakeeruddin, K. Zhu and M. Lira-Cantu, *Nat. Energy*, 2020, **5**, 35–49.
6 E. T. Hoke and M. D. McGehee, *Chem. Sci.*, 2015, **6**, 613–617.
7 R. E. Beal, N. Z. Hagström, J. Barrier, A. Gold-Parker, R. Prasanna, K. A. Bush, D. Passarello, L. T. Schelhas, K. Brüning, C. J. Tassone, H.-G. Steinrück, M. D. McGehee, M. F. Toney and A. F. Nogueira, *Matter*, 2020, **2**, 207–219.
8 Z. Xu, R. A. Kerner, S. P. Harvey, K. Zhu, J. J. Berry and B. P. Rand, *ACS Energy Lett.*, 2023, **8**, 513–520.
9 X. Wang, Y. Fan, L. Wang, C. Chen, Z. Li, R. Liu, H. Meng, Z. Shao, X. Du, H. Zhang, G. Cui and S. Pang, *Chem*, 2020, **6**, 1369–1378.
10 M. Wang, Z. Shi, C. Fei, Z. J. D. Deng, G. Yang, S. P. Dunfield, D. P. Fenning and J. Huang, *Nat. Energy*, 2023, **8**, 1229–1239.
11 M. A. Uddin, P. J. S. Rana, Z. Ni, G. Yang, M. Li, M. Wang, H. Gu, H. Zhang, B. D. Dou and J. Huang, *Nat. Commun.*, 2024, **15**, 1355.
12 H. Zhu, S. Teale, M. N. Lintangpradipto, S. Mahesh, B. Chen, M. D. McGehee, E. H. Sargent and O. M. Bakr, *Nat. Rev. Mater.*, 2023, **8**, 569–586.
13 C. Fei, A. Kuvayskaya, X. Shi, M. Wang, Z. Shi, H. Jiao, T. J. Silverman, M. Owen-Bellini, Y. Dong, Y. Xian, R. Scheidt, X. Wang, G. Yang, H. Gu, N. Li, C. J. Dolan, Z. J. D. Deng, D. N. Cakan, D. P. Fenning, Y. Yan, M. C. Beard, L. T. Schelhas, A. Sellinger and J. Huang, *Science*, 2024, **384**, 1126–1134.
14 Q. Jiang, R. Tirawat, R. A. Kerner, E. A. Gaulding, Y. Xian, X. Wang, J. M. Newkirk, Y. Yan, J. J. Berry and K. Zhu, *Nature*, 2023, **623**, 313–318.
15 J. Ruellou, M. Courty and F. Sauvage, *Adv. Funct. Mater.*, 2023, **33**, 2300811.
16 S. Sun, A. Tiihonen, F. Oviedo, Z. Liu, J. Thapa, Y. Zhao, N. T. P. Hartono, A. Goyal, T. Heumueller, C. Batali, A. Encinas, J. J. Yoo, R. Li, Z. Ren, I. M. Peters, C. J. Brabec, M. G. Bawendi, V. Stevanovic, J. Fisher and T. Buonassisi, *Matter*, 2021, **4**, 1305–1322.
17 W. A. Dunlap-Shohl, Y. Meng, P. P. Sunkari, D. A. C. Beck, M. Meilă and H. W. Hillhouse, *J. Mater. Chem. A*, 2024, 9730–9746.
18 N. T. P. Hartono, J. Thapa, A. Tiihonen, F. Oviedo, C. Batali, J. J. Yoo, Z. Liu, R. Li, D. F. Marrón, M. G. Bawendi, T. Buonassisi and S. Sun, *Nat. Commun.*, 2020, **11**, 4172.
