



Cite this: DOI: 10.1039/d5cs00715a

An AI-accelerated pathway for reproducible and stable halide perovskites

Abigail R. Hering, ^a Carolin M. Sutter-Fella ^b and Marina S. Leite ^{*a}

Halide perovskites (HPs) have remarkable optoelectronic properties, and in the last decade their photovoltaic power conversion efficiency and light-emitting diode efficiency have skyrocketed. Despite the surge in research on these burgeoning materials, two key challenges in the field remain: material irreproducibility and instability. Their behavior is especially dynamic in response to environmental stressors, due to complex interactions with the perovskite crystal lattice. In this review, we survey the latest achievements in HP materials research accomplished with the assistance of artificial intelligence (AI), through the implementation of automated experimentation and machine learning (ML) data analysis. Automated synthesis and characterization tackle problems with material irreproducibility by systematically controlling parameters with very high precision, creating massive datasets, and allowing methodical comparisons from which unbiased conclusions can be drawn. AI can reveal otherwise unnoticed trends, inform future experiments with the highest potential information gain, and forecast future performance. The review concludes with a forward viewpoint of how human-assisted closed-loop laboratories and shared databases allow halide perovskite materials' processing, properties, and performance to be potentially optimized with AI, accelerating the development of highly reproducible and stable optoelectronic devices.

Received 26th June 2025

DOI: 10.1039/d5cs00715a

rsc.li/chem-soc-rev

^a Department of Materials Science and Engineering, University of California, Davis. 1 Shields Ave, Davis, CA, 95616, USA. E-mail: mleite@ucdavis.edu

^b Molecular Foundry Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, CA, 94720, USA

1. Introduction

The rise of halide perovskites (HPs) for optoelectronics has opportunistically coincided with recent artificial intelligence (AI) breakthroughs. Machine learning (ML), which is a subset of



Abigail R. Hering

Abigail R. Hering is a PhD. Candidate in Professor Marina Leite's group in the Department of Materials Science and Engineering at the University of California, Davis, USA. Her research is on applications of machine learning methods to understand degradation of halide perovskites.



Carolin M. Sutter-Fella

Dr. Carolin M. Sutter-Fella is a Staff Scientist at the Molecular Foundry, Lawrence Berkeley National Lab, USA. She received her PhD from the Swiss Federal Institute of Technology in Zurich (ETH, Switzerland) in 2014. Carolin started her independent career in 2017 when she was awarded the Glenn Seaborg Early Career Fellowship. Her research vision is to establish a transformative change in our ability to control matter towards

directed synthesis with real time adaptive control, to enable materials by design for energy applications. Her research is devoted to synthesis science of inorganic and hybrid materials involving robot-assisted synthesis and in situ multimodal characterization.

AI, has contributed to discoveries in the complex, multi-dimensional realm of HPs materials' research. In parallel, the erratic datasets produced by HPs have resulted in challenging and compelling case studies being tackled by ML. HPs are an emerging semiconducting material class that have been the subject of intense research since their photovoltaic response was first discovered in 2009.¹ They have several unique properties, including high absorption coefficients,² tuneable bandgaps,³ defect tolerance,⁴ and ease of processing.⁵ The HPs are fabricated from earth-abundant, low-cost materials, giving them the potential to revolutionize the clean energy market.⁶ The bulk HP structure has an ABX₃ chemical formula, where the A site is a monovalent cation such as Cs, formamidium (FA), or methylammonium (MA), the B site is a divalent metal cation such as Pb or Sn, and the X site is a halide anion, such as I, Br, or Cl. The possible constituent ions may exist in any given ratio, and each composition exhibits unique properties, thus HPs possess a large, convoluted hyperparameter space. HPs are much less stable and reproducible than the conventional Si and III-V semiconductors, especially in the presence of environmental stressors of heat, oxygen, water vapor, light, and electric field bias.⁷ This instability is highly dependent on chemical composition, resulting in several nonlinear trends.^{8,9} AI is being applied to investigate different aspects of these materials' development, including composition screening,^{10–12} automated characterization,^{13–15} device optimization,^{16,17} and device stability performance forecasts.^{18–20}

In this review, we provide a summary of how AI can accelerate the resolution of the remaining challenges in the HP field: reproducibility and stability. First, we present a timeline of notable achievements in HP and AI research, as well as in scientific discoveries and accomplishments that combine the two. We discuss developments in robotic material synthesis and high-throughput characterization as accelerated methods of inferring material properties, by controlling parameters that are difficult or impossible for a human researcher to control

and thus ensuring a higher degree of reproducibility. We outline a survey of ML objectives that have been applied to material irreproducibility and environmental instability. Then, we discuss in detail how AI can make experimental decisions and inform future promising experiments. Finally, we examine and evaluate the concept of an AI-informed, closed-loop laboratory that encompasses each step of HP development, including experimental design, material fabrication and characterization, device testing, and optimization. We anticipate future developments, as well as future challenges, in this field through the expansion of sharable, interpretable databases, with automated data systems and large language model-assisted predictions. We conclude with an outlook that closed-loop labs with human input are expediting scientific breakthroughs and bringing sustainable optoelectronics, specifically HPs, closer to commercialization.

2. AI and halide perovskites: timeline and interdisciplinary achievements

Fig. 1 displays a timeline of developments in the fields of HPs and AI and their merging together. The first perovskite mineral, calcium titanate (CaTiO₃), was discovered in the Ural Mountains of Russia in 1839 by Gustav Rose, and its structure was named after the notable Russian mineralogist Count Lev Perovski.^{21,22} In the following years, several more naturally occurring oxide perovskites were discovered, including barium titanate (BaTiO₃), which has been studied extensively due to its ferroelectricity.²³ HPs, which have a halide instead of an oxygen at the X site, were first synthesized in 1978,²⁴ but their photovoltaic effect was not observed until 2009,¹ when a 3% power conversion efficiency was recorded. This first device degraded within minutes, but in 2012, improved solid-state solar cells were fabricated^{25,26} and recorded by the National Renewable Energy Laboratory (NREL) research cell efficiency chart.²⁷ There has been remarkable research progress on the optoelectronic properties of HPs, with thousands of research articles being published every year. The first HP LED was synthesized in 2013,²⁸ and the first NREL certified HP/Si tandem cell was successfully fabricated in 2018. Tandem solar cells show the greatest potential for commercial use, as they can surpass the theoretical radiative efficiency limit of single junction solar cells. HP power conversion efficiencies surpassed 25% in 2020, and their record efficiency is now 27%²⁷ for single junction cells and 34.6% for tandems, which is on par with the best-performing conventional solar cell materials. HPs are in the early stages of transitioning from research- to commercial-level, though there is still significant progress to be made in the field. Now, Oxford PV in the UK, Tandem PV in the US, and Xianna and Utmo Light in China, are currently beginning commercial development of HP and HP/Si tandem solar modules.^{29–32}

The field of AI has experienced similar exponential growth. The term was first coined in 1950, and the first paper on the topic was published in 1959.³³ AI is broadly defined as the field of computers and robots that can analyse information and



Marina S. Leite

Prof. Marina S. Leite is a Professor in the Department of Materials Science and Engineering and a Chancellor's Fellow at the University of California, Davis, USA. She received her PhD in physics from Campinas State University, followed by a post-doc at CALTECH. Prior joining UC Davis, Marina was an associate professor at the University of Maryland, College Park. Her research interests are centered in materials for sustainability,

encompassing applying AI to accelerate the understanding of halide perovskites' dynamic behavior, materials for solar cells, the advanced design of thermophotovoltaics, and optical materials.

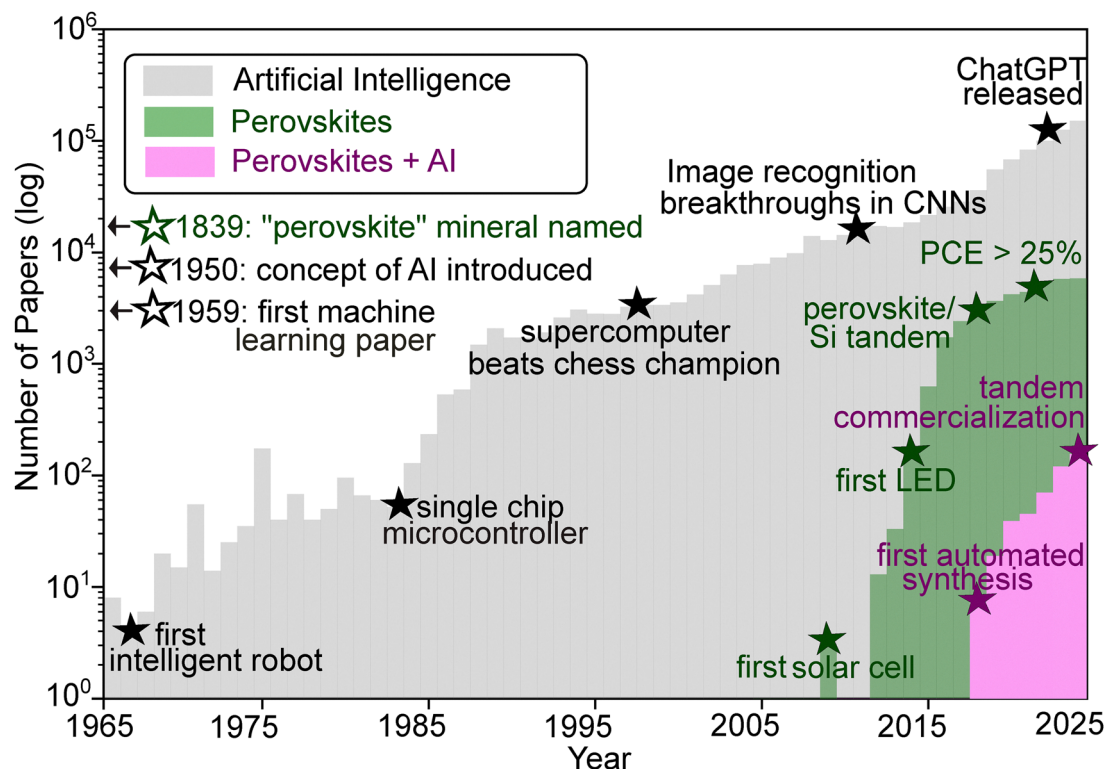


Fig. 1 Number of research papers published per year from 1965–2024 using the SCOPUS search engine, with keywords “perovskite solar cells” or “halide perovskites” in green, keywords “artificial intelligence”, “AI”, “machine learning” or “machine-learning” in grey, and the combination of all keywords in magenta. Notable achievements in each field are highlighted with black, green, and magenta stars, including the first intelligent robot (1966), the report of a supercomputer beating a chess human champion (1997), the first perovskite solar cell and LED (2009 and 2013), and the first automated synthesis of HPs (2018).

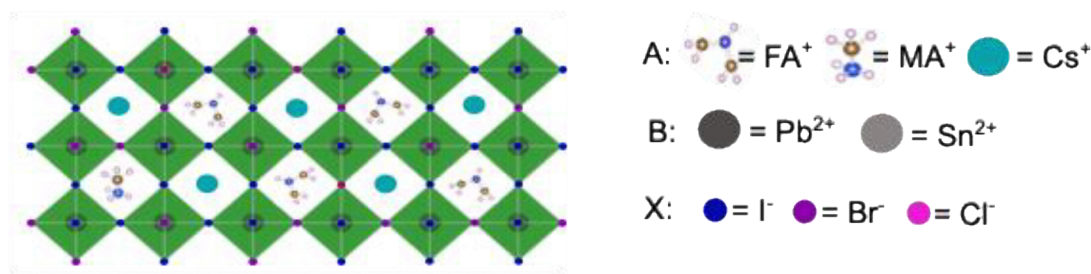
mimic human perception, cognition, and decision making.³⁴ ML, which is a range of algorithms that make predictions based on data,³⁵ is a subfield of AI that will be discussed here in the context of materials research. Several ML algorithms were developed in the 1950s, but computers lacked the required computational power to solve them. The first intelligent robot, Shakey, which had the ability to move and make decisions within its environment, was developed in 1966 by the Stanford Research Institute.³⁶ In 1984, single chip microcontrollers were developed, and computational power grew following Moore's law, which in 1985 stated that the number of transistors on a microchip would double every two years.³⁷ As computational power grew further, a notable milestone included the first supercomputer beating a chess champion, Deep Blue vs. Kasparov, in 1997.³⁸ In the 2000s, image recognition with convolutional neural networks experienced breakthroughs.³⁹ In 2022, ChatGPT was the first large language processing model to be released to the public,⁴⁰ making AI freely accessible for the first time.

In 2018, several research groups began applying ML algorithms to analyse HP behaviour.^{41–43} The first automated synthesis of HPs produced 95 compositions, and ideal wide-bandgap combinations that are promising for tandem cells were identified.⁴⁴ Since then, hundreds of papers have been published applying AI, automation, and ML to HPs, and this combined research interest shows the steepest growth. Several of these papers are discussed in more detail in the following sections, including automation of

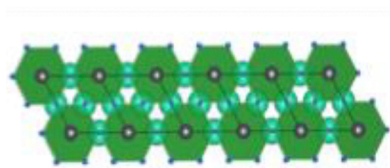
fabrication and characterization, ML paradigms, and reproducibility and stability optimization. There has been a major emphasis on predicting and optimizing solar cell power conversion efficiency (PCE), and these discoveries, including fabrication recipe optimization, device stack optimization, additive and passivation strategies, and identification of ideal storage conditions, have accelerated tandem cell commercial investment and development.^{45–47} Looking forward, as more optoelectronic applications for HPs are explored, these materials may be used in photodetectors, sensors, lasers, transistors, spintronics, and more. If stable HP-based photovoltaics become an established product, they may eventually provide the energy needed to power the rising AI computational costs. Additionally, HP spintronics,⁴⁸ which retain memory states without input power, may be used as an alternative to the extremely energy-intensive cloud data storage centres.⁴⁹ These developments in HP materials science and AI may create a positive feedback loop resulting in a computationally advanced and sustainable future.

3. Halide perovskite materials

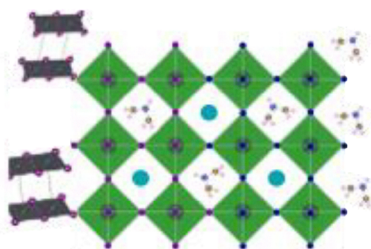
The HP crystal lattice structure, ABX_3 is shown in Fig. 2a. There are several possible compositions, as the A site may be an organic or inorganic cation, the B site may be a lead or tin

a) Halide perovskite structure ABX₃

b) Hexagonal phase



c) Defect evolution



d) Halide segregation

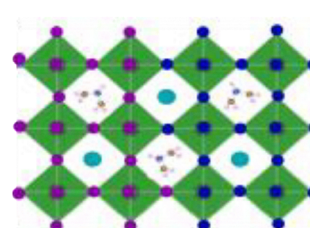


Fig. 2 Cross-section schematic of bulk halide perovskite crystal structure and common degradation pathways. (a) A hybrid organic-inorganic, mixed halide perovskite. (b) Structure of the photo-inactive, hexagonal delta phase. (c) Evolution of defects including PbI_2 formation (black), dangling surface bonds, and organic decomposition. (d) Segregation of halides into Br-rich (purple) and I-rich (blue) regions.

cation, and the X site is a halide anion. Each site may be occupied by a fraction of ions as well, which is displayed in Fig. 2a as a hybrid organic-inorganic mixed-halide composition. The ability of an ion to incorporate into the perovskite structure is predicted by the Goldschmidt⁴⁹ tolerance factor, which is calculated from the relative ionic size. An ideal factor of 0.8–1.0 will result in a cubic HP structure.⁵⁰ There are three phases that form photoactive HPs: cubic, tetragonal, or orthorhombic, which are known as alpha, beta, and gamma phases, respectively. If the strain is too high, or if the structure is stressed in some way, the HP unravels, and the material becomes photo-inactive, which is often a hexagonal, trigonal, or orthorhombic delta phase⁸ (Fig. 2b). In addition to the hundreds of possible compositions for bulk HPs, there are several other possible phases, including double perovskites, vacancy-ordered perovskites, quasi-2D phases such as Ruddlesden-Popper and Dion-Jacobson phases, chiral 2D perovskites, 2D nanosheets, 1D nanowires, 0D quantum dots, and mixed-dimensional phases.⁵¹ The broken symmetry of lower dimensional HPs allows for an even larger composition and property space, as there is no longer a size restriction on the A-site organic cation.⁵² Lower dimensional materials have experienced growing research interest in recent years, due to their extensive range of possible applications in spintronics, memristor, and sensing devices.⁵³

In response to environmental stressors, including light, bias, temperature, oxygen, and relative humidity, HPs react and degrade through a variety of mechanisms.⁵⁴ The degradation depends primarily on the original chemical composition, the severity and combinations of the stressors, and length of

exposure. The most frequently observed degradation pathways are displayed in Fig. 2b–d, although others are possible. The delta phase (Fig. 2b) can often coexist with the photoactive alpha phase at room temperature, meaning both phases are thermodynamically stable. However, only a small activation energy is required to begin the phase transition, which may be induced by defects or exposure to humidity.⁵⁵ This phase transition is reversible, and the HP may ‘heal’ from its delta phase with time or heating.⁵⁶ Additives, such as MACl , are a commonly used strategy to mitigate the delta phase formation. Other strategies include trivalent metal doping, which can relieve residual lattice stress and promote alpha phase formation,⁵⁷ and composition optimization with the appropriate ionic ratios to achieve an ideal tolerance factor.⁴⁷

The second mechanism, defect evolution (Fig. 2c), encompasses point, line, surface, and volume defects. Point defects include vacancies, interstitials, and substitutional defects, which tend to have relatively low formation energies in HPs. Line defects include edge and screw dislocations, which contribute to nonradiative carrier recombination. Surface defects include grain boundaries, twin boundaries, and cracks. Grain boundaries exhibit several unique properties in HPs and can serve as recombination centres due to ion migration and defect trapping at interfaces. However, they can also be a source of self-healing behavior,⁵⁶ and these self-healing properties can be further enhanced with functional additives, which are large molecules that sit inside the grain boundaries, or with dynamic liquid interfaces, which is an interfacial layer between the HP and the transport layer. Thermal energy, which can degrade the HP lattice, activates the reformation of bonds with additive

molecules⁵⁸ or triggers the solid-to-liquid phase transition of the interfacial layer, which passivates surface and grain boundary defects.⁵⁹ Twin boundaries are most often found in single crystals, due to asymmetry in the lower-temperature phases.⁶⁰ Finally, 3D volume defects encompass voids, pores, and inclusions, all of which are induced during the synthesis stage. The defects that most often occur in response to environmental stress are lead iodide formation,⁶¹ dangling surface bonds,⁶² or organic decomposition (Fig. 3c).⁶³

Another commonly observed degradation pathway is halide segregation (Fig. 3d). Ionic migration occurs in HPs under light, heat, or bias exposure, due to the low formation energies of ionic defects. Iodine interstitials and vacancies specifically have the lowest activation energies, as determined by first principles calculations and confirmed by photothermal induced resonance (PTIR).⁶⁴ This migration is more prevalent in grain boundaries than in grain interiors, which is confirmed by conductive atomic force microscopy (c-AFM).⁶⁴ This discrepancy results in localized differences in trap states and electronic band structures, causing nonradiative recombination centres. Halide migration is usually reversible upon removal of the stressor, but is detrimental to device operation when light and bias are present.⁶⁵ Passivation strategies for all of the degradation mechanisms discussed primarily focus on compositional, additive, and interfacial engineering.⁶⁶

Both the strengths and weaknesses of HPs arise from their soft-lattice structure. They are defect tolerant, meaning defects will not be detrimental to their performance in optoelectronic devices.⁶⁷ In contrast, inorganic materials with a rigid crystal-line lattice, such as silicon (diamond cubic) or gallium arsenide (zinc blende) used in conventional semiconductor devices,

require extremely high purity with minimal intrinsic defects or external contaminants.⁶⁸ This difference is due to their strong, covalent bonds, while HPs are held together by a combination of covalent, ionic, hydrogen, and van der Waals bonds, making it easier to break the lattice.⁶⁴ These materials can be synthesized with low-cost manufacturing techniques, rather than the expensive Czochralski single crystal growth used for silicon. Polycrystalline HPs can be fabricated through several different methods, including spin coating, drop casting, blade coating, gas quenching, slot-die coating, and spray coating,^{69,70} but these routes are typically not reproducible or scalable to industrial solar module sizes. Single crystal HPs have been grown with methods including inverse temperature crystallization, in which temperature changes the solubility of a supersaturated solution, and the Bridgman method, which grows an ingot from melted powders *via* a temperature differential in a tube, among others, but these methods also produce challenges with reproducibility and scalability, and are not yet as heavily researched as polycrystalline thin films.⁷¹ Currently, polycrystalline samples result in photovoltaics with higher power conversion efficiencies despite their increased number of defects, because of the detrimental effects of surface defects on single crystal stability.⁷²

HP defect tolerance, property tunability, and ambient processing ability all make controlling reproducibility more difficult. Thus, robotic fabrication techniques have been deployed to address parameters that are difficult or impossible to control during human-led synthesis. The various mitigation strategies introduced, including additives, doping, and interfacial engineering, are often performed on a single composition, and thus are difficult to extrapolate to all HPs. There is a massive

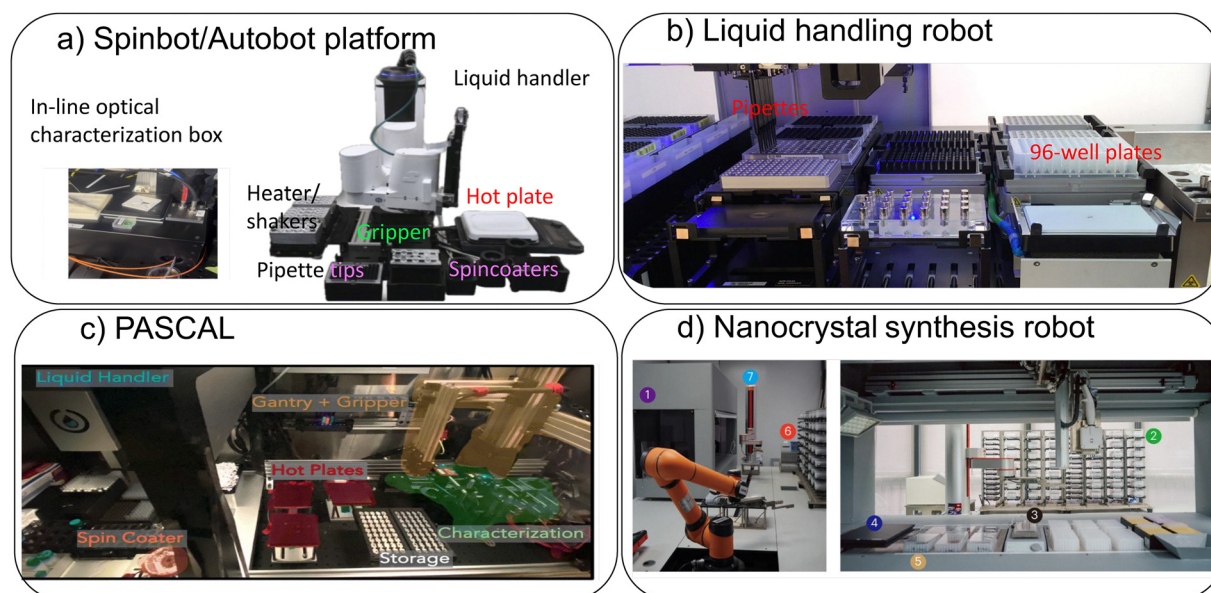


Fig. 3 Examples of robotic fabrication of halide perovskites. (a) SpinBot/Autobot platform. Adapted with permission⁷⁸ © 2023 Wiley Advanced. (b) Liquid handling robot. Photo courtesy of University of Tennessee's Tickle College of Engineering/Shawn Poynter. (c) Perovskite automated spin-coating assembly line (PASCAL). Adapted with permission¹² © 2024 RSC. (d) Nanocrystal synthesis robot, where (1) is the automated pipettes for liquid handling, (2) is storage for samples and consumables, (3) is the synthesis platform, (4) is the light source, (5) is the mobile camera, (6) is the microplate reader, and (7) is the mobile robot for microplate transport. Adapted with permission⁸⁷ © 2023 Nature.

parameter space of potential additive and interfacial molecules to be explored, and therefore understanding of stability strategies may also be greatly improved by implementing automation and ML approaches.

