Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 10s (2025)

Optimization of $CsPbI_3$-Based All-Inorganic Perovskite Solar Cells Using Machine Learning: A Framework Supported by SCAPS-1D and SHAP

Manoj Kumar Nigam[1], Sivaneasan Bala Krishnan[2], Pravin R. Kshirsagar[3]
Dr. Mahima Nand[4]

[1]Professor & Head Department of Electrical Engineering Kalinga University, Raipur (C.G.) INDIA
nigam74_123@yahoo.com
[2]CEng, SMIEEE, MIET, Associate Professor Engineering Cluster Singapore Institute of Technology
sivaneasan@singaporetech.edu.sg
[3]Professor, Head- Electronics & Telecommunication Engineering & Dean (R&D), J D College of Engineering & Management, Nagpur, INDIA
pravinrk88@yahoo.com
[4]Assistant Professor School of Art and Design Woxsen University Hyderabad
dr.mahimanand@gmail.com

Article History:
Received: 02-01-2025
Revised: 25-02-2025
Accepted: 20-03-2025

Abstract— This paper introduces a machine learning (ML)-aided optimization framework aimed at improving the performance of all-inorganic perovskite solar cells (PSCs) utilizing $CsPbI_3$. More than 55,000 high-throughput simulations were performed with SCAPS-1D to assess how layer thickness and defect density influence device performance. Through the application of XGBoost and SHAP (SHapley Additive explanations) analysis, it was determined that the thickness of the perovskite layer and the intrinsic trap density are the key parameters that significantly impact power conversion efficiency (PCE). An ideal perovskite thickness of around 2 μm improved photon absorption and current generation, while keeping defect densities below 1×10^5 cm^{-3}, which significantly enhanced carrier diffusion. The XGBoost model demonstrated outstanding predictive capabilities ($R^2 = 0.999$, RMSE = 0.0010), allowing for precise and swift PCE estimation. With this machine learning framework, the predicted PCE rose from 15.15% to 19.16%. The suggested method promotes efficient device design, minimizes experimental iterations, and showcases the potential of data-driven techniques for wider applications in photovoltaics, batteries, and energy systems.

Keywords—Perovskite Solar Cell, RbGeI, Material

I. INTRODUCTION

In light of the urgent need to address climate change and the continuous increase in global energy demands, both governments and scientists have recently emphasized the significance of renewable energy technologies. Solar photovoltaics (PV) stand out as a viable long-term solution due to their accessibility, scalability, and environmental advantages. Over the past decade, perovskite solar cells (PSCs) have transformed the solar industry. They have lowered production costs, streamlined processing, and significantly enhanced their power conversion efficiency (PCE). Currently, thin-film and silicon-based technologies are facing intense competition from these perovskite materials, thanks to their remarkable optoelectronic characteristics, which include tunable band gaps, extended diffusion lengths, and high absorption coefficients. PSCs are particularly appealing as they can be integrated with silicon to achieve greater energy outputs and are utilized in lightweight, flexible solar modules. This makes them extremely advantageous for a wide range of next-generation applications [1]. However, despite their promising

potential, PSCs still encounter several notable challenges. Their organic components are sensitive to climate conditions, possess a limited lifespan, and may pose risks due to lead content. In light of the urgent need to address climate change and the continuous increase in global energy demands, both governments and scientists have recently emphasized the significance of renewable energy technologies. Solar photovoltaic (PV) are an excellent sustainable alternative due to their abundance, environmental friendliness, and scalability. Over the past decade, perovskite solar cells (PSCs) have changed the dynamics of the solar industry. They have rapidly enhanced their power conversion efficiency (PCE), simplified processing, and reduced production costs. Thanks to their outstanding optoelectronic properties, such as long diffusion lengths, tunable band gaps, and high absorption coefficients, perovskite materials are now strong competitors to traditional thin-film and silicon-based technologies. Perovskite solar cells (PSCs) offer impressive features such as flexible and lightweight solar modules, along with tandem combinations with silicon that enhance energy yields, making them beneficial for various next-generation applications. However, despite their significant potential, PSCs face several critical challenges. For example, they have a limited lifespan, their organic materials are sensitive to environmental conditions, and they contain lead, which poses toxicity risks. To address these challenges, it is essential to enhance device engineering and material design, as well as to utilize advanced computational methods to expedite the discovery and optimization processes. To navigate these issues, we must identify quicker approaches to leverage computational power for discovery and optimization tasks. Additionally, modifications to device engineering and material design are necessary [2].