19 Y. Li, X. Xu, C. Wang, B. Ecker, J. Yang, J. Huang and Y. Gao, *J. Phys. Chem. C*, 2017, **121**, 3904–3910.
20 Q. Wang, B. Chen, Y. Liu, Y. Deng, Y. Bai, Q. Dong and J. Huang, *Energy Environ. Sci.*, 2017, **10**, 516–522.
21 Q. Sun, P. Fassl, D. Becker-Koch, A. Bausch, B. Rivkin, S. Bai, P. E. Hopkinson, H. J. Snaith and Y. Vaynzof, *Adv. Energy Mater.*, 2017, **7**, 1700977.
22 K. P. Goetz, A. D. Taylor, F. Paulus and Y. Vaynzof, *Adv. Funct. Mater.*, 2020, **30**, 1910004.
23 K. Wang, M.-C. Tang, H. X. Dang, R. Munir, D. Barrit, M. De Bastiani, E. Aydin, D.-M. Smilgies, S. De Wolf and A. Amassian, *Adv. Mater.*, 2019, **31**, 1808357.
24 A. D. Taylor, Q. Sun, K. P. Goetz, Q. An, T. Schramm, Y. Hofstetter, M. Litterst, F. Paulus and Y. Vaynzof, *Nat. Commun.*, 2021, **12**, 1878.
25 M. Kumar, V. Pawar, P. K. Jha, P. A. Jha and P. Singh, *J. Solid State Chem.*, 2022, **308**, 122893.
26 B. P. MacLeod, F. G. L. Parlane, T. D. Morrissey, F. Häse, L. M. Roch, K. E. Dettelbach, R. Moreira, L. P. E. Yunker, M. B. Rooney, J. R. Deeth, V. Lai, G. J. Ng, H. Situ, R. H. Zhang, M. S. Elliott, T. H. Haley, D. J. Dvorak, A. Aspuru-Guzik, J. E. Hein and C. P. Berlinguette, *Sci. Adv.*, 2020, **6**, eaaz8867.
27 J. Wagner, C. G. Berger, X. Du, T. Stubhan, J. A. Hauch and C. J. Brabec, *J. Mater. Sci.*, 2021, **56**, 16422–16446.
28 J. Zhang, B. Liu, Z. Liu, J. Wu, S. Arnold, H. Shi, T. Osterrieder, J. A. Hauch, Z. Wu, J. Luo, J. Wagner, C. G. Berger, T. Stubhan, F. Schmitt, K. Zhang, M. Sytnyk, T. Heumueller, C. M. Sutter-Fella, I. M. Peters, Y. Zhao and C. J. Brabec, *Adv. Energy Mater.*, 2023, **13**, 2302594.
29 Y. Zhao, J. Zhang, Z. Xu, S. Sun, S. Langner, N. T. P. Hartono, T. Heumueller, Y. Hou, J. Elia, N. Li, G. J. Matt, X. Du, W. Meng, A. Osvet, K. Zhang, T. Stubhan, Y. Feng, J. Hauch, E. H. Sargent, T. Buonassisi and C. J. Brabec, *Nat. Commun.*, 2021, **12**, 2191.
30 Z. Xu, Y. Zhao, J. Zhang, K. Chen, C. J. Brabec and Y. Feng, *Phys. Rev. Mater.*, 2020, **4**, 095401.
31 T. Osterrieder, F. Schmitt, L. Lüer, J. Wagner, T. Heumüller, J. Hauch and C. J. Brabec, *Energy Environ. Sci.*, 2023, **16**, 3984–3993.
32 A. K. Y. Low, F. Mekki-Berrada, A. Gupta, A. Ostudin, J. Xie, E. Vissol-Gaudin, Y.-F. Lim, Q. Li, Y. S. Ong, S. A. Khan and K. Hippalgaonkar, *npj Comput. Mater.*, 2024, **10**, 104.
33 D. N. Cakan, R. E. Kumar, E. Oberholtz, M. Kodur, J. R. Palmer, A. Gupta, K. Kaushal, H. M. Vossler and D. P. Fenning, *Digital Discovery*, 2024, **3**, 1236–1246.