4. Automation

Automated labs are being implemented for their increased efficiency, accuracy, and precision, as compared with the conventional trial-and-error or institution-based approach of human experimenters. Automation also allows repetitive, menial tasks to be done by a machine, leaving the researchers with more time and energy to invest in tasks such as experimental design, planning, and analysis. Here, we discuss platforms with various levels of automation, and we emphasize that autonomous experiments,⁷³ which automate all experimental steps, including hardware, data collection, data analysis, and AI-informed decision making, are still emerging technologies. Other terms, including “self-driving labs”, “closed-loop labs”, and “materials acceleration platforms (MAPs)”, are sometimes used interchangeably with autonomous experiments in the literature,⁷⁴ though other sources differentiate them. To clarify their main differences in scope, Table 1 provides a possible interpretation of each one. Self-driving labs involve a fully integrated workflow that includes automation of hypothesis formulation, testing, and refinement steps.⁷⁵ Closed-loop labs may include human input and typically involve an objective of exploration or optimization of a parameter space.⁷⁶ MAPs have an underlying goal of accelerating the understanding of materials’ fundamental properties with the assistance of AI.⁷⁷ The robotic synthesis labs and high-throughput characterization experiments discussed below are examples of automated experiments that expand scientific understanding of HPs’ materials properties and synthesis parameter space.

4.1. Robotic synthesis labs

The primary issues with reproducibility from hand-made spin-coated HPs are from lack of fine control over the distances, speeds, volumes, and times of each synthesis step, which may be uncontrolled because the experimenters don’t consider them to be parameters. The SpinBot⁷⁸ is a commercially available liquid handling robotic platform for the synthesis of thin films from chemical precursor solutions (Fig. 3a). This setup

can produce thin films or full solar cell devices through multi-step deposition, in platforms referred to as SpinBot one,⁷⁸ the Autobot,⁷⁹ and a fully automated spin-coating robot.⁷⁰ Other notable materials science research has been done on similar robotic platforms, using “Polybot”⁸⁰ for organic thin film synthesis, “Ada” for optimizing organic and palladium films,⁸¹ the “A-lab” for solid-state synthesis of inorganic powders,⁸² and the “artificial chemist” for quantum dot synthesis.⁸³

With the Spinbot/Autobot, the effects of input parameters such as solvent dispense speed, pipette tip height, spin speed, and antisolvent dispense time and speed, may be explored and optimized.¹⁶ This robotic system can produce dozens of reproducible thin films per day. The HP thin films are characterized with steady-state photoluminescence (PL), UV-Vis absorption, and PL imaging or time-resolved PL. These quick optical characterizations can inform film quality and can be used to create an optimizable metric. This platform combines high-throughput synthesis with high-throughput characterization, which is discussed in the following sections. The Autobot platform has successfully predicted synthesis-property relationships for HPs fabricated under different atmospheric humidity conditions, and demonstrated that the role of the additive molecule during spin-coating and solvent evaporation has a significant effect.⁷⁹

Next, the liquid handling pipetting robot can synthesize and optically characterize hundreds of different HP solutions (Fig. 3b).¹⁰ This method allows rapid compositional screening and combinatorial analysis of up to 96 compositions at a time using a 96-well plate. Overall, the system has extremely precise control of composition and concentration. Decoupled plate readers allow quick optical characterization, creating large “in-house” datasets for ML-based analysis. The PL properties of the vast compositional space can be interpolated, to fully inform a HP cation-space phase diagram of solid solution phases. ML applications based on data from this setup will be discussed in the following sections.

The “PASCAL”, or perovskite automated spin-coating assembly line¹² platform was custom built for liquid handling, thin film spin-coating, annealing, and in-line characterization (Fig. 3c). This platform has a liquid handling robotic arm coupled with a spin coater. HP films can be optically characterized *via* PL and with a GoPro camera.¹⁷ This information creates a large compositional dataset, and the optimal stable compositions can be determined using an ML model, Gaussian

Table 1 Automated laboratories main terminology and corresponding definitions

Term	Definitions
Closed-loop lab	The concept of a lab with “... process automation integrating experimental execution and data acquisition/analysis.” ⁸⁵
Self-driving lab	An application of the closed-loop lab concept, or a “system in which automated experiments are integrated with data-driven decision making.” ⁷⁴
Materials acceleration platform (MAP)	Another application of the closed-loop concept with advanced materials research goals, or platform which “... uses methods of automation and digitalization in material research to accelerate innovation by orders of magnitudes.” ⁸⁴
Autonomous lab	The highest level of AI-control, or a platform which “uses advanced decision algorithms to plan and execute a series of materials experiments iteratively... the system autonomously advances through the iterations of planning, experiment, and analysis.” ⁸⁶

process regression, as the GoPro images are analysed. PASCAL allows characterization of as-spun thin films as well as films that have aged under the influence of environmental stressors such as temperature and illumination, so that fabrication conditions and compositions resulting in the highest stability can be revealed.

The nanocrystal synthesis robot⁸⁷ has been used to fabricate both colloidal metal nanocrystals and double HP nanocrystals. It uses automated pipettes for liquid handling and a robot arm for moving and optically characterizing samples. It also entails an integrated software that uses supervised ML algorithms to recognize structure identifying agents (SDAs), which are parameters that most significantly affect the crystal morphologies. Literature data mining was performed to select the initial synthesis parameters, but this database was expanded with optical measurements of the synthesized nanocrystals. This approach allows tuneable morphologies to optimize for a determined metric, which for the double HPs is PL intensity. Correlations were determined with a supervised ML algorithm called the sure independence screening and sparsifying operator (SISSO) approach. The nanocrystal synthesis robot allowed for inverse design of adjustable nanocrystals with a desired final morphology, which enables tunability of properties and design of the most stable HP compositions and structures.

These robotic synthesis platforms are examples of partial or fully autonomous MAPs,⁸⁸ in which AI unifies the entire experimental process by including coupled characterization, data analysis, ML, and decision making, with the objective of exploring and optimizing HP material properties. They can potentially reduce both intra- and inter-lab bias, allowing more accurate conclusions to be drawn from ML analysis. Inverse design, in which a desired, optimal output can be achieved from altering the experimental inputs, can lead to significantly improved control and understanding of property-processing relationships.⁸⁹ HPs have extremely multifaceted parameter spaces of processing conditions and dynamic properties, affecting their stability and reproducibility, but MAPs have made significant progress in addressing these challenges. By controlling and optimizing fine experimental details which affect the complex perovskite crystallization kinetics, such as composition, concentration, additive molecules and concentrations, solvent evaporation rates, annealing profiles, sequential deposition methods, and additive-mediated nucleation,⁹⁰ these platforms can produce reproducible samples from which robust scientific conclusions can be drawn. They are especially useful when coupled with high-throughput characterization, which is discussed in the following section.

These methods primarily focus on solution-based processes for lab-scale investigations. However, the transition to industrial-scale fabrication will require techniques such as printing, blade-coating, or vacuum processing, in order to grow uniform, larger-area films.⁹¹ Vacuum-based processing requires fine-tuning of deposition rate, temperatures, precursor compositions, and post-treatment conditions. These processes could also benefit from automation and optimization

with AI, as large reproducible films will be necessary for successful commercial products.

4.2. High-throughput material characterization

High-throughput experiments involve multiple tests performed simultaneously, either by measuring many data points per sample or over many samples.⁹² These experiments are often carried out with the assistance of robotics or automated workflows, so that they produce large datasets that are directly comparable. A data science driven approach allows the best-performing and most stable materials to be identified from a large composition and property space. HPs can be characterized through a variety of techniques to measure the optoelectronic, structural, and morphological properties, as well as device stability and performance. Coupled with automated synthesis, high-throughput characterization provides high quality data to train ML models with statistically significant results and minimal bias. Additionally, stability data can be extracted from *in situ* and *operando* experiments, in which a variable is introduced while the material is being characterized or while a device is operating. *In situ* and *operando* experiments are especially useful for determining the effects of HP degradation, as the environmental stressors may be introduced during characterization experiments or while a solar cell device is operating. These types of experiments can elucidate degradation mechanisms of HPs by measuring them in real time.⁹³ The following examples discuss different high-throughput characterization techniques and how they are being applied to furthering the understanding of HP stability and reproducibility.

Photoluminescence (PL) is an optical method that characterizes wavelength-dependent radiative recombination of photo-generated charge carriers, revealing optical bandgap and the presence of multiple phases or band edge trap sites in HPs in <1 s. Fig. 4a displays high-throughput steady-state PL data,⁹⁴ in which ten samples are measured simultaneously for 150 hours while temperature is varied between 15 °C and 55 °C. This is an example of both high-throughput and *in situ* characterization, as the HPs are being characterized while temperature is fluctuating and altering their properties. Thus, the real-time effects of temperature on the thin films could be measured and compared between different compositions of the $\text{Cs}_x\text{FA}_{(1-x)}\text{Pb}(\text{I}_y\text{Br}_{(y-1)})_3$ space. Additionally, these measurements produce thousands of comparable spectra that can be used to train a series of ML models. Then, PL properties of compositions unseen during model training are predicted with a high accuracy, and films with moderate caesium content are revealed to be the most stable to temperature modulation.

PL lifetime can reveal the defect density, and thus the intrinsic instability of different HPs. The time-resolved PL of 2D PEA (phenethylammonium) iodide perovskite samples with mixed tin and lead ratios at the B site is shown in Fig. 4b.⁹⁵ A high-throughput pipetting robot (Fig. 3b) is used to synthesize 96 Pb:Sn ratios between pure lead and tin. The samples are then characterized with steady state and time resolved PL, which shows that higher Pb content prolongs PL lifetimes.

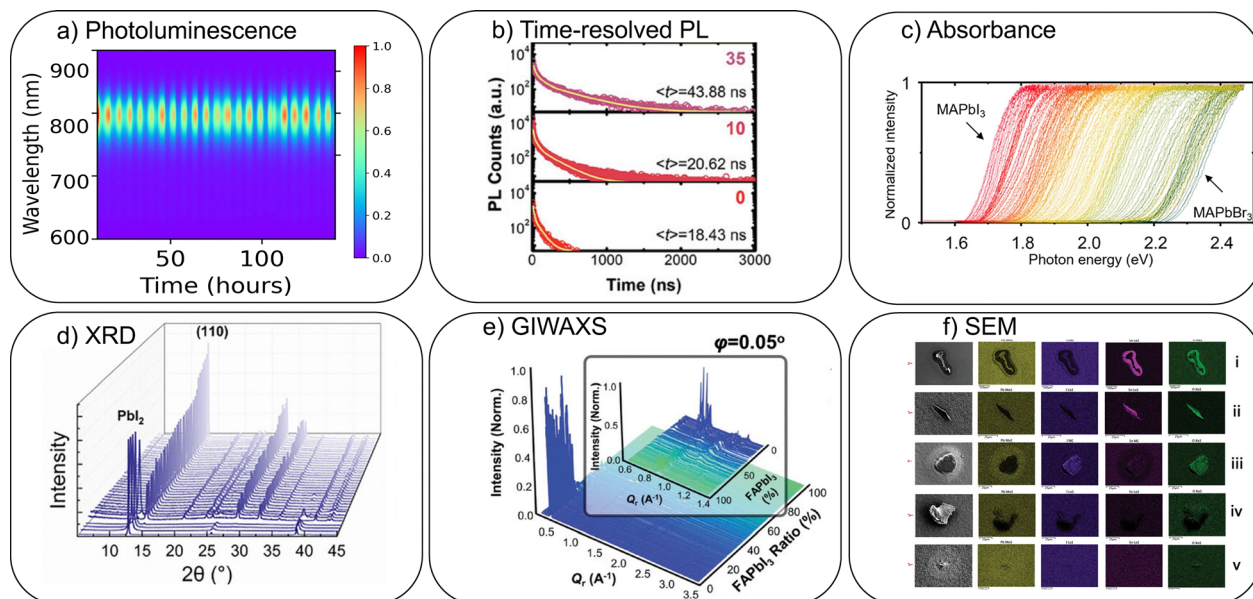


Fig. 4 Examples of high-throughput halide perovskite characterization techniques. (a) *In situ* PL measurements over several days of temperature cycling. Adapted with permission⁹⁴ © 2025 ArXiv. (b) PL lifetime decay of $\text{PEAPb}_x\text{Sn}_{(1-x)}\text{I}_4$ with varying Sn and Pb ratios. Adapted with permission⁹⁵ © 2024 Wiley Advanced. (c) Absorbance spectra of a compositionally graded halide perovskite film from MAPbI_3 to MAPbBr_3 . Adapted with permission⁹⁶ © 2022 Nature. (d) X-ray diffraction of MAPbI_3 thin films with 40 different antisolvents. Adapted with permission⁹⁷ © 2025 Wiley Advanced. (e) Grazing incidence wide-angle X-ray scattering with films of varying FAPbI_3 ratios. Adapted with permission⁹⁸ © 2024 Wiley Advanced. (f) Scanning electron microscopy with energy dispersive X-ray scattering of $\text{FA}_{0.7}\text{MA}_{0.3}\text{Pb}(\text{I}_{0.49}\text{Br}_{0.21}\text{Cl}_{0.3})_3$ films. Adapted with permission⁹⁹ © 2023 Wiley Advanced.

The proposed mechanism is phase segregation of a PEAPbI_4 phase, which establishes charge transfer excitons. 2D HP structures are more stable than 3D ones due to lack of strain and charge carrier confinement. There is also significant interest in replacing Pb with less toxic alternatives, of which Sn is the highest performing, due to its similar electronic structure. While fully eliminating Pb may not be possible without sacrificing device performance, mixed Pb:Sn perovskites show promise for high quality devices while reducing potential negative environmental and neurotoxic effects. These high-throughput experiments investigate the entire Pb:Sn composition space and reveal the defect-rich regions and their effects on phase segregation and optical performance, ultimately connected to material stability.

Absorbance of HPs reveals, in addition to the spectrally-resolved absorbance behaviour of the material, the composition-dependent bandgap, which is another metric that may be used to compare stability or reproducibility between samples. Fig. 4c displays high-throughput absorbance data of a compositionally-graded film,⁹⁶ where the MAPbX_3 film has a gradient in halide concentration from pure iodine to pure bromine. As expected, the bandgap varies nearly linearly with halide composition, which allows for a tuneable bandgap and makes HPs very suitable for tandem solar cells and several colours of LEDs. This high-throughput fabrication and characterization of a compositionally graded film enable continuous rather than discrete measurements, so that hundreds of mixed HPs can be interrogated in a short timeframe. The researchers determined that the degradation of Br-rich regions is driven by hydration, while the I-rich regions' changes originated from loss of organic

components. Intermediate compositions are driven by phase and halide segregation. This experiment allowed multiple degradation methods to be discovered simultaneously, which significantly increased the understanding of the mixed-HP composition space properties and stability.

High-throughput X-ray diffraction (XRD) characterization provides structural information to assign crystalline phases. This measurement can permit a robust screening of materials by comparing the relative intensities of diffraction peaks corresponding to HP and secondary or defect phases. This may be used for evaluation of compositions or synthesis parameters to assess film quality, typically through ensuring only the desired phases are present. Fig. 4d displays XRD patterns of HP films fabricated in a high-throughput platform while screening 40 different anti-solvents.⁹⁷ Whereas the composition space of HPs has now been extensively studied, most recipes rely on highly toxic antisolvents,¹⁰⁰ which work by extracting the solvent from the HP solution, creating supersaturation, and instigating rapid crystallization, nucleation, and growth.¹⁰¹ Antisolvents affect crystallinity and presence of defect phases, therefore impacting sample reproducibility. In this work, the antisolvent space is explored through metrics such as polarity, dispersion, and hydrogen bonding, which are compared within Hansen sphere solubility space. The authors identified sustainable hybrid antisolvent systems as well, revealing superior film properties when mixed methyl acetate and hexane are used. Optimal antisolvent use can ensure higher degree of film reproducibility.

Grazing incidence wide angle X-ray scattering (GIWAXS) is used for probing surface and near-surface structural properties,

making it ideal for investigating samples that have different properties at the surface and within the bulk of the material. Quasi-2D HPs have an even larger composition space than 3D HPs, and the relative ratios of 2D and 3D precursors may result in different phase heterogeneities within a sample. Fig. 4e shows GIWAXS data⁹⁸ for quasi-2D HPs, where the authors employed high-throughput solid-solution synthesis and characterization using the platform discussed in Fig. 3b. The quasi-2D composition space is explored by varying the ratios of 2D and 3D precursors, which creates samples with a vertical phase gradient. This systematic, high-throughput characterization explores phase heterogeneities from the sample surface to deeper into the bulk of the sample. This example advanced the understanding of mixed phase behaviours and revealed that 35–55% 3D FAPbI₃ results in the most stable 2D:3D ratio with a pure alpha phase, and that ratios outside of this range exhibited mixed 2D phases and 3D delta phase.

Microscopy techniques often require advanced data management pipelines to become high-throughput, as a single image contains on the order of tens of megabytes to hundreds of gigabytes or hundreds to thousands of pixels.^{102,103} These methods allow researchers to compare surface morphologies, for example of samples fabricated with different techniques, which is useful to compare the effects of processing on properties and on reproducibility. Fig. 4f shows high-throughput scanning electron micrographs (SEM) with electron dispersion spectroscopy (EDS) analysis.⁹⁹ The researchers collected 2500 images of large-area HP films fabricated *via* vacuum processing, and a combination with EDS analysis was performed on defect regions of the film to categorize them into pinholes or wrinkles. Through this high-throughput exploration of the thin-film surface, there were minimal pinholes observed, as compared with solution-processed films which showed a greater number of pinholes. The minor wrinkling present in the vacuum processed films was determined not to be detrimental to photovoltaic performance. The power conversion efficiency of films fabricated with vacuum processing showed significantly lower standard deviation than those that were solution processed, which confirmed the HT SEM conclusions that high-quality, uniform, and highly reproducible films can be synthesized with this technique.

5. Machine learning paradigms

ML, a subfield of AI, is defined as a set of algorithms that can make predictions or decisions based on data and new information.^{35,104} It can be applied at each step of HP fabrication and device processing, from composition screening, fabrication parameter optimization, characterization of material properties, transport layer selection, and prediction of optoelectronic device performance.^{18,105} To minimize the number of confounding variables, have better control over experimental inputs, and increase the amount of available data, automation has been extremely complimentary to the expansion of ML. AI can inform future experiments through automated data

analysis, real-time feedback loops, and integration of multiple data libraries into single interpretable metrics. AI allows for hypothesis refinement¹⁰⁶ beyond the human domain experts' intuition and guidelines. There is an ascending number of HP research articles employing ML to optimize, predict, or classify material quality, stability, and performance (120 and 156 publications in 2023 and 2024, respectively).

ML can be broadly categorized as supervised learning, unsupervised learning, and reinforcement learning. Additional strategies that do not strictly fit into these categories include semi-supervised learning, active learning, and optimization.¹⁰⁷ Supervised tasks encompass regression or classification with labelled data, and while unsupervised tasks involve clustering of unlabelled data. Reinforcement learning involves maximizing a reward function over multiple learning iterations.¹⁰⁸ Exploration and exploitation are two learning objectives, which seek to find values within a multiparameter space with high uncertainty and find optimal values with low uncertainty, respectively.¹⁰⁹ Optimization algorithms involve both exploration and exploitation, as they find the global minimum or maximum of some defined metric and identify regions of highest and lowest uncertainty. These methods all share the need for high-quality training datasets, from which the algorithms will extract patterns, trends, and correlations between variables. Training data may come from simulated data, high-throughput experiments, or shared databases, and various literature examples applying ML to HP stability and reproducibility are discussed in the following sections.

5.1. Supervised learning

Supervised learning requires labelled training data with known answers, which can be either qualitative or quantitative. One of the simplest supervised regression algorithms is linear regression, which uses a least-squares curve fitting method to fit an input and output. Tree-based algorithms, which are based on decision trees, can be used for both classification and regression tasks and include random forests, gradient boosting, extreme gradient boosting, and more. A vast number of studies have involved deep learning or neural networks, which include many different types of models such as convolutional neural networks, long short-term memory, and artificial neural networks, and these methods can also be used for both classification and regression tasks. Methods such as neural networks require significantly more data than other models, and flexible ensemble learning algorithms tend to outperform others when addressing high-dimensional materials science challenges.¹¹⁰ The majority of ML case studies applied to HPs have used regression to predict metrics such as performance or stability.