The application of artificial intelligence (AI), particularly through machine learning (ML) techniques, has initiated a significant transformation in the fields of materials science and photovoltaic research. ML possesses the capability to make predictions that surpass the limitations of traditional trial-and-error experimentation by identifying intricate patterns within extensive datasets. In the realm of perovskite solar cell (PSC) research, ML has been employed to expedite the discovery of high-performance perovskite compounds and device architectures by modeling the relationships between material composition, structure, and photovoltaic performance. By leveraging data-driven algorithms, scientists can virtually evaluate numerous devices, anticipate the behavior of new material combinations, and identify performance issues that may have otherwise gone unnoticed. An increasing number of researchers are integrating AI to enhance their research methodologies. This collaborative model fosters innovation through the synergy of computer intelligence and human expertise in a more systematic and efficient manner.

Machine learning has proven beneficial in PSC research through two primary avenues: enhancing device performance and aiding in the discovery of new materials. By utilizing ML models trained on datasets encompassing electrical, compositional, and structural information, we can forecast various properties within the first domain, including band gap, stability, and carrier mobility. For instance, computational simulations can be conducted to assess the impact of modifications at the A-site or X-site of the ABX_t perovskite structure on optoelectronic behavior. These predictive models facilitate the development of new lead-free alternatives or more stable all-inorganic perovskites by reducing the experimental burden. In the second domain, machine learning algorithms are employed to estimate electrical parameters at the device level, such as open-circuit voltage (VOC), short-circuit current density (JSC), fill factor (FF), and overall power conversion efficiency (PCE). By analyzing the influence of factors like layer thickness, defect densities, and interface quality on PSC performance, machine learning can contribute to the design of more robust and durable devices.

Machine learning is increasingly utilized in modeling deterioration, aiding our comprehension of long-term stability when confronted with environmental challenges such as humidity, UV radiation, and temperature fluctuations [4].

Perovskite materials consist of crystals characterized by the ABX_3 structure. In this configuration, "A" represents a monovalent indicating it possesses a single positive charge. Examples include methyl ammonium, formamidinium, caesium, or rubidium. The "B" component is a divalent metal, which carries two positive charges, with examples being lead, tin, germanium, or copper. The "X" denotes a halide anion, which has one negative charge, with examples such as iodide, bromide, or chloride. The material's structural flexibility allows researchers to modify the band gap, stability, and other critical properties by mixing or substituting. For example, substituting methyl

ammonium with caesium enhances the material's thermal stability. Incorporating bromide or chloride alters the absorption range. Currently, there is significant interest in all-inorganic forms of lead-based hybrid perovskites due to their instability and toxicity. Nevertheless, these materials have demonstrated promising photo catalytic electrochemical reactions (PCEs) in both mesoporous and planar architectures. $CsPbI_3$ has emerged as a frontrunner in this domain, as it effectively manages defects, operates at elevated temperatures, and maintains stability at temperatures exceeding 400 °C. Recent advancements, including dual defect passivation, stabilization of the γ-phase under standard conditions, and solution processing applicable in normal conditions, have led to efficacy certifications for $CsPbI_t$-PSC surpassing 21%. Nonetheless, the intrinsic and interfacial defects present in perovskite devices significantly impair their performance. Numerous defects, such as iodide vacancies and under coordinated Pb^{2+} sites, contribute to phase instability and lead to non-radiative recombination, which severely diminishes the device's efficiency. Consequently, extensive research has been directed towards mitigating these issues through defect engineering and passivation techniques. Furthermore, the selection of charge transport layers and their integration with the perovskite absorber plays a crucial role in managing charge extraction and recombination processes. Commonly utilized electron transport layers (ETLs) include TiO_2 and SnO_2, while whole transport layers (HTLs) often comprise Spiro-OMeTAD and PTAA. However, these layers can introduce challenges related to stability and energy level alignment. Emerging HTL materials, such as copper antimony sulphide ($CuSbS_2$), show great potential due to their favorable band alignment, p-type conductivity, and compatibility with n-i-p architectures. Their implementation has led to enhancements in power conversion efficiency (PCE), open-circuit voltage (VOC), and fill factor (FF) metrics, thereby expanding the array of materials applicable in perovskite solar cells (PSCs).