34 K. Xu, A. Al-Ashouri, Z.-W. Peng, E. Köhnen, H. Hempel, F. Akhundova, J. A. Marquez, P. Tockhorn, O. Shargaieva, F. Ruske, J. Zhang, J. Dagar, B. Stannowski, T. Unold, D. Abou-Ras, E. Unger, L. Korte and S. Albrecht, *ACS Energy Lett.*, 2022, **7**, 3600–3611.
35 Q. Dong, Y. Yuan, Y. Shao, Y. Fang, Q. Wang and J. Huang, *Energy Environ. Sci.*, 2015, **8**, 2464–2470.

36 M. Kim, G.-H. Kim, T. K. Lee, I. W. Choi, H. W. Choi, Y. Jo, Y. J. Yoon, J. W. Kim, J. Lee, D. Huh, H. Lee, S. K. Kwak, J. Y. Kim and D. S. Kim, *Joule*, 2019, **3**, 2179–2192.
37 X. Wu, S. Wang, J. Zhang, H.-W. Shiu, Y.-J. Hsu, H. Yan, J. Zhu and X. Lu, *Nano Energy*, 2023, **117**, 108907.
38 D. W. de Quilettes, S. M. Vorpahl, S. D. Stranks, H. Nagaoka, G. E. Eperon, M. E. Ziffer, H. J. Snaith and D. S. Ginger, *Science*, 2015, **348**, 683–686.
39 H. Yu, F. Wang, F. Xie, W. Li, J. Chen and N. Zhao, *Adv. Funct. Mater.*, 2014, **24**, 7102–7108.
40 M. Balandat, B. Karrer, D. Jiang, S. Daulton, B. Letham, A. G. Wilson and E. Bakshy, *Advances in Neural Information Processing Systems*, 2020, pp. 21524–21538.
41 P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt and SciPy 1.0 Contributors, *Nat. Methods*, 2020, **17**, 261–272.
42 M. V. Khenkin, E. A. Katz, A. Abate, G. Bardizza, J. J. Berry, C. Brabec, F. Brunetti, V. Bulović, Q. Burlingame, A. Di Carlo, R. Cheacharoen, Y.-B. Cheng, A. Colsmann, S. Cros, K. Domanski, M. Dusza, C. J. Fell, S. R. Forrest, Y. Galagan, D. Di Girolamo, M. Grätzel, A. Hagfeldt, E. von Hauff, H. Hoppe, J. Kettle, H. Köbler, M. S. Leite, S. Liu, Y.-L. Loo, J. M. Luther, C.-Q. Ma, M. Madsen, M. Manceau, M. Matheron, M. McGehee, R. Meitzner, M. K. Nazeeruddin, A. F. Nogueira, Ç. Odabaş, A. Osherov, N.-G. Park, M. O. Reese, F. De Rossi, M. Saliba, U. S. Schubert, H. J. Snaith, S. D. Stranks, W. Tress, P. A. Troshin, V. Turkovic, S. Veenstra, I. Visoly-Fisher, A. Walsh, T. Watson, H. Xie, R. Yldrm, S. M. Zakeeruddin, K. Zhu and M. Lira-Cantu, *Nat. Energy*, 2020, **5**, 35–49.
43 R. E. Kumar, rekumar/mixsol: Release for Zenodo Citation, 2024, **https://github.com/rekumar/mixsol/tree/v0.5.1**.
44 A. Mavlonov, Y. Hishikawa, Y. Kawano, T. Negami, A. Hayakawa, S. Tsujimura, T. Okumura and T. Minemoto, *Sol. Energy Mater. Sol. Cells*, 2024, **277**, 113148.
45 D. P. Kingma and M. Welling, Auto-Encoding Variational Bayes, 2022, **https://arxiv.org/abs/1312.6114**.
46 D. Jimenez Rezende, S. Mohamed and D. Wierstra, *arXiv*, 2014, preprint, arXiv:1401.4082, DOI: **10.48550/arXiv.1401.4082**.
47 J. T. Wilson, R. Moriconi, F. Hutter and M. P. Deisenroth, *The reparameterization trick for acquisition functions*, 2017.