5.1.1. Regression. Regression involves prediction of a numerical test set value based on values of the training dataset, or mapping a continuous input to an output.¹¹¹ High-throughput datasets, as discussed in the automation section, are well-suited for training regression models. While these datasets may limit inter-lab discoveries, they may be very effective in training ML models on a small number of variables and comparable samples. In Fig. 5a, the PL maxima of different

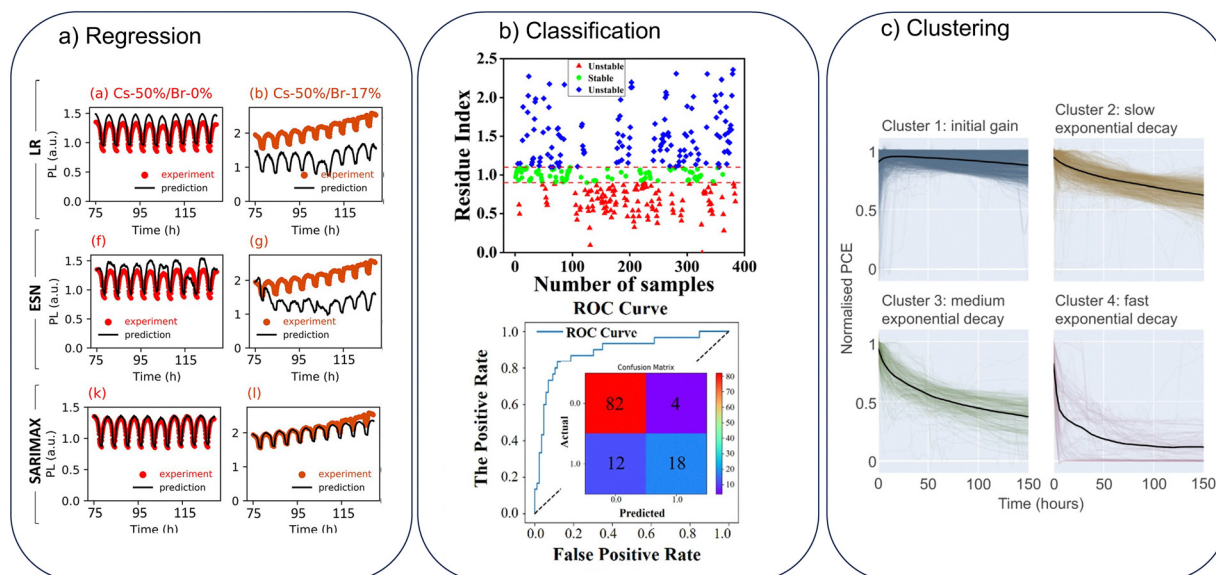


Fig. 5 Different ML paradigms applied to HP stability. (a) Regression models predicting photoluminescence (PL) maximum values with three different ML algorithms (linear regression, echo state network, and seasonal autoregressive integrated moving average with exogenous regressors) using training data from high throughput, *in situ* PL experiments. Adapted with permission¹¹² © 2023 ACS. (b) Classification of HPs as “stable” or “unstable” based on photocurrent measurements of 96 materials. Adapted with permission¹¹⁵ © 2024 Nature. (c) Clustering of power conversion efficiency (PCE) datasets over time based on the similarity of decay curves. Adapted with permission¹¹⁶ © 2023 Nature.

HP compositions under relative humidity fluctuations are predicted¹¹² with a series of supervised ML regression algorithms with increasing computational complexity: linear regression (LR), echo state network (ESN), and seasonal autoregressive integrated moving average with exogenous regression (SARIMAX), of which the latter has the highest accuracy (see Table 2 for descriptions of ML accuracy metrics). This is an example of both a high-throughput and *in situ* experiment, as five samples are characterized simultaneously while relative humidity is varied. The three ML models are all forecasting future behaviour with respect to time, giving significant insight into the intrinsic moisture stability of $\text{Cs}_x\text{FA}_{1-x}\text{Pb}(\text{I}_y\text{Br}_{1-y})_3$ HP composition space.

In other works, regression models trained on high-throughput, *in situ* optical transmittance data revealed changes in carrier diffusion length and quasi-Fermi level splitting in MAPbI_3 and forecast this behaviour over time.¹¹³ Following the ISOS (international summited on organic PV stability) protocols,¹¹⁴ the effects of stressors on material changes could be evaluated in a standardized manner, and the most stable material and device conditions were verified. Regression algorithms, specifically through predictions and forecasting, are ideal for predicting the long-term stability of HPs, especially in response to dynamic environmental stressors.

5.1.2. Classification. The second type of supervised ML is classification, in which labelled data is mapped from an input to an output.¹¹¹ The data itself may be either numerical or categorical, but the predictions will only be categorical.

In Fig. 5b, an extra trees classification model is used to predict whether an HP is stable or not.¹¹⁵ The researchers investigated the aqueous photoelectrochemical stability of

MAPbI_3 HPs with variation in additives, solvents, and post-treatment molecules. Bare HP films degraded very quickly in aqueous solutions and produced no measurable photocurrent. However, the optimally treated samples showed high photocurrent over long immersion time, indicating superior moisture stability. During training, residual indices of photocurrent between 0.9 and 1.1 are considered stable, while others are classified as unstable. The aqueous photoelectrochemical measurements allow several different parameters to be explored, including HP materials, precursor solvent ratios, additives, post-treatments, and water immersion time. The accuracy of classification models is evaluated with the receiver operating characteristic (ROC) curve and area-under-curve (AUC) plots (Fig. 5b), and with a confusion matrix showing the true positive, true negative, false positive, and false negative predictions (see Table 2 for all evaluation metric definitions).

In another example, researchers predicted perovskite synthesizability using graph neural networks.¹¹⁷ By classifying the synthesizability of oxide, halide, hydride, and chalcogenide perovskites, new material candidates can be identified for experimental variation. Classification methods have the benefit of being easily interpretable, but a disadvantage is that the labels must be predefined by the researcher; thus, some prior domain knowledge is beneficial. “Stable” and “unstable” may be clearly defined, but “synthesizable” and “unsynthesizable” may not be, because there is no training data for the “unsynthesizable” case.¹¹⁷ This deep neural network case study outperforms traditional ML methods due to its ability to learn structure-property relationships from large datasets, and to then transfer this knowledge to small, domain-specific datasets.

Table 2 Machine learning accuracy evaluation metrics most used in halide perovskite research, for both regression and classification

Type of model	Evaluation metric	Full name	Definition	Formula
Regression	R^2	Coefficient of determination	Proportion of variance in dependent variable explained by model.	$R^2 = \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2}$
	r	Correlation coefficient	Strength and direction of a linear relationship between two continuous variables.	$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$
	MAE	Mean absolute error	Average absolute difference between the actual and predicted values in the dataset.	$MAE = \frac{1}{N} \sum_{i=1}^N y_i - \hat{y} $
	MSE	Mean-squared error	Average of the squared difference between the actual and predicted values in the dataset.	$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2$
	RMSE	Root-mean squared error	Square root of MSE, standard deviation of residuals.	$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2}$
	NRMSE	Normalized root-mean squared error	Normalizes RMSE by mean of observed values.	$NRMSE = \frac{1}{Y_{\max} - Y_{\min}} \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2}$
Classification	AUC	Area under the curve	Performance metric for binary classification, trade-off between TPR and FPR	Area under ROC curve
	ROC	Receiving operating characteristic	Curve that visualizes a binary classification model's performance across all thresholds	Plot of TPR (true positive rate) vs. FPR (false positive rate)
	TPR (Recall)	True positive rate (recall)	Rate of predicted true positives (TP) over actual positives (true positive plus false negative)	$TPR = \frac{TP}{\text{Actual positive}} = \frac{TP}{TP + FN}$
	FPR	False positive rate	Rate of predicted false positives (FP) over actual negatives (true negative plus false positive)	$FPR = \frac{FP}{\text{Actual negative}} = \frac{FP}{TN + FP}$
	Precision	Precision	Percent of true positives over all positive predictions	$\text{Precision} = \frac{TP}{TP + FP}$
	F1 Score	F1 Score	Harmonic mean of precision and recall.	$F1 \text{ score} = \frac{2(\text{Precision} \times \text{recall})}{\text{Precision} + \text{recall}}$

5.2. Unsupervised learning

Unsupervised learning includes clustering, association, and dimensionality reduction algorithms. In this type of ML, the data is unlabelled, so there are no target outputs or reward functions.¹¹⁸ Unsupervised learning finds patterns in the input data by estimating probability distributions. Association rule learning uncovers hidden relationships by exploring “if-then” relationships between data points. Clustering algorithms group unlabelled data based on similarity between points. Dimensionality reduction techniques, such as principal component analysis (PCA) and *t*-distributed stochastic neighbour embedding (*t*-SNE), transform high-dimensional data into 2D or 3D visualizations, which allows researchers to gain insights on local or global data clusters and distributions across the parameter space.¹¹⁹ These approaches allow researchers to elucidate patterns in multivariate data without needing to know the mathematical relationships between inputs and outputs. Overall, dimensionality reduction models are useful to understanding HP stability and reproducibility, due to the large number of parameters that can influence this material.

5.2.1. Clustering. Clustering groups data on its similarity or dissimilarity by examining the scale of the data within its environment.¹¹⁸ These types of algorithms are beneficial when the input variables affecting the output are either unlabelled or unknown. However, during training, the number of clusters must be set by the programmer, which may require multiple iterations to find the optimal number of clusters. For example, in Fig. 5c, a clustering algorithm groups aging curves of PCE of over 2000 photovoltaic devices into four categories based on the

shape of the curve decay.¹¹⁶ The data shows that after 150 hours of operation under controlled environmental conditions, there is a statistically significant correlation between maximum PCE and stability. In this work, all HP compositions were from the $\text{Cs}_x\text{MA}_y\text{FA}_z\text{PbI}_m\text{Br}_n$ family. The PCE curves are grouped into clusters of initial gain and slow, medium, and fast exponential decay. This study reveals that the most stable HP samples correspond to the highest efficiency devices, as seen in the “initial gain” cluster. This analysis also reveals the intrinsic HP irreproducibility, as there are a variety of possible intertwined degradation mechanisms occurring during the aging experiments. While the researchers cannot correlate these observations directly with physical degradation mechanisms, they hypothesize that devices that already have imperfections during material synthesis will experience faster degradation due to evolution of defects. Therefore, unsupervised ML shows that stability and reproducibility of HP materials are inherently linked.

5.3. Reinforcement learning

In reinforcement learning (RL), the machine or agent learns from its environment over many iterations, or actions, to maximize a reward function.¹¹⁸ This learning method involves a trial and error approach to make decisions within an environment, and it may either follow a model or be model-less. It seeks to explicitly solve a goal,¹⁰⁸ rather than simply identify data patterns. There has been very limited research done using RL methods in the HP field, and we anticipate an opportunity to leverage the knowledge achieved by other materials science

fields. For instance, RL has been used for the inverse design of inorganic materials while considering their enthalpy of formation, electronegativity balance, and charge neutrality,¹²⁰ for identifying mechanically tough 2D materials,¹²¹ and for nanostructures entailing optimal emission for thermophotovoltaics.¹²² In a recent robotics example, a scanning probe microscope was trained as an autonomous robot with RL to remove molecules from a supramolecular structure.¹²³ Similar to the HP synthesis space, the supramolecular one is extremely complex, with large uncertainties and sparse feedback. The authors used model-based RL with a positive reward for lifting a molecule successfully and a negative reward for rupturing bonds, which trains the agent to perform nanofabrication without human intervention. This type of ML may be extended to the HP space if an informative reward function and learning goal are defined. For instance, one could implement reinforcement learning to discover thermodynamically stable HPs with specific spectral response (*e.g.*, well-defined photoluminescence features) for multi-junction all-HP photovoltaics, LEDs, and other applications.

ML results are evaluated through scores such as R², MAE, MSE, RMSE, NRMSE, *r*, and MRE for regression tasks¹²⁴ and TPR, FPR, AUC, ROC, and F1 for classification tasks, which are described in more detail in Table 2. We note that currently, most studies applying ML to HPs have chosen a single evaluation metric to assess the performance of the ML models used. Yet, we advocate that publications display at least three metrics for a comprehensive judgment of all models implemented. This strategy could prevent some of the common flaws observed when using ML, such as underestimated variance, and could provide a more robust comparison between different models test sets.¹²⁵

6. AI for material and device reproducibility

The irreproducibility of HPs creates large intrinsic uncertainties within the synthesis and property parameter spaces. This irreproducibility is present from HP film fabrication to device performance. In Fig. 6a, the difference in PL peak energy between batches of films fabricated with the PASCAL¹² platform is displayed. Parallel processing, in which the synthesis steps are performed together, results in significantly higher variation than serial processing, in which the synthesis steps are performed sequentially. This is likely due nitrogen purging of the glove box in between the synthesis steps of the sequential method. Atmospheric fluctuations from evaporating solvents can result in higher variations between films, so precise control is necessary, even within an inert glove box environment.

In Fig. 6b, device performance is compared between robot-fabricated and human-expert fabricated devices,⁷⁰ where the robot-fabricated batch has the lowest standard deviation in PCE. In this robotic platform, all pipetting and spin-coating steps are fully automated. The system is inside a nitrogen glovebox, and there is constant pumping of solvent vapours

to control the atmosphere during crystallization. The variation is attributed to the human experts' inability to achieve the same level of antisolvent dispense and timing control. To demonstrate the device irreproducibility across many different labs, with various levels of automation, the Perovskite Database^{126,127} shows the PCE of all submitted devices. The increasing PCE trend is seen, but there is still significant variation. Often, only the highest performing device from an experimental batch is reported, making the true level of irreproducibility difficult to determine.

Environmental factors affect both HP stability and reproducibility. In Fig. 6c, the PCE of 150 devices operating in a humid and a dry environment are compared.¹²⁸ The coefficient of variation (CV) is significantly higher in the humid environment. These studies demonstrate the variation in HP materials, devices, and device operation and reveal the need for precise experimental control and optimization. ML models that take uncertainty into account tend to perform the best for addressing this challenge. Autonomous experiments are also well-suited, as they provide high-quality ML training data. Optimization algorithms, and particularly Bayesian optimization, are useful in the field of HP materials science, as they allow a parameter space with high uncertainty to be explored and optimized.¹²⁹ Further experiments can be completed, and more advanced predictions can be made if autonomous labs make decisions based on ML-inferred results. An essential part of the self-driving lab involves AI-informed decisions, in which several iterations of experiments and ML are repeated for maximum information gained and a known uncertainty threshold. Through AI-informed experiments, the path from fundamental science to commercial manufacturing of reproducible HP optoelectronics can be streamlined.

This section presents three examples of Bayesian optimization, which find optimal solutions for HP synthesis, device synthesis, and tandem PV operation, by finding areas of maximum uncertainty and exploring them with further experiments. Challenges such as synthesis variability, domain shift, measurement noise, and out-of-distribution generalization may be mitigated with appropriately defined sampling windows, adequate noise in models, and continuous sampling and refining of ML models upon expansion of datasets. Random train-test splits is standard practice for in-distribution performance, and out-of-distribution performance may be defined by materials science criteria such as elemental groups, space groups, point groups, or crystal systems, in which case leave-one-out cross validation methods can improve results.¹³⁰ Gaussian process regression is especially useful for sparse or noisy datasets, and is an effective surrogate model for optimization, due to the function taking inherent noise and uncertainty into account.

6.1. Synthesis optimization

Optimization is a frequent ML objective that seeks to minimize or maximize a certain metric within a parameter space. Several algorithms are used for optimization, such as gradient descent, stochastic optimization, Monte Carlo optimization,

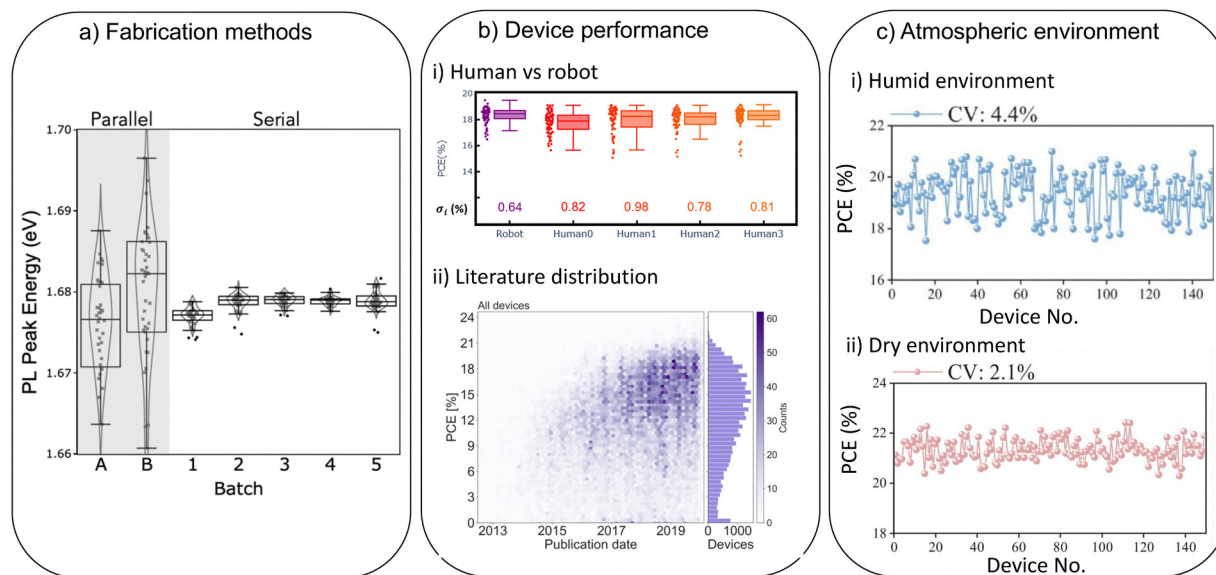


Fig. 6 Examples of statistical distributions within HP research showcasing high irreproducibility and need for high levels of control through automation and optimization with AI. (a) Box plots of PL peak location of HP thin films fabricated in parallel vs. serial processing, in which fabrication steps are performed together *versus* sequentially. Adapted with permission.¹² © 2025 RSC. (b) Device performance when fabricated with a robot vs. human experts. Standard deviation is lowest with robotic fabrication. Variation across all HP devices submitted to the perovskite database. Adapted with permission.^{70,126,127} © 2024 ACS and © 2021 Nature. (c) Variation in PCE across 150 devices operating in (i) humid *versus* (ii) dry environments. Adapted with permission.¹²⁸ © 2025 Wiley.

and Bayesian optimization (BO).¹³¹ BO is exceptionally well-suited for systems with unknown experimental responses,¹³² due to the large, multidimensional parameter spaces, noisy data, and unknown underlying properties involved. It searches for a global optimum efficiently by identifying experiments that will give the maximum information gain. Each iteration of the experiments refines the optimization search, as the total correlation acquisition function reveals points in the parameter space with the lowest correlations to each other and to previously tested ones, which are then quantified with Kullback–Leibler divergence.⁷⁹

In Fig. 7a, the researchers explored the synthesis parameter space with the AutoBot (Fig. 3a) and identified the optimal conditions for a high-quality thin film, by maximizing a unitless metric known as the *total score*, which represents material quality.⁷⁹ Here, a high-quality film is defined as having high PL intensity, high above-bandgap absorption, low below-bandgap absorption, and uniform PL images. This case aims at understanding the underlying physical and chemical properties of HPs under different synthesis conditions. The total score is based on optical measurements of thin films, and it allows film quality to be accessed quickly and efficiently, without the need for full solar cell device fabrication. This metric increases the fundamental understanding of the role of synthesis parameters, including antisolvent drop time, relative humidity during spin coating, annealing temperature, and annealing time, on the materials' optical properties. A high feature importance of spin coater relative humidity was discovered, which prompted the further exploration of HP hydration interactions *via in situ* PL. In turn, these experiments revealed the interactions between

the methylammonium chloride (MACl) additive and water, which ultimately affect film nucleation and grain growth during the spin coating processing step. This synthesis space exploration gives vital insight to HP formation kinetics and shows that high-quality HP films can be fabricated in ambient atmosphere.

6.2. Device optimization

HP devices also suffer from irreproducibility, so the performance parameter space must also be optimized for reproducible, high-performing devices to become commercialization ready. For instance, a common optimization metric for solar cells is photovoltaic PCE, but different metrics such as quantum yield, transmission, or other figures of merit may be useful as well. In Fig. 7b, an optimized solar cell efficiency of 23.7% was achieved in a six-dimensional experimental parameter space using the automated platform SPINBOT (Fig. 3a).¹⁶ The experiments involved 77 trials of unique process parameter sets, which included spin speed 1, spin speed 2, spin duration 1, spin duration 3, dripping speed, and spin speed 3, which are varied across the four iterative rounds informed by Bayesian optimization. Speed and duration 1 refer to the electron transport layer (ETL), speed 2 and dripping speed refer to the HP solution, and speed and duration 3 refer to the hole transport layer (HTL). Spin speed of the SnO₂ ETL layer is revealed to be the most important feature for device performance, which reveals that it is necessary to optimize each layer of the device instead of just the HP layer alone. The optimal processing parameter conditions were discovered efficiently with minimal human input. The results also show decreasing variance with

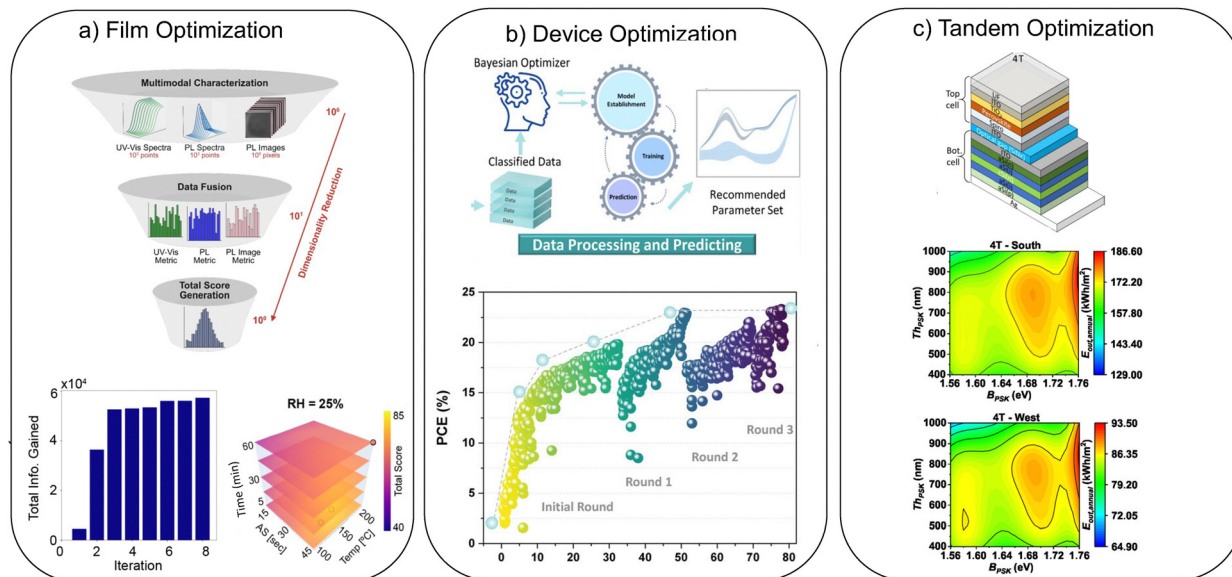


Fig. 7 Examples of ML addressing halide perovskites' reproducibility with optimization objectives. (a) Film optimization of total score by exploring synthesis parameter space with Gaussian process regression and Bayesian optimization. Total score is generated from multimodal characterization via UV-vis spectroscopy, photoluminescence, and PL imaging. Data fusion combines these measurements into a single, total score. Total information gained or cumulative Kullback–Leibler divergence for subsequent experimental iterations. Map shows total score synthesis space. Adapted with permission.⁷⁹ © 2025 Wiley. (b) Photovoltaic device power conversion efficiency (PCE) optimization with Bayesian optimization over four experimental rounds. Adapted with permission.¹⁶ © 2025 Wiley. (c) Tandem solar cell configurations with Bayesian optimization of monthly energy output and predicted annual energy output. Adapted with permission.⁴⁵ © 2024 Wiley.

increasing PCE across the experimental iterations, indicating higher reproducibility with control of these synthesis parameters.