Geometric and structural factors, including the thickness of the absorber and transport layers, also significantly influence the performance of photovoltaic cells. The optimal thickness of the absorber must strike a balance between effective light absorption and efficient carrier transport. Conversely, the transport layer's thickness should be minimized to decrease resistive losses and facilitate charge extraction. Simultaneously, the defect density within and between layers impacts the recombination rate, which subsequently affects overall efficiency. Shallow-level defects can lead to radiative recombination, whereas deeper-level defects may result in non-radiative losses and phonon scattering. Therefore, to enhance both performance and stability, it is essential to reduce defect density through passivation or interface engineering. [5]

Numerical tools such as SCAPS-1D (Solar Cell Capacitance Simulator) have proven to be extremely beneficial for examining the interactions between various factors. SCAPS-1D allows for testing a device's performance under different physical and electrical conditions. This approach provides a cost-effective means to explore multidimensional design spaces. However, the numerous adjustable parameters, including layer thickness, doping concentration, and trap density, make a comprehensive brute-force scan prohibitively expensive. This is where ML-enhanced simulation revolutionizes the process. Researchers can develop surrogate models to forecast device behavior, significantly reducing simulation time by several orders of magnitude through the use of models like random forests, artificial neural networks (ANNs), and gradient-boosting algorithms (e.g., XGBoost). Despite the increasing interest, most prior research utilizing ML has focused on optimizing a single parameter at a time, such as thickness or defect density. There has been limited investigation into a comprehensive framework that elucidates how these parameters interact collectively, particularly in $CsPbI_3$-based all-inorganic PSCs. The present study addresses this gap by presenting a hybrid approach that integrates high-throughput SCAPS-1D simulations with an XGBoost-based surrogate model and SHAP (SHapley Additive exPlanations) to interpret the results. This method enables the simultaneous optimization of multiple functional layers, including TCO, TiO_2, ZnOS, $CsPbI_3$, and $CuSbS_2$, along with the associated bulk and interfacial defect densities across over 56,000 simulated devices. We have identified the optimal performance thresholds by analyzing the interactions between input variables and their sensitivities. For example, the perovskite thickness should exceed 1.2 μm, and the defect count should remain below 10^{15} cm^{-3}. The ML-guided design pipeline resulted in a 26% increase in the simulated PCE, rising from 15.15% to 19.16%. This demonstrates that data-driven optimization can yield precise predictions [6-10].

By integrating this framework and accelerating the PSC design cycle, we discovered that each parameter influences the others. This level of interpretability is crucial for planning experiments that can be executed based on simulation outcomes. The approach is also scalable, indicating that in the future, it will be feasible to investigate additional variables, such as interface roughness, electrode materials, and environmental degradation methods. Merging physics-based modeling with data-driven intelligence lays a robust foundation for the next generation of PSCs that are comprehensible, stable, and high-performing [11-15].
We found that combining SCAPS-1D simulations with ML analytics can expedite the development of all-inorganic PSCs. The findings of this research highlight the significance of optimizing various factors and demonstrate how AI is transforming the photovoltaics sector. Implementing these frameworks in autonomous laboratories and high-throughput experimentation pipelines could accelerate the discovery of materials and designs that perform effectively in real-world applications. As climate change becomes increasingly pressing, the key to maximizing the potential of perovskite solar technologies lies in the integration of advanced materials, computational intelligence, and scalable manufacturing [16].