48 S. Seo, M. Wallat, T. Graepel and K. Obermayer, *Proceedings of the International Joint Conference on Neural Networks*, 2000, **vol. 3**, pp. 241–246.
49 X. Chen and Q. Zhou, Proceedings of the 2014 Winter Simulation Conference, 2014, pp. 3821–3832.
50 M. Binois, J. Huang, R. B. Gramacy and M. Ludkovski, *Technometrics*, 2018, **61**, 7–23.
51 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.
52 T. K. Ho, *IEEE Trans. Pattern Anal. Mach. Intell.*, 1998, **20**, 832–844.
53 N. Li, Y. Luo, Z. Chen, X. Niu, X. Zhang, J. Lu, R. Kumar, J. Jiang, H. Liu, X. Guo, B. Lai, G. Brocks, Q. Chen, S. Tao, D. P. Fenning and H. Zhou, *Joule*, 2020, **4**, 1743–1758.
54 D. N. Cakan, C. J. Dolan, E. Oberholtz, M. Kodur, J. R. Palmer, H. M. Vossler, Y. Luo, R. E. Kumar, T. Zhou, Z. Cai, B. Lai, M. V. Holt, S. P. Dunfield and D. P. Fenning, *RSC Adv.*, 2024, **14**, 21065–21074.
55 K. Frohna, C. Chosy, A. Al-Ashouri, F. Scheler, Y.-H. Chiang, M. Dubajic, J. E. Parker, J. M. Walker, L. Zimmermann, T. A. Selby, Y. Lu, B. Roose, S. Albrecht, M. Anaya and S. D. Stranks, *Multimodal operando microscopy reveals that interfacial chemistry and nanoscale performance disorder dictate perovskite solar cell stability*, 2024.
56 J. Xu, C. C. Boyd, Z. J. Yu, A. F. Palmstrom, D. J. Witter, B. W. Larson, R. M. France, J. Werner, S. P. Harvey, E. J. Wolf, W. Weigand, S. Manzoor, M. F. A. M. van Hest, J. J. Berry, J. M. Luther, Z. C. Holman and M. D. McGehee, *Science*, 2020, **367**, 1097–1104.
57 V. M. Goldschmidt, *Naturwissenschaften*, 1926, **14**, 477–485.
58 R. E. Beal, D. J. Slotcavage, T. Leijtens, A. R. Bowring, R. A. Belisle, W. H. Nguyen, G. F. Burkhard, E. T. Hoke and M. D. McGehee, *J. Phys. Chem. Lett.*, 2016, **7**, 746–751.
59 S. M. Lundberg and S.-I. Lee, *Advances in Neural Information Processing Systems 30*, Curran Associates, Inc., 2017, pp. 4765–4774.
60 J. Wang, L. Zeng, D. Zhang, A. Maxwell, H. Chen, K. Datta, A. Caiazzo, W. H. Remmerswaal, N. R. Schipper, Z. Chen, K. Ho, A. Dasgupta, G. Kusch, R. Ollearo, L. Bellini, S. Hu, Z. Wang, C. Li, S. Teale, L. Grater, B. Chen, M. M. Wienk, R. A. Oliver, H. J. Snaith, R. A. Janssen and E. H. Sargent, *Nat. Energy*, 2024, **9**, 70–80.
61 Z. Li, Q. Xu, Q. Sun, Z. Hou and W.-J. Yin, *Stability Engineering of Halide Perovskite via Machine Learning*, 2018.
62 M. Ahmadi, M. Ziatdinov, Y. Zhou, E. A. Lass and S. V. Kalinin, *Joule*, 2021, **5**, 2797–2822.
63 M. A. Ziatdinov, A. Ghosh and S. V. Kalinin, *Mach. Learn.: Sci. Technol.*, 2022, **3**, 015003.
64 M. Ziatdinov, *Active Learning with Fully Bayesian Neural Networks for Discontinuous and Nonstationary Data*, 2024.