Carrier transport layer optimization is a vital part of device optimization, and there are several fabrication techniques and potential materials for both HTLs and ETLs, making this another large parameter space to consider. Recent developments in transport layer fabrication have included atomic layer deposition,¹³³ spin-coating,¹³⁴ and self-assembled monolayers (SAMs).¹³⁵ SAMs are ultrathin layers that passivate interfaces to mitigate nonradiative carrier recombination, and their dipole moments facilitate hole extraction. They also possess an enormous molecular design space of different anchoring, inking, and head groups.¹³⁶ Therefore, AI is very suitable for accelerating the development of SAMs by optimizing for a desired property, such as stability, flexibility, rigidity, or device performance.

6.3. Tandem optimization

Because HP-Si tandem cells have surpassed the efficiency of single junction HP and Si solar cells, they are the closest to industrial and commercial distribution (see Fig. 7c⁴⁵ for optimized tandem cell). Multi-terminal cells include multiple junctions, with higher energy bandgap cells in the top of the stack and lower-energy bandgap cells in the bottom of the stack. While these architectures are more complex, they allow absorption of a broader range of the solar spectrum, which increases the PCE beyond the theoretical radiative efficiency limit.¹³⁷ As an example, it has been shown that annual output power and energy yield of tandem solar cells in Japan's outdoor

conditions could be predicted using five ML models: ensembles of trees, Gaussian process regression, regression trees, support vector machines, and an artificial neural network. By comparing simulations of 2-terminal, 3-terminal, and 4-terminal tandem cells operating in five different Japanese cities, it was found that the latter rooftop structure has the highest performance for building-integrated photovoltaics in blue-rich solar spectrum zones, as three HP bandgaps may be used to capture the largest portion of solar irradiance. The input parameters were HP thickness, HP bandgap, incident solar spectrum angle, and irradiance. Their methods included 5-fold cross validation and Bayesian optimization of monthly energy output predictions to ensure the highest accuracy of their annual energy output predictions.

ML has been invaluable to furthering fundamental scientific understanding of HPs, and it has been expanded to device engineering and product quality, stability, and reproducibility testing. While novel optoelectronic applications of HPs are still being explored, photovoltaics is the closest to market readiness. ML has been applied to photovoltaic devices for interfacial engineering, tandem optimization, performance forecasting, and outdoor testing. Optimization algorithms have greatly improved HP synthesis processes.

7. AI for material and device stability

Once reproducible HP materials are achieved, then reproducible photovoltaics, tandem cells, and light-emitting devices may be optimized for stability and performance. High-throughput

experiments, *in situ* and *operando* experiments, and calculations of intrinsic thermodynamic stability are methods being applied to solving HP stability. Methods for improving the understanding of material and device stability include composition engineering, additive engineering, interfacial engineering, dynamic liquid interfaces,⁵⁹ and environmental stressing experiments. When stability is finally achieved, HPs will be able to replace or supplement the existing PV materials on the market.

7.1. Composition engineering

Compositional engineering can be achieved by fabricating entire HP composition space families and characterizing their properties, or their properties can be calculated as a method of material screening. First principles methods⁴³ break down complex problems to the fundamental physics of a system. Density functional theory (DFT), the most extensively used method, solves the unique functional of the probability density of the ground state and can be used to calculate the band structure of materials.¹³⁸ First principles methods have been used to calculate theoretical thermodynamic stability of HP compositions, which can save experimental time, as they allow experiments to be directed to theoretically stable compositions. These computational techniques can be used to create simulated data to determine various materials properties, such as bandgap, decomposition energy, theoretical defect-limited photovoltaic PCE.^{139,140} In Fig. 8a, thermodynamic stability of all-inorganic HPs with B-site alloying is calculated using a combination of DFT and ML with crystal graph convolutional neural networks (CGCNNs), which are trained on 41 400 compositions from 3159 DFT datasets.¹³⁹ Alloying of the B site with up to four elements creates more stable structures due to large configurational entropy. From this screening, compositions with high stability and optimal bandgaps were identified while avoiding the toxicity of Pb and the instability of Sn-based HPs.

In other cases, high-throughput DFT has been adopted to compute the properties of 495 pseudo-HP structures, including bandgap, lattice parameters, decomposition energy, and theoretical photovoltaic efficiency.¹⁴¹ Another computational report explored the bandgaps and stability of 5158 lead-free alternatives.⁴¹ Combining first-principle calculations and high-throughput experimental data, optimally stable HPs composition can be discovered.¹⁴² Additional first principles methods, such as linear combinations of atomic orbitals¹⁴³ are necessary to improve DFT by including calculations for both local and delocalized bonds. However, these approaches are limited to hundreds of atoms, which may limit their accuracy in complex bulk materials.

Despite the benefits of computational datasets, the simulated data are limited in accuracy and are still computationally expensive. First principles methods can be complementary to experimental approaches but cannot fully replace experiments. There are often discrepancies between theoretical and experimental values, due to confounding properties such as defect electronic levels, charge carrier transport layer properties, and interfacial properties not being accounted for in calculations.¹³⁹

The irreproducibility of HPs makes it difficult to computationally model the complex electronic structures of native defects.¹⁴⁴ While there are challenges to this approach, combined DFT and ML studies can allow for higher ML accuracy by significantly expanding limited datasets.¹⁴⁵

7.2. Interfacial engineering

The surface of an HP film often contributes to nonradiative recombination losses, leading to surface degradation, due to the greater number of defects at the surface.¹⁴⁶ To solve this issue, surface passivation strategies include pseudo-halide anion engineering,¹⁴⁷ surface chemical polishing,¹⁴⁸ additives,¹⁴⁹ passivation layers,⁶⁶ and interfacial engineering.¹⁵⁰ An archetypal HP solar cell architecture is shown in Fig. 8b, where a buffer sits between the active layer and the carrier transport ones.¹⁵¹ These layers are required for effective electron and hole transport, and the interfaces between them may either accelerate degradation or passivate the surface defects in the HP itself. In this work, an ML screening of 175 molecules for functional groups finds that PAPzO ((2-(5,5-dioxido-10*H*-phenothiazin-10-yl)ethyl)phosphonic) acid and PAPz ((2-(10*H*-phenothiazin-10-yl)ethyl)phosphonic acid) are the most suitable, through decreasing trap state densities and increasing carrier lifetimes. The performance of these molecules is confirmed experimentally through 1200 hours of *operando* measurement of maximum power point, after which they maintain over 90% of their original PCE. Thus, interfacial engineering is an established and valuable method for enhancing HP device stability.

7.3. Photovoltaics outdoor operation

While lab-scale degradation testing extracts useful information, outdoor testing is necessary to fully understand the effects of real environmental factors on HPs.¹⁵² With expanding datasets, forecasts of PCE over time have been an area of growing interest as well. This value can be predicted from the HP composition and device architecture, but it will change and possibly degrade under real outdoor operating conditions, making accurate forecasts essential. An effective ML pipeline is also necessary for complex predictions and forecasts.¹⁵³ First it is necessary to acquire the dataset, which may be experimental, computational, or shared literature values. The next steps are to validate and clean the data, and to select the parameters, which can include structural, compositional, device, or environmental variables. The next task is model selection, which may include several different options so that comparisons can be made between them. Model training includes hyperparameter training and cross-validation. Finally, during the evaluation step, unseen test data should be assessed, and the accuracy metrics and feature importance should be quantified. Using this pipeline, Kernel ridge regression (KRR) was determined to have superior forecasting ability for HP power output. In Fig. 8c, the ML predictions of maximum power point over 150 hours of operation are compared with the true operational values. The algorithm is trained in different accelerated indoor environmental test conditions, so that the parameters with the highest contribution to material degradation can be extracted through

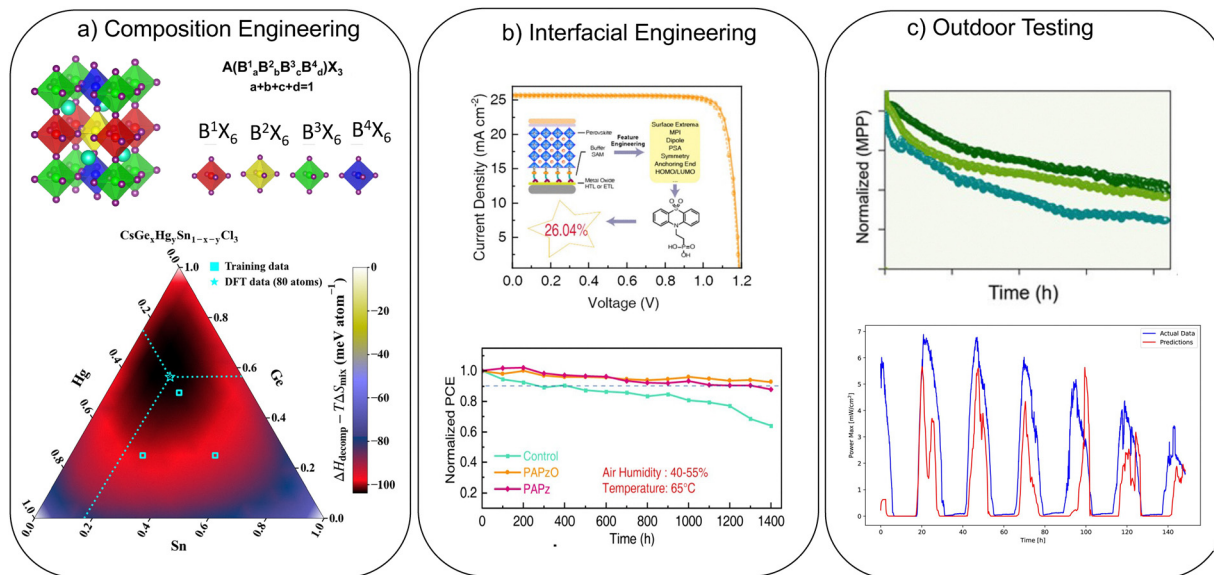


Fig. 8 Examples of ML applied to HP stability. (a) Thermodynamic stability determined by DFT and ML. Adapted with permission.¹³⁹ © 2024 Nature. (b) Interfacial engineering and optimization of HP solar cells PCE. Adapted with permission.¹⁵¹ © 2024 ACS. (c) Outdoor HP device testing. Adapted with permission.¹⁵² © 2024 ACS.

relative feature importance. The indoor tests are performed following ISOS¹⁵⁴ protocols to ensure generalizability and interpretability between labs. Six device types are used to train a bidirectional long short-term memory neural network, which achieves high accuracy and reveals that a combination of illumination and air are the most important environmental features when performing outdoor performance testing. This means that degradation mechanisms under illumination and ambient conditions must be further understood, so that they may be passivated.

While there is no “ideal” stable HP composition, the studies discussed above reveal the most stable HP materials under environmental stressors are those with tolerance factors near 1 and with device stacks incorporating passivated interfaces, as the majority of nonradiative recombination occurs at the perovskite surface. Future studies may combine the multiple composition, interfacial, carrier transport layer, and additive parameters in their analysis.

8. Accessibility and data sharing

For ML methods to be broadly established across HP and materials science research, the results and predictions must be explainable to the general scientific community. This information may be attained by reporting multiple accuracy metrics, incorporating standardized data sharing across the community, emphasizing relative feature importance, and highlighting the importance of databases and repositories for further discovery.

8.1. FAIR principles

The FAIR (findable, accessible, interoperable, reproducible)¹⁵⁵ data principles lay out a set of guidelines for sharing scientific

data. HP research can be expanded if self-driving labs follow these principles for sharing data, code, and data analysis details. Databases also increase the accessibility of AI, by providing training data for computational-based research groups that may not have the resources or expertise for expensive experimental laboratory equipment to perform their own experiments. While the robotic synthesis platforms discussed in the sections above require high initial investments and expertise in robotics and AI, open-access code databases and simpler platforms are allowing autonomous labs to become more accessible. A simple self-driving lab can be built for as little as \$100,¹⁵⁶ which means that labs with less resources can begin building skills and code repositories that are relevant to the growing AI field, increasing accessibility.

8.2. Interpretability tools

Feature importance is an “XAI”, or explainable artificial intelligence method, that allows interpretation of ML results.¹⁵⁷ It can be extracted from several ML algorithms, and it informs which experimental variables have the largest effect on the output, so that they can be further explored. This versatile technique can be applied to datasets at each stage of HP development, as seen in Fig. 9. While ML models are typically thought of as “closed-box” systems, where the underlying mathematical relationships are unknown, the relative feature importance of the input variables can give insight into the empirical relationships between variables.^{158,159} SHapley additive exPlanations (SHAP) analysis and local interpretable model agnostic explanation (LIME) are the most commonly used feature importance tools, and they can be extracted to inform the strength of each input on the predicted output, improving the interpretability of ML results.

SHAP values are model-agnostic and are determined from game theory and assigned importance values.¹⁶⁰ Game theory is

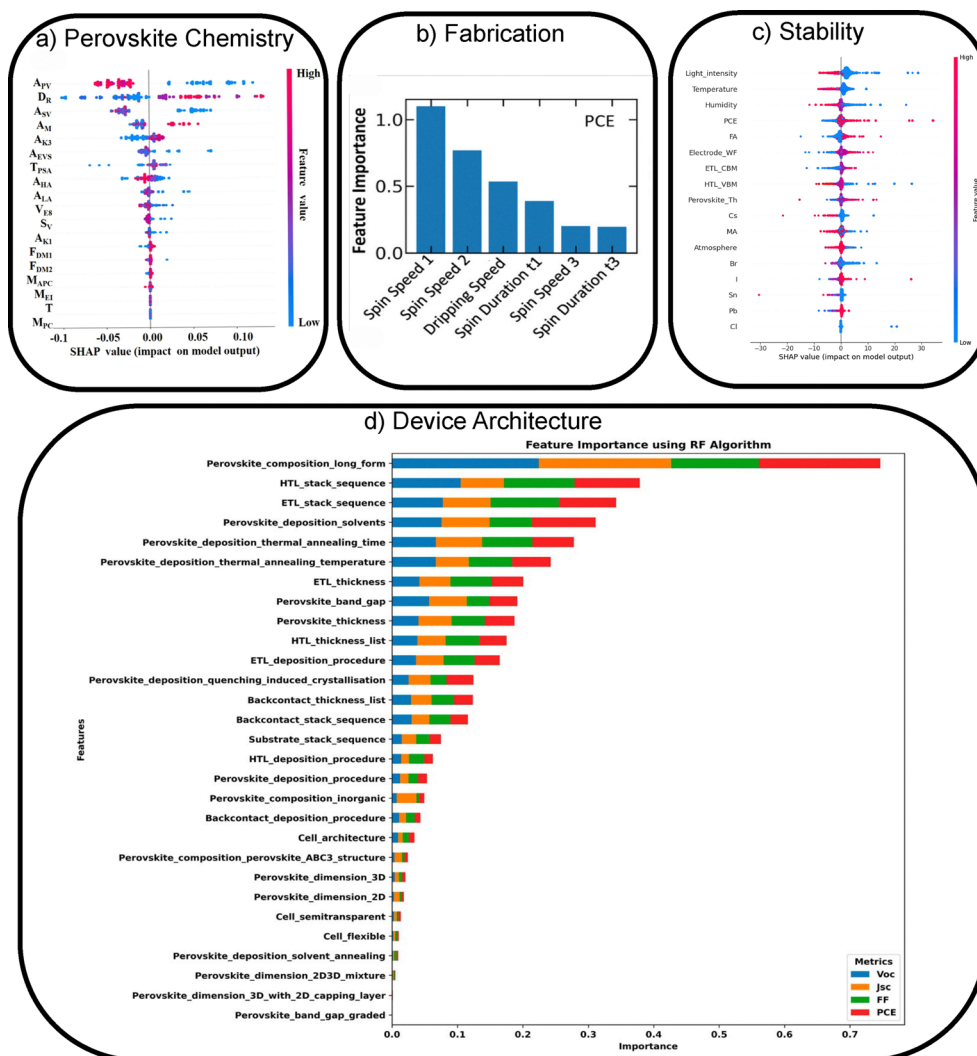


Fig. 9 Examples of relative feature importance on halide perovskite chemistry, thin film fabrication, environmental stability, and photovoltaic architecture. (a) SHAP values of additive chemical features and their effect on aqueous photocurrent. Adapted with permission.¹¹⁵ © 2024 Nature. (b) Feature importance of synthesis parameters on device power conversion efficiency. Adapted with permission.¹⁶ © 2025 Wiley (c) Environmental stressors and their effects on material stability. Adapted with permission.¹⁶³ © 2025 ACS (d) Relative feature importance of device architecture metrics on solar cell figures of merit. Adapted with permission.¹⁶⁴ © 2023 AIP Publishing.

a branch of mathematics that considers optimization, specifically by players choices of actions, and whether or not players actions affect each other's results or payouts.¹⁶¹ It considers the input features as players, and the output targets as payouts. Higher positive SHAP values indicate an enhanced positive contribution to the output, while more negative scores denote adverse contributions to the output. LIME, or local interpretable model-agnostic explanations, offers local explanations, which may be beneficial for explaining specific individual predictions, while SHAP offers global explanations for the entire models' predictions. Because it only considers local values, LIME computations are significantly faster than SHAP calculations. It is based on feature perturbation method, which is another mathematical theory which finds approximate solutions to complex problems by finding true solutions to simpler problems.¹⁶² The average absolute value of these metrics is

often taken so that the features can be easily compared. The feature importance extraction is an extremely versatile technique that can help researchers decide on next experiments in each step of the material development process, in precursor preparation, thin film synthesis, material stability testing, and device testing.

An analysis of 96 HP compositions with varying additives, solvents, and post-treatment molecules (Fig. 9a)¹¹⁵ evaluated the effects of this vast chemical space on aqueous photocurrent, which informed both stability and performance. The SHAP analysis revealed that atomic charge distribution (A_{PV}) was the most important feature, followed by solvent DMF:DMSO ratio (D_R) and hydrophilicity (A_{SV}). This method combines ML and first principles calculations, to understand the extremely complex chemical space of molecule-material interactions. After a SHAP analysis of synthesis input variables, it is

shown that spin speed 1 and spin duration has the strongest and lowest effects on device efficiency, respectively (Fig. 9b).¹⁶ In future experiments, spin speed can be kept constant while other fabrication parameters are explored. A similar SHAP analysis from the AutoBot reveals the most important feature on material synthesis is relative humidity during spincoating.⁷⁹

An investigation of film stability showed that light intensity, temperature, and humidity are the strongest factors contributing to HP instability (Fig. 9c).¹⁶³ Here, T80, or the time it takes for PCE to decrease to 80% of its original value, is adopted as a metric of stability. The researchers trained a multi-head SEResNet model on 906 perovskite datasets from the literature. The relative feature importance of the environmental stressors and composition inputs elucidated that the environmental stressors matter more than the HP composition when predicting device stability, emphasizing the need for passivation strategies for long term device performance.

Fig. 9d shows the feature importance of several device architecture metrics on the key solar cell figures of merit (open-circuit voltage: V_{OC} , short-circuit current: J_{SC} , fill factor: FF, and PCE).¹⁶⁵ Twenty-nine features from 26 000 Sn-based HP experiments recorded in the perovskite database¹²⁷ are used to train and screen a series of ML models, and random forest had the highest performance for figure-of-merit prediction. The relative feature importance, which is extracted from the decision trees within the random forest model, uncovers which optimal parameter combinations result in the highest performing device. In sum, the HP composition, hole transport layer stack sequence, and electron transport layer stack sequence, are the top three features informing photovoltaic figures of merit. Uncovering the relative importance of input parameters can aid with interpretation of ML results and can inform future experiments.

8.3. Databases

The perovskite database project,¹²⁶ which is part of the novel materials discovery laboratory (NOMAD)¹⁶⁶ repository for sharing materials science data, follows the FAIR data sharing principles and seeks to increase data sharing within the research community. The goal of this database is to document the massive amounts of data produced in a universal manner, to inform data science and ML-based discoveries. The database contains over 100 experimental parameters and data from over 42 000 photovoltaic perovskite devices. Despite the massive parameter space this entails, its utilization has been very modest, as it is limited to photovoltaic devices and requires continuous updates from researchers.

There is a considerable amount of data involving pure material characterization missing. Additionally, most experiments involve only one or two experimental variables, leaving most of the database table entries blank. There are also challenges in defining stability metrics,¹⁶⁷ which can contribute to misleading conclusions. T80, the time for PCE to reach 80% of its initial value, is a common FOM (figure of merit), but this is again limited to full PV devices and cannot characterize thin film stability. Even definitions of T80 vary, as some

researchers do not consider initial “burn-in” decay time. Therefore, there is a pressing need for higher data standardization when reporting to databases. A proposed solution to these limitations is to share complete, unanalysed or raw datasets, including undesirable values, to ensure that each shared dataset is pre-processed and interpreted in the same way, ensuring higher reproducibility.¹⁶⁸ Training ML algorithms on a variety of datasets can reduce bias and improve accuracy and transparency for the entire HP community.

As one successful example of the database usage, an ML study explored the optimal materials combinations, deposition methods, and storage conditions for efficient and stable solar cell devices, with the HP and hole transport layers each having 15% relative feature importance.^{102,169} Other groups have explored the T80 metric in the database *via* clustering algorithms,¹¹⁶ though it is noted that limited data availability in the database is restricting the quality of stability predictions.¹⁶⁷ In Fig. 10a, the bandgap of mixed HPs is predicted with a gradient boosted regression tree with optimized parameters (GBRT-P) algorithm.¹⁷⁰ The optimized parameters are found to be characteristics such as Pauling's electronegativity, dipole polarizability, electron affinity, and mobility, which allow the bandgap to be predicted with much higher accuracy than simply inputting the HP composition.

Other materials science databases, such as the inorganic crystallographic structure database (ICSD), have been used to forecast the bandgaps of 75 low dimensional lead-free HPs.¹⁴ Thus far, inter-lab breakthroughs have been limited, due to a lack of standardization, sharing of only “good” results, and limited parameter spaces. However, databases produced “in-house” tend to be computationally expensive¹⁷¹ and limited in their interoperability. Most ML applications have used a combination of data from the literature, data simulated with first principles methods, and experimental input.

In Fig. 10b,¹⁷² the prediction of PCE as a function of energy levels of the electron transport layer, HP, and hole transport material band structure is shown. This study used ML to optimize for composition and predict performance based on the highest occupied molecular orbital and lowest unoccupied molecular orbital levels of each layer, the bandgap of the HP, the exciton binding energy, crystallinity, carrier mobility, and grain size. The ML models were informed by 333 data points from 2000 papers from the literature, which included both experimental and computational values. The survey of ML methods includes linear regression, *k*-nearest neighbours, support vector regression, random forest, and an artificial neural network. The ML predictions were complemented with newly fabricated material compositions that were not in the literature training and testing datasets. The researchers fabricated compositions with mixed Cs-MA cations, Pb or Sn metal, and mixed I-Br halides, which experimentally confirmed the ML model's PCE predictions, and confirmed that the highest PCE occurs at a bandgap of 1.2 eV, which matches the theoretical radiative efficiency limit predictions. This result demonstrates that the ML models can extract physical trends from relatively small datasets.

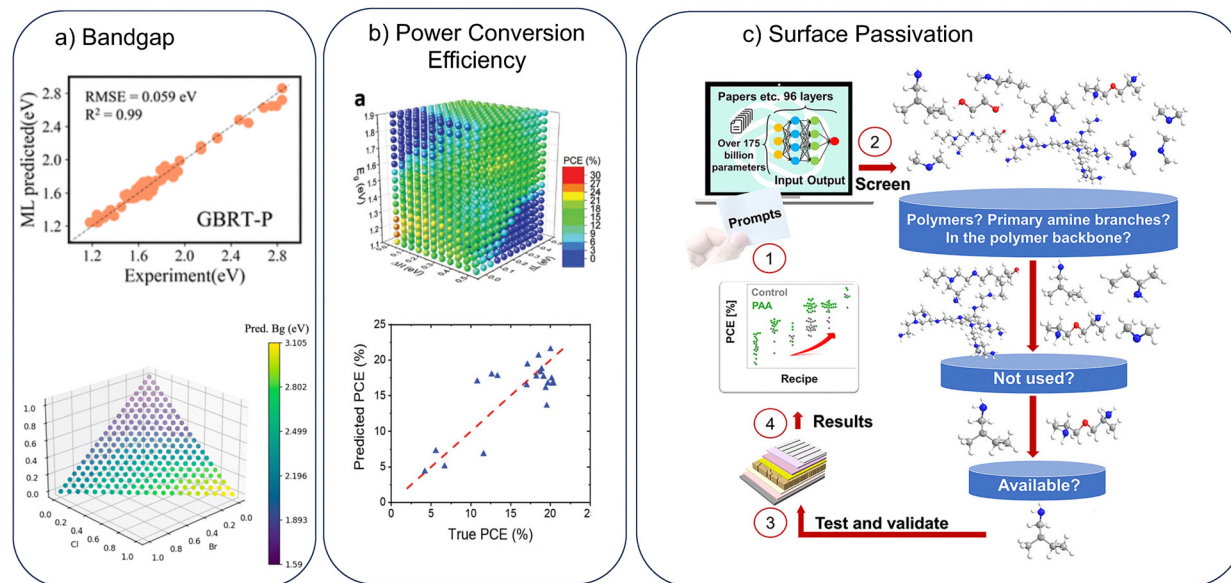


Fig. 10 Predictions with databases. (a) Bandgap predicted from perovskite database project. Adapted with permission.¹⁷⁰ © 2023 ACS. (b) Power conversion efficiency (PCE) based on shared literature values. Adapted with permission.¹⁷⁴ © 2019 Wiley. (c) Surface passivation predicted with ChatGPT hypothesis. Adapted with permission.¹⁷³ © 2025 Elsevier.

8.4. Large language models

Large language models (LLMs) may be able to accelerate data and database interpretation, experimental design, and hypothesis formation. In Fig. 10c, a hypothesis driven by ChatGPT was created from several thousand literature papers and billions of possible parameters.¹⁷³ The authors used ChatGPT to brainstorm hypotheses to identify surface passivation molecules for HP solar cells. They identified polyallylamine as a potential surface modifier, from a massive space of potential molecules, and experimentally verified its capability. While this area of materials science is extremely novel, it demonstrates that LLMs, when prompted appropriately by human experts, can reveal insights that may have been otherwise overlooked. The rapid growth in AI research to address HP stability and reproducibility requires oversight by human experts in the field to ensure scientifically sound results. The details of all datasets discussed in this review are summarized in Table 3.

MaterialsBERT,¹⁷⁵ an LLM, has been used to extract chemical relationships from the literature. Similar scientific LLMs include PolyBERT for polymers and PubMedBERT for medical publications. These models are trained specifically on scientific texts, in materials science, chemistry, and medical fields, respectively. They are more likely to produce accurate results than generalized LLMs such as ChatGPT, because their training datasets, though smaller, are better controlled. ChatGPT, in contrast, is trained on public information on the internet.¹⁷⁶

Despite these novel applications to materials science, LLMs are still limited to text prediction, so further insights are required to extract true physical and chemical relationships. Models that are physics-informed, meaning that they combine physical laws with their learning algorithms, allow prior

domain knowledge to be leveraged with data-driven analyses. This approach may include built-in restraints or boundary conditions on the ML models, which can increase their accuracy and applicability to real experimental datasets. Looking forward, we anticipate a significant expansion of physics-informed ML applied to HP research. In turn, these models will lead to more accurate, scalable, and reliable results.¹⁷⁷

9. Challenges and concerns

The recent growth of the AI field has revealed significant challenges that must be addressed. Some major concerns include the massive energy and water consumption by data storage centres, potential biases and inaccuracy, and unintentional plagiarism and theft of intellectual property. Data storage centres consume up to 5 million gallons of water per day to cool their processing units and avoid overheating.¹⁷⁸ In 2023, data centres used 4% of electricity consumed in the U. S, and this is expected to quadruple by 2030, because they require a constant supply of energy.¹⁷⁹ They are also currently producing around 2% of the U. S. greenhouse gas emissions, which is also projected to consistently increase. Therefore, there is an urgent need for renewable energy sources and alternative data storage devices as the demand for AI continues to grow daily. Novel optoelectronic devices such as HPs have the potential to either power the high energy demands of AI through PV or to store high density data in a way that does not require continuous energy input through memory devices such as spintronics, memristors, and artificial synapses.¹⁸⁰ Ethical concerns about data privacy, intellectual property, and potential for misinformation can be mitigated by using specialized models such as MATBert rather than ChatGPT, so that the training dataset is

Table 3 HP datasets discussed

Dataset figure	Source	Size	Features (independent variables)	Targets (dependent variables)
Fig. 4a	High-throughput photoluminescence spectra	137 000 spectra	Time, temperature, cation ratio, halide ratio	$PL_{\max}$, PL_{area} , PL_{fwhm} , PL_{loc}
Fig. 4b	High-throughput time-resolved photoluminescence spectra	96 spectra	Lead to tin ratio	Lifetime
Fig. 4c	High-throughput absorbance spectra	215 spectra	Halide ratio	Bandgap
Fig. 4d	High-throughput X-ray diffraction patterns	40 patterns	Antisolvent properties	Alpha peak ratio
Fig. 4e	High-throughput grazing-incidence wide-angle x-ray scattering patterns	96 patterns	2D/3D ratio	Alpha peak ratio
Fig. 4f	High-throughput scanning electron microscopy images	2500 images	Fabrication technique	Surface defects
Fig. 5a	High-throughput photoluminescence spectra	7200 spectra	Time, relative humidity	$PL_{\max}$
Fig. 5b	High-throughput photocurrent measurements	96 measurements	Additives, solvents, post-treatment molecules	Stability
Fig. 5c	The perovskite database project	2000 decay curves	NA	Decay curve shape
Fig. 6a	Photoluminescence spectra	7 batches, 45 sub-cells/batch	Deposition method	PL peak energy
Fig. 6b	High-throughput power conversion efficiency measurements; the perovskite database project	Nine batches of devices, >40 000 PCE measurements	Year, human vs. robot	PCE
Fig. 6c	High-throughput power conversion efficiency measurements	150 devices	Relative humidity	PCE coefficient of variation
Fig. 7a	High-throughput photoluminescence spectra, UV-vis spectra, and PL images	40 batches, 4 samples/batch, 17 measurements/sample	Relative humidity, antisolvent drop time, annealing time, annealing temperature	Total score
Fig. 7b	High-throughput power conversion efficiency measurements	77 devices	Spin speed 1, spin speed 2, spin duration, dripping speed, and spin speed 3	PCE
Fig. 7c	Annual energy output calculations	42 000 simulations run	Thickness, bandgap, incident angle, irradiance, number of terminals, location, bandgap	Energy output
Fig. 8a	DFT simulations thermodynamic stability	41 400 compositions	B site quaternary alloy ratio	$\Delta H_{\text{decomp}} - T\Delta S_{\text{mix}}$
Fig. 8b	High-throughput power conversion efficiency measurements	175 devices, 1200 hours of operando measurement	175 molecules	Maximum power point
Fig. 8c	High-throughput maximum power point measurements	6 devices, 150 hours of operando measurement	Temperature, relative humidity, time, atmosphere	PCE
Fig. 9a	High-throughput aqueous photocurrent measurements	96 compositions	Additive, solvent, post-treatment molecules	Stability
Fig. 9b	High-throughput power conversion efficiency measurements	77 devices	Spin speed 1, spin speed 2, spin duration, dripping speed, and spin speed 3	PCE
Fig. 9c	Literature papers published 2016–2023	906 perovskite datasets	Light intensity, temperature, humidity, composition	T80
Fig. 9d	The perovskite database project	26 000 experimental records	Device architecture	Power conversion efficiency, short-circuit current, open-circuit voltage, and fill factor
Fig. 10a	The perovskite database project	42 000 perovskites	Pauling's electronegativity, dipole polarizability, electron affinity, mobility	Bandgap
Fig. 10b	Power conversion efficiency	333 perovskites	Halide ratio, cation ratio, metal ratio, bandgap	PCE
Fig. 10c	Surface passivation	Several thousand papers	Billions of parameters	Surface passivation molecules

well controlled.¹⁸¹ Increasing the quality and quantity of training data can significantly improve the accuracy of predictions, and oversight by domain experts can provide the critical need for curated datasets. Equally important, databases must have adequate security practices, to ensure that only the correct information is shared. Researchers may need to relinquish their rights to data ownership to the databases, for the overall benefit of the HP community.

10. Conclusions and outlook

In conclusion, we showcased the promise of AI to tackle the two major constraints for HP commercialization: irreproducibility and instability. First, we outlined a timeline of major accomplishments in the fields of AI, HPs, and recent reports of their combined efforts. Next, we discussed material-based reasons for HP instability and irreproducibility. An outline of automated laboratories was presented, where the power of AI for

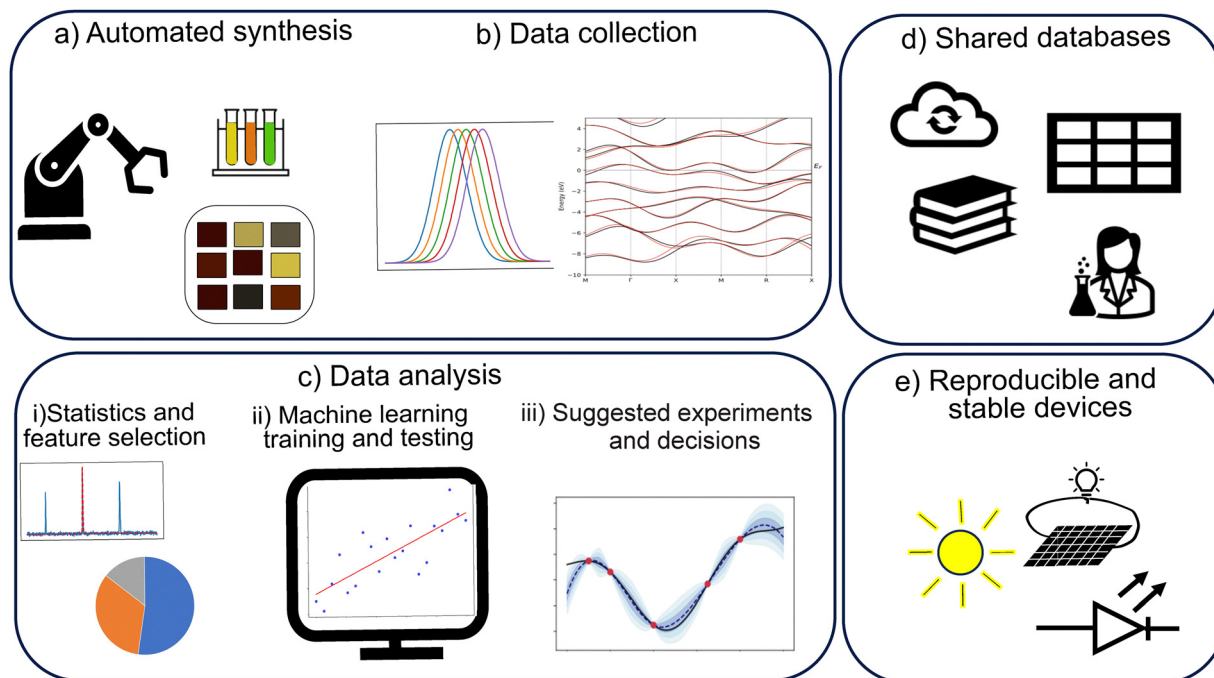


Fig. 11 Proposed workflow for autonomous closed-loop laboratories based on AI-informed halide perovskite development. (a) Automated perovskite solution, thin film synthesis, or device processing. (b) Data collection via (i) high-throughput characterization, or (ii) theoretical simulations. (c) Data analysis involves (i) statistical methods and feature selection, (ii) machine learning model training and testing, (iii) suggested experiments and decisions. (d) Shared databases and human expert decision making with the assistance of AI. After this step, steps (a)–(c) are repeated until (e) reproducible and stable optoelectronic devices are achieved.

materials development is just starting to be implemented. Examples from the literature of automated synthesis and characterization platforms are described in detail for how they contribute to the understanding of HP reproducibility and stability. A survey of machine learning approaches, including supervised, unsupervised, and reinforcement learning, as well as optimization tasks was presented with various examples from the HP literature.

Evidence of HP irreproducibility is discussed to motivate the use of AI in experiments. The use of ML to achieve stable compositions, interfaces, and operational solar and tandem cells were discussed. To improve accessibility and interpretability, we encourage the more widespread use of databases, reporting relative feature importance, and incorporating LLM models for database interpretation.

Looking forward, AI and materials research are becoming increasingly integrated, but this combined field is still in its early days. The platforms discussed here are not fully autonomous, and their scientific accomplishments are still basic compared to their potential. Autonomous closed-loop laboratories, represented in Fig. 11, involve decision making for the task of optimization.¹⁸² Advancements in materials science research are being accelerated with AI, which, in turn, strengthens the discovery of new materials for increased computational power that drives AI itself. Both fields are becoming codependent, as computational power is once again limited by material choices. HP research illustrates the significant impact of AI, by reducing the amount of time researchers spend on routine research tasks and

allowing them to focus more on higher-level tasks such as experimental design, theory, and data interpretation. We anticipate that databases will become an essential component of closed-loop laboratories. While these workflows are becoming increasingly automated, it is still necessary to have a human scientist driving the process.¹⁸³ Even in the case of LLMs, a human expert is providing the prompts to the language model. Human-AI teams continuously outperform fully AI-driven approaches. Through closed-loop laboratories with AI-driven data analysis, HPs are on the brink of commercialization and widespread adoption as highly reproducible and stable optoelectronic devices.

Author contributions

A. R. H. wrote the first manuscript version. C. M. S.-F and M. S. L. supervised. All authors contributed to the discussion and manuscript revision.

Conflicts of interest

There are no conflicts to declare.

Data availability

No primary research results, software, or code have been included, and no new data were generated or analysed as part of this review.

Acknowledgements

M. S. L acknowledges the financial support from the National Science Foundation (EPMD program, awards #2023974 and #2415023). A. R. H acknowledges the U. S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program. The SCGSR program is administered by the Oak Ridge Institute for Science and Education for the DOE under contract number DE SC0014664. C. M. S-F. acknowledges support by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences in the Early Career Research Program “Accelerated Robotic Design of Energy Materials (ACE Lab)” and the Molecular Foundry supported by the Office of Science, office of Basic Energy Sciences, of the U. S. Department of Energy Under Contract No. DE-AC02-05CH11231.

References

- 1 A. Kojima, K. Teshima, Y. Shirai and T. Miyasaka, Organometal Halide Perovskites as Visible-Light Sensitizers for Photovoltaic Cells, *J. Am. Chem. Soc.*, 2009, **131**(17), 6050–6051, DOI: [10.1021/ja809598r](#).
- 2 C. Tablero Crespo, Absorption Coefficients Data of Lead Iodine Perovskites Using 14 Different Organic Cations, *Data Brief*, 2019, **27**, 104636, DOI: [10.1016/j.dib.2019.104636](#).
- 3 G. E. Eperon, S. D. Stranks, C. Menelaou, M. B. Johnston, L. M. Herz and H. J. Snaith, Formamidinium Lead Trihalide: A Broadly Tunable Perovskite for Efficient Planar Heterojunction Solar Cells, *Energy Environ. Sci.*, 2014, **7**(3), 982–988, DOI: [10.1039/C3EE43822H](#).
- 4 A. Zakutayev, C. M. Caskey, A. N. Fioretti, D. S. Ginley, J. Vidal, V. Stevanovic, E. Tea and S. Lany, Defect Tolerant Semiconductors for Solar Energy Conversion, *J. Phys. Chem. Lett.*, 2014, **5**(7), 1117–1125, DOI: [10.1021/jz5001787](#).
- 5 M. Abdelsamie, J. Xu, K. Bruening, C. J. Tassone, H.-G. Steinrück and M. F. Toney, Impact of Processing on Structural and Compositional Evolution in Mixed Metal Halide Perovskites during Film Formation, *Adv. Funct. Mater.*, 2020, **30**(38), 2001752, DOI: [10.1002/adfm.202001752](#).
- 6 Z. Song, C. L. McElvany, A. B. Phillips, I. Celik, P. W. Krantz, S. C. Watthage, G. K. Liyanage, D. Apul and M. J. Heben, A Technoeconomic Analysis of Perovskite Solar Module Manufacturing with Low-Cost Materials and Techniques, *Energy Environ. Sci.*, 2017, **10**(6), 1297–1305, DOI: [10.1039/C7EE00757D](#).
- 7 Y. Zhou and Y. Zhao, Chemical Stability and Instability of Inorganic Halide Perovskites, *Energy Environ. Sci.*, 2019, **12**(5), 1495–1511, DOI: [10.1039/C8EE03559H](#).
- 8 Y. An, C. A. R. Perini, J. Hidalgo, A.-F. Castro-Méndez, J. N. Vagott, R. Li, W. A. Saidi, S. Wang, X. Li and J.-P. Correa-Baena, Identifying High-Performance and Durable Methylammonium-Free Lead Halide Perovskites via High-Throughput Synthesis and Characterization, *Energy Environ. Sci.*, 2021, **14**(12), 6638–6654, DOI: [10.1039/D1EE02691G](#).
- 9 D. Zhang, D. Li, Y. Hu, A. Mei and H. Han, Degradation Pathways in Perovskite Solar Cells and How to Meet International Standards, *Commun. Mater.*, 2022, **3**(1), 58, DOI: [10.1038/s43246-022-00281-z](#).
- 10 K. Higgins, S. M. Valletti, M. Ziatdinov, S. V. Kalinin and M. Ahmadi, Chemical Robotics Enabled Exploration of Stability in Multicomponent Lead Halide Perovskites via Machine Learning, *ACS Energy Lett.*, 2020, **5**(11), 3426–3436, DOI: [10.1021/acsenergylett.0c01749](#).
- 11 J. C. Stanley, F. Mayr and A. Gagliardi, Machine Learning Stability and Bandgaps of Lead-Free Perovskites for Photovoltaics, *Adv. Theory Simul.*, 2020, **3**(1), 1900178, DOI: [10.1002/adts.201900178](#).
- 12 D. N. Cakan, R. E. Kumar, E. Oberholtz, M. Kodur, J. R. Palmer, A. Gupta, K. Kaushal, H. M. Vossler and D. P. Fenning, PASCAL: The Perovskite Automated Spin Coat Assembly Line Accelerates Composition Screening in Triple-Halide Perovskite Alloys, *Digital Discovery*, 2024, **3**(6), 1236–1246, DOI: [10.1039/D4DD00075G](#).
- 13 J. M. Howard, E. M. Tennyson, B. R. A. Neves and M. S. Leite, Machine Learning for Perovskites' Reap-Rest-Recovery Cycle, *Joule*, 2019, **3**(2), 325–337, DOI: [10.1016/j.joule.2018.11.010](#).
- 14 S. Sun, N. T. P. Hartono, Z. D. Ren, F. Oviedo, A. M. Buscemi, M. Layurova, D. X. Chen, T. Ogunfunmi, J. Thapa, S. Ramasamy, C. Settens, B. L. DeCost, A. G. Kusne, Z. Liu, S. I. P. Tian, I. M. Peters, J.-P. Correa-Baena and T. Buonassisi, Accelerated Development of Perovskite-Inspired Materials via High-Throughput Synthesis and Machine-Learning Diagnosis, *Joule*, 2019, **3**(6), 1437–1451, DOI: [10.1016/j.joule.2019.05.014](#).
- 15 A. Wiczorek, A. G. Kuba, J. Sommerhäuser, L. N. Caceres, C. M. Wolff and S. Siol, Advancing High-Throughput Combinatorial Aging Studies of Hybrid Perovskite Thin Films via Precise Automated Characterization Methods and Machine Learning Assisted Analysis, *J. Mater. Chem. A*, 2024, **12**(12), 7025–7035, DOI: [10.1039/d3ta07274f](#).
- 16 J. Zhang, V. M. Le Corre, J. Wu, T. Du, T. Osterrieder, K. Zhang, H. Zhang, L. Lüer, J. Hauch and C. J. Brabec, Autonomous Optimization of Air-Processed Perovskite Solar Cell in a Multidimensional Parameter Space, *Adv. Energy Mater.*, 2025, **15**, 2404957, DOI: [10.1002/aenm.202404957](#).
- 17 D. N. Cakan, E. Oberholtz, K. Kaushal, S. Dunfield and D. P. Fenning, Bayesian Optimization and Prediction of the Durability of Triple-Halide Perovskite Thin Films under Light and Heat Stressors, *Mater. Adv.*, 2025, **6**(2), 598–606, DOI: [10.1039/D4MA00747F](#).
- 18 M. Srivastava, J. M. Howard, T. Gong, M. Rebello Sousa Dias and M. S. Leite, Machine Learning Roadmap for Perovskite Photovoltaics, *J. Phys. Chem. Lett.*, 2021, **12**(32), 7866–7877, DOI: [10.1021/acs.jpclett.1c01961](#).
- 19 Ç. Odabaşı and R. Yildirim, Assessment of Reproducibility, Hysteresis, and Stability Relations in Perovskite Solar Cells

- Using Machine Learning, *Energy Technol.*, 2020, **8**(12), 1901449, DOI: [10.1002/ente.201901449](https://doi.org/10.1002/ente.201901449).
- 20 M. Mammeri, H. Bencherif, L. Dehimi, A. Hajri, P. Sasikumar, A. Syed and H. A. AL-Shwaiman, Stability Forecasting of Perovskite Solar Cells Utilizing Various Machine Learning and Deep Learning Techniques, *J. Opt.*, 2025, **54**, 930–938, DOI: [10.1007/s12596-024-01819-9](https://doi.org/10.1007/s12596-024-01819-9).
 - 21 A Stand out Family, *Nat. Mater.*, 2021, **20**(10), 1303, DOI: [10.1038/s41563-021-01127-8](https://doi.org/10.1038/s41563-021-01127-8).
 - 22 “Rose, Gustav: De novis quibusdam fossilibus quae in montibus Uraliis inveniuntur”, Image 1 of 12 | MDZ. <https://www.digitale-sammlungen.de/en/view/bsb10978561?page=1> (accessed 2025-04-23).
 - 23 C. N. Berglund and W. S. Baer, Electron Transport in Single-Domain, Ferroelectric Barium Titanate, *Phys. Rev.*, 1967, **157**(2), 358–366, DOI: [10.1103/PhysRev.157.358](https://doi.org/10.1103/PhysRev.157.358).
 - 24 H. Arend, W. Huber, F. H. Mischgofsky and G. K. Richter-Van Leeuwen, Layer Perovskites of the $(\text{CnH}_{2n} + 1\text{NH}_3)_2\text{MX}_4$ and $\text{NH}_3(\text{CH}_2)_m\text{NH}_3\text{MX}_4$ Families with $\text{M} = \text{Cd}, \text{Cu}, \text{Fe}, \text{Mn}$ OR Pd and $\text{X} = \text{Cl}$ OR Br : Importance, Solubilities and Simple Growth Techniques, *J. Cryst. Growth*, 1978, **43**(2), 213–223, DOI: [10.1016/0022-0248\(78\)90170-7](https://doi.org/10.1016/0022-0248(78)90170-7).
 - 25 M. M. Lee, J. Teuscher, T. Miyasaka, T. N. Murakami and H. J. Snaith, Efficient Hybrid Solar Cells Based on Meso-Superstructured Organometal Halide Perovskites, *Science*, 2012, **338**(6107), 643–647, DOI: [10.1126/science.1228604](https://doi.org/10.1126/science.1228604).
 - 26 H.-S. Kim, C.-R. Lee, J.-H. Im, K.-B. Lee, T. Moehl, A. Marchioro, S.-J. Moon, R. Humphry-Baker, J.-H. Yum, J. E. Moser, M. Grätzel and N.-G. Park, Lead Iodide Perovskite Sensitized All-Solid-State Submicron Thin Film Mesoscopic Solar Cell with Efficiency Exceeding 9%, *Sci. Rep.*, 2012, **2**(1), 591, DOI: [10.1038/srep00591](https://doi.org/10.1038/srep00591).
 - 27 Best Research-Cell Efficiency Chart. <https://www.nrel.gov/pv/cell-efficiency.html> (accessed 2024-08-13).
 - 28 Z.-K. Tan, R. S. Moghaddam, M. L. Lai, P. Docampo, R. Higler, F. Deschler, M. Price, A. Sadhanala, L. M. Pazos, D. Credgington, F. Hanusch, T. Bein, H. J. Snaith and R. H. Friend, Bright Light-Emitting Diodes Based on Organometal Halide Perovskite, *Nat. Nanotech.*, 2014, **9**(9), 687–692, DOI: [10.1038/nnano.2014.149](https://doi.org/10.1038/nnano.2014.149).
 - 29 P. Jowett, Oxford PV starts commercial distribution of perovskite solar modules. *pv magazine International*. <https://www.pv-magazine.com/2024/09/05/oxford-pv-starts-commercial-distribution-of-perovskite-solar-modules/>.
 - 30 J. T. Jacobo, Tandem PV raises US\$50 million to commercialise US-made perovskite modules. *PV Tech*. <https://www.pv-tech.org/tandem-pv-raises-us50-million-to-commercialise-us-made-perovskite-modules/> (accessed 2025-05-12).
 - 31 C. U. L. Shanghai, Showcased its first full perovskite P. module at this week's S. P. trade fair in; Willuhn, underscoring the technology's ongoing shift toward commercialization M. Commercial perovskite solar modules at SNEC 2024 trade show. *pv magazine International*. <https://www.pv-magazine.com/2024/06/13/commercial-perovskite-solar-modules-at-snec-2024-trade-show/> (accessed 2025-05-12).
 - 32 China's First Commercial Four-Terminal Perovskite-Silicon Tandem Modules Delivered for 50 MW Project - *EnergyTrend*. <https://www.energystrend.com/news/20241218-48889.html> (accessed 2025-05-12).
 - 33 D. H. Rai, Artificial Intelligence Through Time: A Comprehensive Historical Review, 2024, DOI: [10.13140/RG.2.2.22835.03364](https://doi.org/10.13140/RG.2.2.22835.03364).
 - 34 Y. Xu, X. Liu, X. Cao, C. Huang, E. Liu, S. Qian, X. Liu, Y. Wu, F. Dong, C.-W. Qiu, J. Qiu, K. Hua, W. Su, J. Wu, H. Xu, Y. Han, C. Fu, Z. Yin, M. Liu, R. Roepman, S. Dietmann, M. Virta, F. Kengara, Z. Zhang, L. Zhang, T. Zhao, J. Dai, J. Yang, L. Lan, M. Luo, Z. Liu, T. An, B. Zhang, X. He, S. Cong, X. Liu, W. Zhang, J. P. Lewis, J. M. Tiedje, Q. Wang, Z. An, F. Wang, L. Zhang, T. Huang, C. Lu, Z. Cai, F. Wang and J. Zhang, Artificial Intelligence: A Powerful Paradigm for Scientific Research, *Innovation*, 2021, **2**(4), 100179, DOI: [10.1016/j.xinn.2021.100179](https://doi.org/10.1016/j.xinn.2021.100179).
 - 35 J. A. Nichols, H. W. Herbert Chan and M. A. B. Baker, Machine Learning: Applications of Artificial Intelligence to Imaging and Diagnosis, *Biophys. Rev.*, 2018, **11**(1), 111–118, DOI: [10.1007/s12551-018-0449-9](https://doi.org/10.1007/s12551-018-0449-9).
 - 36 Stanford's robotics legacy. <https://news.stanford.edu/stories/2019/01/stanfords-robotics-legacy> (accessed 2025-05-12).
 - 37 G. E. Moore, Cramming More Components onto Integrated Circuits, Reprinted from *Electronics*, Volume 38, Number 8, April 19, 1965, Pp.114 Ff. *IEEE Solid-State Circuits Society Newsletter*, 2006, **11** (3), pp. 33–35, DOI: [10.1109/NSSC.2006.4785860](https://doi.org/10.1109/NSSC.2006.4785860).
 - 38 Deep Blue, IBM. <https://www.ibm.com/history/deep-blue> (accessed 2025-05-12).
 - 39 D. Steinkraus, I. Buck and P. Y. Simard, Using GPUs for machine learning algorithms, *Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*, Seoul, Korea (South), 2005, vol. 2, pp. 1115–1120, DOI: [10.1109/ICDAR.2005.251](https://doi.org/10.1109/ICDAR.2005.251).
 - 40 Introducing ChatGPT. <https://openai.com/index/chatgpt/> (accessed 2025-05-12).
 - 41 S. Lu, Q. Zhou, Y. Ouyang, Y. Guo, Q. Li and J. Wang, Accelerated Discovery of Stable Lead-Free Hybrid Organic-Inorganic Perovskites via Machine Learning, *Nat. Commun.*, 2018, **9**(1), 3405, DOI: [10.1038/s41467-018-05761-w](https://doi.org/10.1038/s41467-018-05761-w).
 - 42 J.-P. Correa-Baena, K. Hippalgaonkar, J. van Duren, S. Jaffer, V. R. Chandrasekhar, V. Stevanovic, C. Wadia, S. Guha and T. Buonassisi, Accelerating Materials Development via Automation, Machine Learning, and High-Performance Computing, *Joule*, 2018, **2**(8), 1410–1420, DOI: [10.1016/j.joule.2018.05.009](https://doi.org/10.1016/j.joule.2018.05.009).
 - 43 O. Allam, C. Holmes, Z. Greenberg, K. C. Kim and S. S. Jang, Density Functional Theory – Machine Learning Approach to Analyze the Bandgap of Elemental Halide Perovskites and Ruddlesden-Popper Phases, *Chem-PhysChem*, 2018, **19**(19), 2559–2565, DOI: [10.1002/cphc.201800382](https://doi.org/10.1002/cphc.201800382).
 - 44 S. Chen, Y. Hou, H. Chen, X. Tang, S. Langner, N. Li, T. Stubhan, I. Levchuk, E. Gu, A. Osvet and C. J. Brabec, Exploring the Stability of Novel Wide Bandgap Perovskites

- by a Robot Based High Throughput Approach, *Adv. Energy Mater.*, 2018, **8**(6), 1701543, DOI: [10.1002/aenm.201701543](https://doi.org/10.1002/aenm.201701543).
- 45 D. C. Nguyen, T. Asada, I. Raifuku and Y. Ishikawa, Analysis and Selection of Optimal Perovskite/Silicon Tandem Configuration for Building Integrated Photovoltaics Based on Their Annual Outdoor Energy Yield Predicted by Machine Learning, *Sol. RRL*, 2024, **8**(9), 2400072, DOI: [10.1002/solr.202400072](https://doi.org/10.1002/solr.202400072).
 - 46 Z.-Z. Li, C. Guo, W. Lv, P. Huang and Y. Zhang, Machine Learning-Enabled Optical Architecture Design of Perovskite Solar Cells, *J. Phys. Chem. Lett.*, 2024, **15**(14), 3835–3842, DOI: [10.1021/acs.jpcclett.4c00320](https://doi.org/10.1021/acs.jpcclett.4c00320).
 - 47 R. Luo, X. Jia, X. Niu, S. Liu, X. Guo, J. Li, Z.-J. Zhao, Y. Hou and J. Gong, Machine Learning-Driven Insights for Phase-Stable $\text{FAXCs1-xPb(IyBr1-y)}_3$ Perovskites in Tandem Solar Cells, *JACS Au*, 2025, **5**(4), 1771–1780, DOI: [10.1021/jacsau.5c00033](https://doi.org/10.1021/jacsau.5c00033).
 - 48 J. Xu, K. Li, U. N. Huynh, M. Fadel, J. Huang, R. Sundararaman, V. Vardeny and Y. Ping, How Spin Relaxes and Dephases in Bulk Halide Perovskites, *Nat. Commun.*, 2024, **15**(1), 188, DOI: [10.1038/s41467-023-42835-w](https://doi.org/10.1038/s41467-023-42835-w).
 - 49 DOE Releases New Report Evaluating Increase in Electricity Demand from Data Centers. *Energy.gov*. <https://www.energy.gov/articles/doe-releases-new-report-evaluating-increase-electricity-demand-data-centers> (accessed 2025-04-25).
 - 50 Z. Li, M. Yang, J.-S. Park, S.-H. Wei, J. J. Berry and K. Zhu, Stabilizing Perovskite Structures by Tuning Tolerance Factor: Formation of Formamidinium and Cesium Lead Iodide Solid-State Alloys, *Chem. Mater.*, 2016, **28**(1), 284–292, DOI: [10.1021/acs.chemmater.5b04107](https://doi.org/10.1021/acs.chemmater.5b04107).
 - 51 X. Li, S. Aftab, S. Hussain, F. Kabir, A. M. A. Henaish, A. G. Al-Sehemi, M. R. Pallavolu and G. Koyyada, Dimensional Diversity (0D, 1D, 2D, and 3D) in Perovskite Solar Cells: Exploring the Potential of Mixed-Dimensional Integrations, *J. Mater. Chem. A*, 2024, **12**(8), 4421–4440, DOI: [10.1039/D3TA06953B](https://doi.org/10.1039/D3TA06953B).
 - 52 R. Lyu, C. E. Moore, T. Liu, Y. Yu and Y. Wu, Predictive Design Model for Low-Dimensional Organic-Inorganic Halide Perovskites Assisted by Machine Learning, *J. Am. Chem. Soc.*, 2021, **143**(32), 12766–12776, DOI: [10.1021/jacs.1c05441](https://doi.org/10.1021/jacs.1c05441).
 - 53 X. Zhao, H. Xu, Z. Wang, Y. Lin and Y. Liu, Memristors with Organic-Inorganic Halide Perovskites, *InfoMat*, 2019, **1**(2), 183–210, DOI: [10.1002/inf2.12012](https://doi.org/10.1002/inf2.12012).
 - 54 C. C. Boyd, R. Cheacharoen, T. Leijtens and M. D. McGehee, Understanding Degradation Mechanisms and Improving Stability of Perovskite Photovoltaics, *Chem. Rev.*, 2018, **119**, DOI: [10.1021/acs.chemrev.8b00336](https://doi.org/10.1021/acs.chemrev.8b00336).
 - 55 T. Chen, J. Xie, B. Wen, Q. Yin, R. Lin, S. Zhu and P. Gao, Inhibition of Defect-Induced α -to- δ Phase Transition for Efficient and Stable Formamidinium Perovskite Solar Cells, *Nat. Commun.*, 2023, **14**(1), 6125, DOI: [10.1038/s41467-023-41853-y](https://doi.org/10.1038/s41467-023-41853-y).
 - 56 H. Liu, N. Li, Z. Chen, S. Tao, C. Li, L. Jiang, X. Niu, Q. Chen, F. Wang, Y. Zhang, Z. Huang, T. Song and H. Zhou, Reversible Phase Transition for Durable Formamidinium-Dominated Perovskite Photovoltaics, *Adv. Mater.*, 2022, **34**(39), 2204458, DOI: [10.1002/adma.202204458](https://doi.org/10.1002/adma.202204458).
 - 57 Q. Guo, J. Dou, Y. Mei, Y. Peng, F. Liu, Y. Liu, Q. Chen, Y. Zhao, Y. Wang, X. Zhang, J. Duan and Q. Tang, Toward α -Phase Stabilization of Formamidinium Lead Iodide Perovskites with Dual-Trivalent Metal Regulation, *Chem. Eng. J.*, 2025, **504**, 158875, DOI: [10.1016/j.cej.2024.158875](https://doi.org/10.1016/j.cej.2024.158875).
 - 58 Q. Zhang, J. Duan, Q. Guo, J. Zhang, D. Zheng, F. Yi, X. Yang, Y. Duan and Q. Tang, Thermal-Triggered Dynamic Disulfide Bond Self-Heals Inorganic Perovskite Solar Cells, *Angew. Chem., Int. Ed.*, 2022, **61**(8), e202116632, DOI: [10.1002/anie.202116632](https://doi.org/10.1002/anie.202116632).
 - 59 Q. Guo, J. Duan, J. Zhang, Q. Zhang, Y. Duan, X. Yang, B. He, Y. Zhao and Q. Tang, Universal Dynamic Liquid Interface for Healing Perovskite Solar Cells, *Adv. Mater.*, 2022, **34**(26), 2202301, DOI: [10.1002/adma.202202301](https://doi.org/10.1002/adma.202202301).
 - 60 X. Xiao, W. Li, Y. Fang, Y. Liu, Y. Shao, S. Yang, J. Zhao, X. Dai, R. Zia and J. Huang, Benign Ferroelastic Twin Boundaries in Halide Perovskites for Charge Carrier Transport and Recombination, *Nat. Commun.*, 2020, **11**(1), 2215, DOI: [10.1038/s41467-020-16075-1](https://doi.org/10.1038/s41467-020-16075-1).
 - 61 D. Bryant, Light and Oxygen Induced Degradation Limits the Operational Stability of Methylammonium Lead Triiodide Perovskite Solar Cells, *Energy Environ. Sci.*, 2016, **9**, DOI: [10.1039/C6EE00409A](https://doi.org/10.1039/C6EE00409A).
 - 62 B. Chen, P. N. Rudd, S. Yang, Y. Yuan and J. Huang, Imperfections and Their Passivation in Halide Perovskite Solar Cells, *Chem. Soc. Rev.*, 2019, **48**(14), 3842–3867, DOI: [10.1039/C8CS00853A](https://doi.org/10.1039/C8CS00853A).
 - 63 L. Ma, D. Guo, M. Li, C. Wang, Z. Zhou, X. Zhao, F. Zhang, Z. Ao and Z. Nie, Temperature-Dependent Thermal Decomposition Pathway of Organic-Inorganic Halide Perovskite Materials, *Chem. Mater.*, 2019, **31**(20), 8515–8522, DOI: [10.1021/acs.chemmater.9b03190](https://doi.org/10.1021/acs.chemmater.9b03190).
 - 64 H. Zai, Y. Ma, Q. Chen and H. Zhou, Ion Migration in Halide Perovskite Solar Cells: Mechanism, Characterization, Impact and Suppression, *J. Energy Chem.*, 2021, **63**, 528–549, DOI: [10.1016/j.jechem.2021.08.006](https://doi.org/10.1016/j.jechem.2021.08.006).
 - 65 N. Phung, A. Al-Ashouri, S. Meloni, A. Mattoni, S. Albrecht, E. L. Unger, A. Merdasa and A. Abate, The Role of Grain Boundaries on Ionic Defect Migration in Metal Halide Perovskites, *Adv. Energy Mater.*, 2020, **10**(20), 1903735, DOI: [10.1002/aenm.201903735](https://doi.org/10.1002/aenm.201903735).
 - 66 Z. Li, J. Dong, C. Liu, J. Guo, L. Shen and W. Guo, Surface Passivation of Perovskite Solar Cells Toward Improved Efficiency and Stability, *Nano-Micro Lett.*, 2019, **11**(1), 50, DOI: [10.1007/s40820-019-0282-0](https://doi.org/10.1007/s40820-019-0282-0).
 - 67 D. Meggiolaro, S. G. Motti, E. Mosconi, A. J. Barker, J. Ball, C. Andrea Riccardo Perini, F. Deschler, A. Petrozza and F. De Angelis, Iodine Chemistry Determines the Defect Tolerance of Lead-Halide Perovskites, *Energy Environ. Sci.*, 2018, **11**(3), 702–713, DOI: [10.1039/C8EE00124C](https://doi.org/10.1039/C8EE00124C).
 - 68 P. Kopperschmidt, S. Senz, G. Kästner, D. Hesse and U. M. Gösele, Materials Integration of Gallium Arsenide and Silicon by Wafer Bonding, *Appl. Phys. Lett.*, 1998, **72**(24), 3181–3183, DOI: [10.1063/1.121586](https://doi.org/10.1063/1.121586).
 - 69 J. Li, J. Dagar, O. Shargaieva, O. Maus, M. Remec, Q. Emery, M. Khenkin, C. Ulbrich, F. Akhundova, J. A. Márquez,

- T. Unold, M. Fenske, C. Schultz, B. Stegemann, A. Al-Ashouri, S. Albrecht, A. T. Esteves, L. Korte, H. Köbler, A. Abate, D. M. Többsen, I. Zizak, E. J. W. List-Kratochvil, R. Schlattmann and E. Unger, Ink Design Enabling Slot-Die Coated Perovskite Solar Cells with >22% Power Conversion Efficiency, Micro-Modules, and 1 Year of Outdoor Performance Evaluation, *Adv. Energy Mater.*, 2023, 2203898, DOI: [10.1002/aenm.202203898](https://doi.org/10.1002/aenm.202203898).
- 70 D. O. Baumann, F. Laufer, J. Roger, R. Singh, M. Gholipour and U. W. Paetzold, Repeatable Perovskite Solar Cells through Fully Automated Spin-Coating and Quenching, *ACS Appl. Mater. Interfaces*, 2024, 16(40), 54007–54016, DOI: [10.1021/acsami.4c13024](https://doi.org/10.1021/acsami.4c13024).
- 71 K. Prakash, N. J. Valeti, P. Jain, C. S. Pathak, M. K. Singha, S. Gupta, E. Edri and S. Mukhopadhyay, Single-Crystal Perovskite Halide: Crystal Growth to Devices Applications, *Energy Technol.*, 2025, 13(7), 2400618, DOI: [10.1002/ente.202400618](https://doi.org/10.1002/ente.202400618).
- 72 H.-H. Fang, S. Adjokatse, H. Wei, J. Yang, G. R. Blake, J. Huang, J. Even and M. A. Loi, Ultrahigh Sensitivity of Methylammonium Lead Tribromide Perovskite Single Crystals to Environmental Gases, *Sci. Adv.*, 2016, 2(7), e1600534, DOI: [10.1126/sciadv.1600534](https://doi.org/10.1126/sciadv.1600534).
- 73 M. Noack and D. Ushizima, *Methods and Applications of Autonomous Experimentation*, Chapman and Hall/CRC, New York, 1st edn, 2023. , DOI: [10.1201/9781003359593](https://doi.org/10.1201/9781003359593).
- 74 G. Tom, S. P. Schmid, S. G. Baird, Y. Cao, K. Darvish, H. Hao, S. Lo, S. Pablo-García, E. M. Rajaonson, M. Skreta, N. Yoshikawa, S. Corapi, G. D. Akkoc, F. Strieth-Kalthoff, M. Seifrid and A. Aspuru-Guzik, Self-Driving Laboratories for Chemistry and Materials Science, *Chem. Rev.*, 2024, 124(16), 9633–9732, DOI: [10.1021/acs.chemrev.4c00055](https://doi.org/10.1021/acs.chemrev.4c00055).
- 75 F. Delgado-Licona and M. Abolhasani, Research Acceleration in Self-Driving Labs: Technological Roadmap toward Accelerated Materials and Molecular Discovery, *Adv. Intell. Syst.*, 2023, 5(4), 2200331, DOI: [10.1002/aisy.202200331](https://doi.org/10.1002/aisy.202200331).
- 76 A. G. Kusne, H. Yu, C. Wu, H. Zhang, J. Hatrick-Simpers, B. DeCost, S. Sarker, C. Oses, C. Toher, S. Curtarolo, A. V. Davydov, R. Agarwal, L. A. Bendersky, M. Li, A. Mehta and I. Takeuchi, On-the-Fly Closed-Loop Materials Discovery via Bayesian Active Learning, *Nat. Commun.*, 2020, 11(1), 5966, DOI: [10.1038/s41467-020-19597-w](https://doi.org/10.1038/s41467-020-19597-w).
- 77 J. Zhang, J. Wu, V. M. Le Corre, J. A. Hauch, Y. Zhao and C. J. Brabec, Advancing Perovskite Photovoltaic Technology through Machine Learning-Driven Automation, *Info-Mat*, 2025, 7(5), e70005, DOI: [10.1002/inf2.70005](https://doi.org/10.1002/inf2.70005).
- 78 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, Optimizing Perovskite Thin-Film Parameter Spaces with Machine Learning-Guided Robotic Platform for High-Performance Perovskite Solar Cells, *Adv. Energy Mater.*, 2023, 13(48), 2302594, DOI: [10.1002/aenm.202302594](https://doi.org/10.1002/aenm.202302594).
- 79 A. Halder, M. B. Alghalayini, S. Cheng, N. Thalanki, T. M. Nguyen, A. R. Hering, D.-K. Lee, S. Arnold, M. S. Leite, E. Barnard, A. Razumtcev, M. Wall, A. Gashi, Y.-R. Liu, M. M. Noack, S. Sun and C. M. Sutter-Fella, AI-Driven Robot Enables Synthesis-Property Relation Prediction for Metal Halide Perovskites in Humid Atmosphere, *Adv. Energy Mater.*, 2025, 15(34), 2502294, DOI: [10.1002/aenm.202502294](https://doi.org/10.1002/aenm.202502294).
- 80 C. Wang, Y.-J. Kim, A. Vriza, R. Batra, A. Baskaran, N. Shan, N. Li, P. Darancet, L. Ward, Y. Liu, M. K. Y. Chan, S. K. R. S. Sankaranarayanan, H. C. Fry, C. S. Miller, H. Chan and J. Xu, Autonomous Platform for Solution Processing of Electronic Polymers, *Nat. Commun.*, 2025, 16(1), 1498, DOI: [10.1038/s41467-024-55655-3](https://doi.org/10.1038/s41467-024-55655-3).
- 81 B. P. MacLeod, F. G. L. Parlange, 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, Self-Driving Laboratory for Accelerated Discovery of Thin-Film Materials, *Sci. Adv.*, 2020, 6(20), eaaz8867, DOI: [10.1126/sciadv.aaz8867](https://doi.org/10.1126/sciadv.aaz8867).
- 82 N. J. Szymanski, B. Rendy, Y. Fei, R. E. Kumar, T. He, D. Milsted, M. J. McDermott, M. Gallant, E. D. Cubuk, A. Merchant, H. Kim, A. Jain, C. J. Bartel, K. Persson, Y. Zeng and G. Ceder, An Autonomous Laboratory for the Accelerated Synthesis of Novel Materials, *Nature*, 2023, 624(7990), 86–91, DOI: [10.1038/s41586-023-06734-w](https://doi.org/10.1038/s41586-023-06734-w).
- 83 R. W. Epps, M. S. Bowen, A. A. Volk, K. Abdel-Latif, S. Han, K. G. Reyes, A. Amassian and M. Abolhasani, Artificial Chemist: An Autonomous Quantum Dot Synthesis Bot, *Adv. Mater.*, 2020, 32(30), 2001626, DOI: [10.1002/adma.202001626](https://doi.org/10.1002/adma.202001626).
- 84 S. P. Stier, C. Kreisbeck, H. Ihssen, M. A. Popp, J. Hauch, K. Malek, M. Reynaud, T. P. M. Goumans, J. Carlsson, I. Todorov, L. Gold, A. Räder, W. Wenzel, S. T. Bandesha, P. Jacques, F. Garcia-Moreno, O. Arcelus, P. Friederich, S. Clark, M. Maglione, A. Laukkanen, I. E. Castelli, J. Carrasco, M. C. Cabanas, H. S. Stein, O. Ozcan, D. Elbert, K. Reuter, C. Scheurer, M. Demura, S. S. Han, T. Vegge, S. Nakamae, M. Fabrizio and M. Kozdras, Materials Acceleration Platforms (MAPs): Accelerating Materials Research and Development to Meet Urgent Societal Challenges, *Adv. Mater.*, 2024, 36(45), 2407791, DOI: [10.1002/adma.202407791](https://doi.org/10.1002/adma.202407791).
- 85 S. Sadeghi, F. Bateni, T. Kim, D. Y. Son, J. A. Bennett, N. Orouji, V. S. Punati, C. Stark, T. D. Cerra, R. Awad, F. Delgado-Licona, J. Xu, N. Mukhin, H. Dickerson, K. G. Reyes and M. Abolhasani, Autonomous Nanomanufacturing of Lead-Free Metal Halide Perovskite Nanocrystals Using a Self-Driving Fluidic Lab, *Nanoscale*, 2024, 16(2), 580–591, DOI: [10.1039/D3NR05034C](https://doi.org/10.1039/D3NR05034C).
- 86 E. Stach, B. DeCost, A. G. Kusne, J. Hatrick-Simpers, K. A. Brown, K. G. Reyes, J. Schrier, S. Billinge, T. Buonassisi, I. Foster, C. P. Gomes, J. M. Gregoire, A. Mehta, J. Montoya, E. Olivetti, C. Park, E. Rotenberg, S. K. Saikin, S. Smullin, V. Stanev and B. Maruyama, Autonomous Experimentation Systems for Materials

- Development: A Community Perspective, *Matter*, 2021, 4(9), 2702–2726, DOI: [10.1016/j.matt.2021.06.036](https://doi.org/10.1016/j.matt.2021.06.036).
- 87 H. Zhao, W. Chen, H. Huang, Z. Sun, Z. Chen, L. Wu, B. Zhang, F. Lai, Z. Wang, M. L. Adam, C. H. Pang, P. K. Chu, Y. Lu, T. Wu, J. Jiang, Z. Yin and X.-F. Yu, A Robotic Platform for the Synthesis of Colloidal Nanocrystals, *Nat. Synth.*, 2023, 2(6), 505–514, DOI: [10.1038/s44160-023-00250-5](https://doi.org/10.1038/s44160-023-00250-5).
 - 88 F. Häse, L. M. Roch and A. Aspuru-Guzik, Next-Generation Experimentation with Self-Driving Laboratories, *Trends Chem.*, 2019, 1(3), 282–291, DOI: [10.1016/j.trechm.2019.02.007](https://doi.org/10.1016/j.trechm.2019.02.007).
 - 89 B. Sanchez-Lengeling and A. Aspuru-Guzik, Inverse Molecular Design Using Machine Learning: Generative Models for Matter Engineering, *Science*, 2018, 361(6400), 360–365, DOI: [10.1126/science.aat2663](https://doi.org/10.1126/science.aat2663).
 - 90 Y. Peng, Q. Guo, Y. Liu, Q. Chen, W. Lang, Y. Yang, J. Dou, Y. Wang, X. Zhang, J. Duan, Y. Zhao, X. Yang, W. Chen and Q. Tang, Dual-Intermediator Guided Methodical Molecular Exchange towards Optimized Crystallization Kinetics of Advanced Perovskite Solar Cells, *Nano Energy*, 2025, 142, 111169, DOI: [10.1016/j.nanoen.2025.111169](https://doi.org/10.1016/j.nanoen.2025.111169).
 - 91 C. Zhou, W. Meng, L. Kong, C. Zhang, J. Zhang, F. Liu, H. Li, G. Jia and X. Yang, Vacuum Processed Metal Halide Perovskite Light-Emitting Diodes, *Adv. Funct. Mater.*, 2024, 34(8), 2307682, DOI: [10.1002/adfm.202307682](https://doi.org/10.1002/adfm.202307682).
 - 92 High-Throughput Studies. College of Computing and Informatics. <https://cci.charlotte.edu/research/bioinformatics-research/high-throughput-studies/> (accessed 2025-05-30).
 - 93 Monitoring the stability and degradation mechanisms of perovskite solar cells by *in situ* and operando characterization, *APL Energy*, AIP Publishing. <https://pubs.aip.org/aip/aep/article/1/1/011501/2884924/Monitoring-the-stability-and-degradation> (accessed 2025-05-02).
 - 94 A. R. Hering; M. Dubey; E. Hosseini; M. Srivastava; Y. An; J.-P. Correa-Baena; H. Homayoun and M. S. Leite, *Machine Learning Reveals Composition Dependent Thermal Stability in Halide Perovskites*, *arXiv*, preprint, arXiv.2504.04002, 2025, DOI: [10.48550/arXiv.2504.04002](https://doi.org/10.48550/arXiv.2504.04002).
 - 95 E. Foadian, J. Yang, S. B. Harris, Y. Tang, C. M. Rouleau, S. Joy, K. R. Graham, B. J. Lawrie, B. Hu and M. Ahmadi, Decoding the Broadband Emission of 2D Pb-Sn Halide Perovskites through High-Throughput Exploration, *Adv. Funct. Mater.*, 2024, 34(52), 2411164, DOI: [10.1002/adfm.202411164](https://doi.org/10.1002/adfm.202411164).
 - 96 S. Moradi, S. Kundu, M. Rezazadeh, V. Yeddu, O. Voznyy and M. I. Saidaminov, High-Throughput Exploration of Halide Perovskite Compositionally-Graded Films and Degradation Mechanisms, *Commun. Mater.*, 2022, 3(1), 1–5, DOI: [10.1038/s43246-022-00235-5](https://doi.org/10.1038/s43246-022-00235-5).
 - 97 Y. Shang, W. Meng, F. Lu, H. Li, Z. Yuan and N. Li, A High-Throughput Approach to Identifying Environment-Friendly Artificial Antisolvents for Efficient Perovskite Solar Cells, *Adv. Mater.*, 2025, 37, 2504602, DOI: [10.1002/adma.202504602](https://doi.org/10.1002/adma.202504602).
 - 98 J. Yang, J. Hidalgo, D. Song, S. V. Kalinin, J.-P. Correa-Baena and M. Ahmadi, Accelerating Materials Discovery by High-Throughput GIWAXS Characterization of Quasi-2D Formamidinium Metal Halide Perovskites, *Adv. Funct. Mater.*, 2024, 34(49), 2409293, DOI: [10.1002/adfm.202409293](https://doi.org/10.1002/adfm.202409293).
 - 99 L. Tan, J. Zhou, X. Zhao, S. Wang, M. Li, C. Jiang, H. Li, Y. Zhang, Y. Ye, W. Tress, L. Ding, M. Grätzel and C. Yi, Combined Vacuum Evaporation and Solution Process for High-Efficiency Large-Area Perovskite Solar Cells with Exceptional Reproducibility, *Adv. Mater.*, 2023, 35(13), 2205027, DOI: [10.1002/adma.202205027](https://doi.org/10.1002/adma.202205027).
 - 100 E. Gu, X. Tang, S. Langner, P. Duchstein, Y. Zhao, I. Levchuk, V. Kalancha, T. Stubhan, J. Hauch, H. J. Egelhaaf, D. Zahn, A. Osvet and C. J. Brabec, Robot-Based High-Throughput Screening of Antisolvents for Lead Halide Perovskites, *Joule*, 2020, 4(8), 1806–1822, DOI: [10.1016/j.joule.2020.06.013](https://doi.org/10.1016/j.joule.2020.06.013).
 - 101 M. Konstantakou, D. Perganti, P. Falaras and T. Stergiopoulos, Anti-Solvent Crystallization Strategies for Highly Efficient Perovskite Solar Cells, *Crystals*, 2017, 7(10), 291, DOI: [10.3390/cryst7100291](https://doi.org/10.3390/cryst7100291).
 - 102 S. V. Kalinin, D. Mukherjee, K. Roccapriore, B. J. Blaiszik, A. Ghosh, M. A. Ziatdinov, A. Al-Najjar, C. Doty, S. Akers, N. S. Rao, J. C. Agar and S. R. Spurgeon, Machine Learning for Automated Experimentation in Scanning Transmission Electron Microscopy, *npj Comput. Mater.*, 2023, 9(1), 1–16, DOI: [10.1038/s41524-023-01142-0](https://doi.org/10.1038/s41524-023-01142-0).
 - 103 P. Ercius, I. J. Johnson, P. Pelz, B. H. Savitzky, L. Hughes, H. G. Brown, S. E. Zeltmann, S.-L. Hsu, C. C. S. Pedroso, B. E. Cohen, R. Ramesh, D. Paul, J. M. Joseph, T. Stezelberger, C. Czarnik, M. Lent, E. Fong, J. Ciston, M. C. Scott, C. Ophus, A. M. Minor and P. Denes, The 4D Camera: An 87 kHz Direct Electron Detector for Scanning/Transmission Electron Microscopy, *Microsc. Microanal.*, 2024, 30(5), 903–912, DOI: [10.1093/mam/ozae086](https://doi.org/10.1093/mam/ozae086).
 - 104 M. I. Jordan and T. M. Mitchell, Machine Learning: Trends, Perspectives, and Prospects, *Science*, 2015, 349(6245), 255–260, DOI: [10.1126/science.aaa8415](https://doi.org/10.1126/science.aaa8415).
 - 105 A. R. Hering, M. Dubey and M. S. Leite, Emerging Opportunities for Hybrid Perovskite Solar Cells Using Machine Learning, *APL Energy*, 2023, 1(2), 020901, DOI: [10.1063/5.0146828](https://doi.org/10.1063/5.0146828).
 - 106 E. Foadian, S. Sanchez, S. V. Kalinin and M. Ahmadi, From Sunlight to Solutions: Closing the Loop on Halide Perovskites, *ACS Mater. Au*, 2025, 5(1), 11–23, DOI: [10.1021/acsmaterialsau.4c00096](https://doi.org/10.1021/acsmaterialsau.4c00096).
 - 107 T. Jo, *Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning*, Springer International Publishing, Cham, 2021. , DOI: [10.1007/978-3-030-65900-4](https://doi.org/10.1007/978-3-030-65900-4).
 - 108 R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, The MIT Press, 2nd edn, 2015.
 - 109 R. C. Wilson, E. Bonawitz, V. D. Costa and R. B. Ebitz, Balancing Exploration and Exploitation with Information and Randomization, *Curr. Opin. Behav. Sci.*, 2021, 38, 49–56, DOI: [10.1016/j.cobeha.2020.10.001](https://doi.org/10.1016/j.cobeha.2020.10.001).
 - 110 N. A. Chaudhary, P. Jain, S. Thakor, V. A. Rana and A. N. Prajapati, Exploring Molecular Interactions and

- Dielectric Relaxation in N-Octanol/DMF Binary Mixtures: A Machine Learning-Enhanced VNA Study, *Spectrochim. Acta, Part A*, 2025, **339**, 126271, DOI: [10.1016/j.saa.2025.126271](https://doi.org/10.1016/j.saa.2025.126271).
- 111 K. P. Murphy, *Machine Learning - A Probabilistic Perspective; Adaptive Computation and Machine Learning*, MIT Press, Cambridge, 2014.
 - 112 M. Srivastava, A. R. Hering, Y. An, J.-P. Correa-Baena and M. S. Leite, Machine Learning Enables Prediction of Halide Perovskites' Optical Behavior with > 90% Accuracy, *ACS Energy Lett.*, 2023, **8**(4), 1716–1722, DOI: [10.1021/acsenerylett.2c02555](https://doi.org/10.1021/acsenerylett.2c02555).
 - 113 R. J. Stoddard, W. A. Dunlap-Shohl, H. Qiao, Y. Meng, W. F. Kau and H. W. Hillhouse, Forecasting the Decay of Hybrid Perovskite Performance Using Optical Transmittance or Reflected Dark-Field Imaging, *ACS Energy Lett.*, 2020, **5**(3), 946–954, DOI: [10.1021/acsenerylett.0c00164](https://doi.org/10.1021/acsenerylett.0c00164).
 - 114 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. Colmann, S. Cros, K. Domanski, M. Dusz, 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. (F.) 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. Yıldırım, S. M. Zakeeruddin, K. Zhu and M. Lira-Cantu, Consensus Statement for Stability Assessment and Reporting for Perovskite Photovoltaics Based on ISOS Procedures, *Nat. Energy*, 2020, **5**(1), 35–49, DOI: [10.1038/s41560-019-0529-5](https://doi.org/10.1038/s41560-019-0529-5).
 - 115 S. Wang, Y. Huang, W. Hu and L. Zhang, Data-Driven Optimization and Machine Learning Analysis of Compatible Molecules for Halide Perovskite Material, *npj Comput. Mater.*, 2024, **10**(1), 1–10, DOI: [10.1038/s41524-024-01297-4](https://doi.org/10.1038/s41524-024-01297-4).
 - 116 N. T. P. Hartono, H. Köbler, P. Graniero, M. Khenkin, R. Schlattmann, C. Ulbrich and A. Abate, Stability Follows Efficiency Based on the Analysis of a Large Perovskite Solar Cells Ageing Dataset, *Nat. Commun.*, 2023, **14**(1), 4869, DOI: [10.1038/s41467-023-40585-3](https://doi.org/10.1038/s41467-023-40585-3).
 - 117 G. H. Gu, J. Jang, J. Noh, A. Walsh and Y. Jung, Perovskite Synthesizability Using Graph Neural Networks, *npj Comput. Mater.*, 2022, **8**(1), 71, DOI: [10.1038/s41524-022-00757-z](https://doi.org/10.1038/s41524-022-00757-z).
 - 118 O. Bousquet; U. Luxburg and G. Rätsch, *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003, Tübingen, Germany, August 4-16, 2003, Revised Lectures; Lecture Notes in Computer Science*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.
 - 119 H. Huang, Y. Wang, C. Rudin and E. P. Browne, Towards a Comprehensive Evaluation of Dimension Reduction Methods for Transcriptomic Data Visualization, *Commun. Biol.*, 2022, **5**(1), 719, DOI: [10.1038/s42003-022-03628-x](https://doi.org/10.1038/s42003-022-03628-x).
 - 120 C. Karpovich, E. Pan and E. A. Olivetti, Deep Reinforcement Learning for Inverse Inorganic Materials Design, *npj Comput. Mater.*, 2024, **10**(1), 287, DOI: [10.1038/s41524-024-01474-5](https://doi.org/10.1038/s41524-024-01474-5).
 - 121 B. Zheng, Z. Zheng and G. X. Gu, Designing Mechanically Tough Graphene Oxide Materials Using Deep Reinforcement Learning, *npj Comput. Mater.*, 2022, **8**(1), 225, DOI: [10.1038/s41524-022-00919-z](https://doi.org/10.1038/s41524-022-00919-z).
 - 122 S. Yu, Z. Chen, W. Liao, C. Yuan, B. Shang and R. Hu, Enhancing Overall Performance of Thermophotovoltaics via Deep Reinforcement Learning-Based Optimization, *J. Appl. Phys.*, 2024, **136**(2), 023101, DOI: [10.1063/5.0213211](https://doi.org/10.1063/5.0213211).
 - 123 P. Leinen, M. Esders, K. T. Schütt, C. Wagner, K.-R. Müller and F. S. Tautz, Autonomous Robotic Nanofabrication with Reinforcement Learning, *Sci. Adv.*, 2020, **6**(36), eabb6987, DOI: [10.1126/sciadv.abb6987](https://doi.org/10.1126/sciadv.abb6987).
 - 124 P. Jain, S. Thakor, A. Joshi, K. V. Chauhan and C. R. Vaja, Machine Learning-Driven Analysis of Dielectric Response in Polymethyl Methacrylate Nanocomposites Reinforced with Multi-Walled Carbon Nanotubes, *J. Mater. Sci. Mater. Electron.*, 2024, **35**(20), 1419, DOI: [10.1007/s10854-024-13188-x](https://doi.org/10.1007/s10854-024-13188-x).
 - 125 O. Rainio, J. Teuho and R. Klén, Evaluation Metrics and Statistical Tests for Machine Learning, *Sci. Rep.*, 2024, **14**(1), 6086, DOI: [10.1038/s41598-024-56706-x](https://doi.org/10.1038/s41598-024-56706-x).
 - 126 T. J. Jacobsson, A. Hultqvist, A. García-Fernández, A. Anand, A. Al-Ashouri, A. Hagfeldt, A. Crovetto, A. Abate, A. G. Ricciardulli, A. Vijayan, A. Kulkarni, A. Y. Anderson, B. P. Darwich, B. Yang, B. L. Coles, C. A. R. Perini, C. Rehmann, D. Ramirez, D. Fairen-Jimenez, D. Di Girolamo, D. Jia, E. Avila, E. J. Juarez-Perez, F. Baumann, F. Mathies, G. S. A. González, G. Boschloo, G. Nasti, G. Paramasivam, G. Martínez-Denegri, H. Näsström, H. Michaels, H. Köbler, H. Wu, I. Benesperi, M. I. Dar, I. Bayrak Pehlivan, I. E. Gould, J. N. Vagott, J. Dagar, J. Kettle, J. Yang, J. Li, J. A. Smith, J. Pascual, J. J. Jerónimo-Rendón, J. F. Montoya, J.-P. Correa-Baena, J. Qiu, J. Wang, K. Sveinbjörnsson, K. Hirslandt, K. Dey, K. Frohna, L. Mathies, L. A. Castriotta, M. H. Aldamasy, M. Vasquez-Montoya, M. A. Ruiz-Preciado, M. A. Flatken, M. V. Khenkin, M. Grischek, M. Kedia, M. Saliba, M. Anaya, M. Veldhoen, N. Arora, O. Shargaieva, O. Maus, O. S. Game, O. Yudilevich, P. Fassel, Q. Zhou, R. Betancur, R. Munir, R. Patidar, S. D. Stranks, S. Alam, S. Kar, T. Unold, T. Abzieher, T. Edvinsson, T. W. David, U. W. Paetzold, W. Zia, W. Fu, W. Zuo, V. R. F. Schröder, W. Tress, X. Zhang, Y.-H. Chiang, Z. Iqbal, Z. Xie and E. Unger, An Open-Access Database and Analysis Tool for Perovskite Solar Cells Based on the FAIR Data Principles, *Nat. Energy*, 2021, **7**(1), 107–115, DOI: [10.1038/s41560-021-00941-3](https://doi.org/10.1038/s41560-021-00941-3).
 - 127 *The Perovskite Database*. <https://www.perovskitedatabase.com/> (accessed 2025-06-20).
 - 128 C. Deng, L. Tang, P. Luo, H. Li, L. Yang, Z. Liu, B. Liu, X. Lu, Y. Song, X. Sun and Y. Zhao, Unveiling the Statistical Behaviors of Metal-Halide Perovskites from Films to Devices through a High-Throughput Experimental Platform, *InfoMat*, 2025, e70039, DOI: [10.1002/inf2.70039](https://doi.org/10.1002/inf2.70039).

- 129 X. Li, Y. Che, L. Chen, T. Liu, K. Wang, L. Liu, H. Yang, E. O. Pyzer-Knapp and A. I. Cooper, Sequential Closed-Loop Bayesian Optimization as a Guide for Organic Molecular Metallophotocatalyst Formulation Discovery, *Nat. Chem.*, 2024, **16**(8), 1286–1294, DOI: [10.1038/s41557-024-01546-5](https://doi.org/10.1038/s41557-024-01546-5).
- 130 K. Li, A. N. Rubungo, X. Lei, D. Persaud, K. Choudhary, B. DeCost, A. B. Dieng and J. Hattrick-Simpers, Probing Out-of-Distribution Generalization in Machine Learning for Materials, *Commun. Mater.*, 2025, **6**(1), 9, DOI: [10.1038/s43246-024-00731-w](https://doi.org/10.1038/s43246-024-00731-w).
- 131 G. F. Gomes; B. Kouider and J. L. J. Pereira, Optimization and Artificial Intelligence: An in-Depth Analysis of Multi-Objective Optimization, Sampling Methods, and Regression Algorithms Applied to Structural Design, *Mechanics Based Design of Structures and Machines*, 2025, vol. 0 (0), pp. 1–28, DOI: [10.1080/15397734.2025.2476041](https://doi.org/10.1080/15397734.2025.2476041).
- 132 M. Christensen, L. P. E. Yunker, F. Adediji, F. Häse, L. M. Roch, T. Gensch, G. dos Passos Gomes, T. Zepel, M. S. Sigman, A. Aspuru-Guzik and J. E. Hein, Data-Science Driven Autonomous Process Optimization, *Commun. Chem.*, 2021, **4**(1), 1–12, DOI: [10.1038/s42004-021-00550-x](https://doi.org/10.1038/s42004-021-00550-x).
- 133 H. Cheng, Y. Liu, B. Cai, C. Häggglund, T. Kubart, G. Boschloo and H. Tian, Atomic Layer Deposition of SnO₂ as an Electron Transport Material for Solid-State P-Type Dye-Sensitized Solar Cells, *ACS Appl. Energy Mater.*, 2022, **5**(10), 12022–12028, DOI: [10.1021/acsaem.2c01328](https://doi.org/10.1021/acsaem.2c01328).
- 134 D. Koo, Y. Choi, U. Kim, J. Kim, J. Seo, E. Son, H. Min, J. Kang and H. Park, Mesoporous Structured MoS₂ as an Electron Transport Layer for Efficient and Stable Perovskite Solar Cells, *Nat. Nanotechnol.*, 2025, **20**(1), 75–82, DOI: [10.1038/s41565-024-01799-8](https://doi.org/10.1038/s41565-024-01799-8).
- 135 S. Hu, S. Zeng, X. Deng, P. Hou, H. Du, Y. Dou, W. Xiong, J. Pan, Y. Peng, Y.-B. Cheng and Z. Ku, Scalable Impregnation Method for Preparing a Self-Assembled Monolayer in High-Performance Vapor-Deposited Lead-Halide Perovskite Solar Cells, *ACS Nano*, 2025, **19**(15), 15018–15029, DOI: [10.1021/acsnano.5c01479](https://doi.org/10.1021/acsnano.5c01479).
- 136 J. Yang, G. Qu, Y. Qiao, S. Cai, J. Hu, S. Geng, Y. Li, Y. Jin, N. Shen, S. Chen, A. K.-Y. Jen and Z.-X. Xu, Flexibility Meets Rigidity: A Self-Assembled Monolayer Materials Strategy for Perovskite Solar Cells, *Nat. Commun.*, 2025, **16**(1), 6968, DOI: [10.1038/s41467-025-62388-4](https://doi.org/10.1038/s41467-025-62388-4).
- 137 Z. Shokrollahi, M. Piralaee and A. Asgari, Performance and Optimization Study of Selected 4-Terminal Tandem Solar Cells, *Sci. Rep.*, 2024, **14**(1), 11515, DOI: [10.1038/s41598-024-62085-0](https://doi.org/10.1038/s41598-024-62085-0).
- 138 T. Das, G. Di Liberto and G. Pacchioni, Density Functional Theory Estimate of Halide Perovskite Band Gap: When Spin Orbit Coupling Helps, *J. Phys. Chem. C*, 2022, **126**(4), 2184–2198, DOI: [10.1021/acs.jpcc.1c09594](https://doi.org/10.1021/acs.jpcc.1c09594).
- 139 J.-S. Kim, J. Noh and J. Im, Machine Learning-Enabled Chemical Space Exploration of All-Inorganic Perovskites for Photovoltaics, *npj Comput. Mater.*, 2024, **10**(1), 1–10, DOI: [10.1038/s41524-024-01270-1](https://doi.org/10.1038/s41524-024-01270-1).
- 140 S. Kim, J. A. Márquez, T. Unold and A. Walsh, Upper Limit to the Photovoltaic Efficiency of Imperfect Crystals from First Principles, *Energy Environ. Sci.*, 2020, **13**(5), 1481–1491, DOI: [10.1039/D0EE00291G](https://doi.org/10.1039/D0EE00291G).
- 141 J. Yang, P. Manganaris and A. Mannodi-Kanakkithodi, A High-Throughput Computational Dataset of Halide Perovskite Alloys, *Digital Discovery*, 2023, **2**(3), 856–870, DOI: [10.1039/D3DD00015J](https://doi.org/10.1039/D3DD00015J).
- 142 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, A Data Fusion Approach to Optimize Compositional Stability of Halide Perovskites, *Matter*, 2021, **4**(4), 1305–1322, DOI: [10.1016/j.matt.2021.01.008](https://doi.org/10.1016/j.matt.2021.01.008).
- 143 A. H. Slavney, B. A. Connor, L. Leppert and H. I. Karunadasa, A Pencil-and-Paper Method for Elucidating Halide Double Perovskite Band Structures, *Chem. Sci.*, 2019, **10**(48), 11041–11053, DOI: [10.1039/C9SC03219C](https://doi.org/10.1039/C9SC03219C).
- 144 D. Meggiolaro and F. De Angelis, First-Principles Modeling of Defects in Lead Halide Perovskites: Best Practices and Open Issues, *ACS Energy Lett.*, 2018, **3**(9), 2206–2222, DOI: [10.1021/acsenerylett.8b01212](https://doi.org/10.1021/acsenerylett.8b01212).
- 145 F. De Angelis, The Impact of Machine Learning in Energy Materials Research: The Case of Halide Perovskites, *ACS Energy Lett.*, 2023, **8**(2), 1270–1272, DOI: [10.1021/acsenerylett.3c00182](https://doi.org/10.1021/acsenerylett.3c00182).
- 146 D. Cahen, Y. Rakita, D. A. Egger and A. Kahn, Surface Defects Control Bulk Carrier Densities in Polycrystalline Pb-Halide Perovskites, *Adv. Mater.*, 2024, **36**(50), 2407098, DOI: [10.1002/adma.202407098](https://doi.org/10.1002/adma.202407098).
- 147 J. Xu, H. Chen, L. Grater, C. Liu, Y. Yang, S. Teale, A. Maxwell, S. Mahesh, H. Wan, Y. Chang, B. Chen, B. Rehl, S. M. Park, M. G. Kanatzidis and E. H. Sargent, Anion Optimization for Bifunctional Surface Passivation in Perovskite Solar Cells, *Nat. Mater.*, 2023, **22**(12), 1507–1514, DOI: [10.1038/s41563-023-01705-y](https://doi.org/10.1038/s41563-023-01705-y).
- 148 Y. Pan, J. Wang, Z. Sun, J. Zhang, Z. Zhou, C. Shi, S. Liu, F. Ren, R. Chen, Y. Cai, H. Sun, B. Liu, Z. Zhang, Z. Zhao, Z. Cai, X. Qin, Z. Zhao, Y. Ji, N. Li, W. Huang, Z. Liu and W. Chen, Surface Chemical Polishing and Passivation Minimize Non-Radiative Recombination for All-Perovskite Tandem Solar Cells, *Nat. Commun.*, 2024, **15**(1), 7335, DOI: [10.1038/s41467-024-51703-0](https://doi.org/10.1038/s41467-024-51703-0).
- 149 B. Yang, J. Suo, D. Bogachuk, W. Kaiser, C. Baretzky, O. Er-Raji, G. Loukeris, A. A. Allothman, E. Mosconi, M. Kohlstädt, U. Würfel, F. D. Angelis and A. Hagfeldt, A Universal Ligand for Lead Coordination and Tailored Crystal Growth in Perovskite Solar Cells, *Energy Environ. Sci.*, 2024, **17**(4), 1549–1558, DOI: [10.1039/D3EE02344C](https://doi.org/10.1039/D3EE02344C).
- 150 M. Noman, A. H. H. Khan and S. T. Jan, Interface Engineering and Defect Passivation for Enhanced Hole Extraction, Ion Migration, and Optimal Charge Dynamics in Both Lead-Based and Lead-Free Perovskite Solar Cells, *Sci. Rep.*, 2024, **14**(1), 5449, DOI: [10.1038/s41598-024-56246-4](https://doi.org/10.1038/s41598-024-56246-4).
- 151 Machine-Learning-Assisted Design of Buried-Interface Engineering Materials for High-Efficiency and Stable

- Perovskite Solar Cells, *ACS Energy Lett.*, 2024, 9(12), 5924–5934, DOI: [10.1021/acsenenergylett.4c02610](https://doi.org/10.1021/acsenenergylett.4c02610).
- 152 I. Kouroudis, K. T. Tanko, M. Karimipour, A. B. Ali, D. K. Kumar, V. Sudhakar, R. K. Gupta, I. Visoly-Fisher, M. Lira-Cantu and A. Gagliardi, Artificial Intelligence-Based, Wavelet-Aided Prediction of Long-Term Outdoor Performance of Perovskite Solar Cells, *ACS Energy Lett.*, 2024, 9(4), 1581–1586, DOI: [10.1021/acsenenergylett.4c00328](https://doi.org/10.1021/acsenenergylett.4c00328).
 - 153 D. Mohanty and A. K. Palai, Comprehensive Machine Learning Pipeline for Prediction of Power Conversion Efficiency in Perovskite Solar Cells, *Adv. Theory Simul.*, 2023, 6(12), 2300309, DOI: [10.1002/adts.202300309](https://doi.org/10.1002/adts.202300309).
 - 154 J. Kettle, Using ISOS Consensus Test Protocols for Development of Quantitative Life Test Models in Ageing of Organic Solar Cells, *Sol. Energy Mater. Sol. Cells*, 2017, 167, DOI: [10.1016/j.solmat.2017.04.005](https://doi.org/10.1016/j.solmat.2017.04.005).
 - 155 M. D. Wilkinson, M. Dumontier, Ij. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. G. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. C. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao and B. Mons, The FAIR Guiding Principles for Scientific Data Management and Stewardship, *Sci. Data*, 2016, 3(1), 160018, DOI: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).
 - 156 S. G. Baird and T. D. Sparks, What Is a Minimal Working Example for a Self-Driving Laboratory, *Matter*, 2022, 5(12), 4170–4178, DOI: [10.1016/j.matt.2022.11.007](https://doi.org/10.1016/j.matt.2022.11.007).
 - 157 A. M. Salih, Z. Raisi-Estabragh, I. B. Galazzo, P. Radeva, S. E. Petersen, K. Lekadir and G. Menegaz, A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME, *Adv. Intell. Syst.*, 2025, 7(1), 2400304, DOI: [10.1002/aisy.202400304](https://doi.org/10.1002/aisy.202400304).
 - 158 O. Fajardo-Fontiveros, I. Reichardt, H. R. De Los Ríos, J. Duch, M. Sales-Pardo and R. Guimerà, Fundamental Limits to Learning Closed-Form Mathematical Models from Data, *Nat. Commun.*, 2023, 14(1), 1043, DOI: [10.1038/s41467-023-36657-z](https://doi.org/10.1038/s41467-023-36657-z).
 - 159 Y. Gebreyesus, D. Dalton, S. Nixon, D. De Chiara and M. Chinnici, Machine Learning for Data Center Optimizations: Feature Selection Using Shapley Additive exPlanation (SHAP), *Future Internet*, 2023, 15(3), 88, DOI: [10.3390/fi15030088](https://doi.org/10.3390/fi15030088).
 - 160 S. M. Lundberg and S.-I. Lee, *A Unified Approach to Interpreting Model Predictions*. In *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017, Vol. 30.
 - 161 *Game Theory: An Overview*. In *Production Planning in Production Networks: Models for Medium and Short-term Planning*, ed P. Argoneto, G. Perrone, P. Renna, G. Lo Nigro, M. Bruccoleri, S. N. La Diega, Springer, London, 2008, pp. 13–24, DOI: [10.1007/978-1-84800-058-2_2](https://doi.org/10.1007/978-1-84800-058-2_2).
 - 162 P. M. Stevenson, Optimized Perturbation Theory, *Phys. Rev. D*, 1981, 23(12), 2916–2944, DOI: [10.1103/PhysRevD.23.2916](https://doi.org/10.1103/PhysRevD.23.2916).
 - 163 S. Zhao, S. Zhou, Z. Guo, H. Luo, Z. Jiang, N. Lin, M. Chen, L. Li and C. Li, Machine Learning-Assisted Analysis of Perovskite Solar Cell Long-Term Stability under Multiple Environmental Factors, *ACS Sustainable Chem. Eng.*, 2025, 13(19), 7155–7165, DOI: [10.1021/acssuschemeng.5c01361](https://doi.org/10.1021/acssuschemeng.5c01361).
 - 164 A. H. Rumman, M. A. Sahriar, M. T. Islam, K. M. Shorowordi, J. Carbonara, S. Broderick and S. Ahmed, Data-Driven Design for Enhanced Efficiency of Sn-Based Perovskite Solar Cells Using Machine Learning, *APL Mach. Learn.*, 2023, 1(4), 046117, DOI: [10.1063/5.0177271](https://doi.org/10.1063/5.0177271).
 - 165 A. H. Rumman, M. A. Sahriar, M. T. Islam, K. M. Shorowordi, J. Carbonara, S. Broderick and S. Ahmed, Data-Driven Design for Enhanced Efficiency of Sn-Based Perovskite Solar Cells Using Machine Learning, *APL Mach. Learn.*, 2023, 1(4), 046117, DOI: [10.1063/5.0177271](https://doi.org/10.1063/5.0177271).
 - 166 C. Draxl and M. Scheffler, The NOMAD Laboratory: From Data Sharing to Artificial Intelligence, *J. Phys. Mater.*, 2019, 2(3), 036001, DOI: [10.1088/2515-7639/ab13bb](https://doi.org/10.1088/2515-7639/ab13bb).
 - 167 P. Graniero, M. Khenkin, H. Köbler, N. T. P. Hartono, R. Schlatmann, A. Abate, E. Unger, T. J. Jacobsson and C. Ulbrich, The Challenge of Studying Perovskite Solar Cells' Stability with Machine Learning, *Front. Energy Res.*, 2023, 11.
 - 168 E. Unger and T. J. Jacobsson, The Perovskite Database Project: A Perspective on Collective Data Sharing, *ACS Energy Lett.*, 2022, 7(3), 1240–1245, DOI: [10.1021/acsenenergylett.2c00330](https://doi.org/10.1021/acsenenergylett.2c00330).
 - 169 M. Mammeri, L. Dehimi, H. Bencherif and F. Pezzimenti, Paths towards High Perovskite Solar Cells Stability Using Machine Learning Techniques, *Sol. Energy*, 2023, 249, 651–660, DOI: [10.1016/j.solener.2022.12.002](https://doi.org/10.1016/j.solener.2022.12.002).
 - 170 C. Yang, X. Chong, M. Hu, W. Yu, J. He, Y. Zhang, J. Feng, Y. Zhou and L.-W. Wang, Accelerating the Discovery of Hybrid Perovskites with Targeted Band Gaps via Interpretable Machine Learning, *ACS Appl. Mater. Interfaces*, 2023, 15(34), 40419–40427, DOI: [10.1021/acsami.3c06392](https://doi.org/10.1021/acsami.3c06392).
 - 171 C. Chen, Y. Zuo, W. Ye, X. Li, Z. Deng and S. P. Ong, A Critical Review of Machine Learning of Energy Materials, *Adv. Energy Mater.*, 2020, 10(8), 1903242, DOI: [10.1002/aenm.201903242](https://doi.org/10.1002/aenm.201903242).
 - 172 J. Li, B. Pradhan, S. Gaur and J. Thomas, Predictions and Strategies Learned from Machine Learning to Develop High-Performing Perovskite Solar Cells, *Adv. Energy Mater.*, 2019, 9(46), 1901891, DOI: [10.1002/aenm.201901891](https://doi.org/10.1002/aenm.201901891).
 - 173 *The use of ChatGPT to generate experimentally testable hypotheses for improving the surface passivation of perovskite solar cells: Cell Reports Physical Science*. [https://www.cell.com/cell-reports-physical-science/fulltext/S2666-3864\(24\)00327-8](https://www.cell.com/cell-reports-physical-science/fulltext/S2666-3864(24)00327-8) (accessed 2025-05-20).
 - 174 J. Li, B. Pradhan, S. Gaur and J. Thomas, Predictions and Strategies Learned from Machine Learning to Develop High-Performing Perovskite Solar Cells, *Adv. Energy Mater.*, 2019, 9(46), 1901891, DOI: [10.1002/aenm.201901891](https://doi.org/10.1002/aenm.201901891).
 - 175 P. Shetty, A. C. Rajan, C. Kuenneth, S. Gupta, L. P. Panchumarti, L. Holm, C. Zhang and R. Ramprasad,

- A General-Purpose Material Property Data Extraction Pipeline from Large Polymer Corpora Using Natural Language Processing, *npj Comput. Mater.*, 2023, **9**(1), 52, DOI: [10.1038/s41524-023-01003-w](https://doi.org/10.1038/s41524-023-01003-w).
- 176 How ChatGPT and our foundation models are developed. OpenAI Help Center. <https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed> (accessed 2025-09-02).
- 177 G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang and L. Yang, Physics-Informed Machine Learning, *Nat. Rev. Phys.*, 2021, **3**(6), 422–440, DOI: [10.1038/s42254-021-00314-5](https://doi.org/10.1038/s42254-021-00314-5).
- 178 M. A. B. Siddik, A. Shehabi and L. Marston, The Environmental Footprint of Data Centers in the United States, *Environ. Res. Lett.*, 2021, **16**(6), 064017, DOI: [10.1088/1748-9326/abfb1](https://doi.org/10.1088/1748-9326/abfb1).
- 179 N. S. A., *The Environmental Cost of Data Centers*. Net Zero Insights. <https://netzeroinsights.com/resources/data-centers-environmental-cost/> (accessed 2025-09-02).
- 180 Y. Wang, Z. Lv, L. Zhou, X. Chen, J. Chen, Y. Zhou, V. A. L. Roy and S.-T. Han, Emerging Perovskite Materials for High Density Data Storage and Artificial Synapses, *J. Mater. Chem. C*, 2018, **6**(7), 1600–1617, DOI: [10.1039/C7TC05326F](https://doi.org/10.1039/C7TC05326F).
- 181 L. A. Castriotta, ChatGPT Integration in Perovskite Research: Unveiling Pros and Cons of AI Integration for Scientific Advancements, *Sustainable Energy Fuels*, 2024, **8**(4), 697–699, DOI: [10.1039/D3SE01562A](https://doi.org/10.1039/D3SE01562A).
- 182 T. Dai, S. Vijayakrishnan, F. T. Szczypiński, J.-F. Ayme, E. Simaei, T. Fellowes, R. Clowes, L. Kotopantov, C. E. Shields, Z. Zhou, J. W. Ward and A. I. Cooper, Autonomous Mobile Robots for Exploratory Synthetic Chemistry, *Nature*, 2024, **635**(8040), 890–897, DOI: [10.1038/s41586-024-08173-7](https://doi.org/10.1038/s41586-024-08173-7).
- 183 F. Adams, A. McDannald, I. Takeuchi and A. G. Kusne, Human-in-the-Loop for Bayesian Autonomous Materials Phase Mapping, *Matter*, 2024, **7**(2), 697–709, DOI: [10.1016/j.matt.2024.01.005](https://doi.org/10.1016/j.matt.2024.01.005).