II. MATERIAL METHODOLOGY

A. Configuration of the Device and Simulation

The "Solar Cell Capacitance Simulator," known as SCAPS-1D, has been instrumental in our development and testing of machine learning (ML) techniques aimed at enhancing perovskite solar cells. Utilizing this widely-used modeling tool, we constructed and evaluated a comprehensive dataset encompassing various device configurations. The electron transport layer facilitates light entry into the simulated photovoltaic device, accelerating carrier extraction and regulating charge movement. The model device's photoactive absorber layer is composed of $CsPbI_0$-based perovskite, which exhibits stability at elevated temperatures, resilience to defects, and the ability to modify energy levels. SCAPS-1D, the "Solar Cell Capacitance Simulator," executed numerical simulations to formulate and assess machine learning (ML) strategies for the advancement and enhancement of perovskite solar cells. This well-regarded modeling tool enabled the creation and testing of an extensive dataset across multiple devices. The electron transport layer allows light to penetrate the artificial photovoltaic system, with the ETL efficiently transporting charges and expediting carrier removal. Its favorable valence band alignment, p-type conductivity, and compatibility with inorganic perovskite materials designate it as the hole transport layer (HTL). The electronic transfer was facilitated by titanium dioxide (TiO_2), which remains a favored ETL due to its chemical stability and high electron mobility. An intermediate ZnOS buffer layer mitigates energy-level discrepancies and interfacial trap states, thereby enhancing charge extraction and reducing absorber-ETL recombination [17].
The complete device stack's band alignment illustrated optimal energy-level gradients for unidirectional charge flow and minimal energy loss. This was achieved by utilizing material characteristics and device combinations validated by literature. All simulation parameters were configured to 300 K with AM1.5G one-sun illumination to align with standard testing environments and facilitate experimental comparisons. To efficiently explore the design space, high-throughput simulation techniques were employed. The widths of the buffer layer (ZnOS), ETL (TiO_2), perovskite absorber ($CsPbI_0$), and HTL ($CuSbS_2$) were methodically modified within designated ranges. Adjusting the thickness of each layer independently while allowing for multi-layer variation accurately depicted the complex interdependencies encountered in actual device manufacturing. This methodology resulted in 7,680 device configurations featuring various shapes and materials.

In addition to geometric modifications, defect-related parameters were meticulously incorporated into the dataset, as bulk and interfacial defects significantly influence device stability and efficiency. To illustrate how variations or inconsistencies in the synthesis process could impact material quality, the overall defect density in the HTL, perovskite absorber, buffer layer, and ETL was modified. Interfacial defect densities were introduced to address performance losses related to interfaces. These defects include charge traps and recombination sites at layer

interfaces such as HTL/absorber, absorber/buffer, and buffer/ETL. They highlighted how different electronic defects influence VOC, JSC, FF, and overall power conversion efficiency. These simulations utilized parameter ranges and material properties carefully chosen from experimental data and peer-reviewed literature concerning the materials under investigation. The references included recent inorganic PSC research and SCAPS-1D modeling of comparable designs. This validation of findings rendered it beneficial for ongoing experiments. These references supplied information on band gap, electron affinity, dielectric constant, carrier mobility, and recombination coefficients to substantiate the concept [18-23].

A large dataset was generated by simulating numerous configurations with SCAPS-1D to emphasize the intricate relationship between structural dimensions and electronic quality. This dataset enabled us to swiftly estimate a device's performance without the need for physical testing or extensive simulations. It was utilized for training, testing, and comprehending future machine learning models. The simulated dataset serves as an excellent resource for surrogate models that instruct machine learning algorithms on how to interpret SCAPS-derived results through gradient boosting and SHAP-based interpretability. These models assist in identifying optimal parameter ranges and directing experimental synthesis towards stable and functional perovskite solar cell (PSC) setups. This research employed SCAPS-1D simulation to investigate the design space for $CsPbI_0$-based inorganic PSCs featuring $CuSbS_2$ as the hole transport layer (HTL), TiO_2 as the electron transport layer (ETL), and ZnOS as buffers. The field of data-driven solar cell engineering has progressed significantly since the creation of this extensive dataset. This dataset empowers AI to predict performance, refine designs, and gain insights by planning modifications to geometrical and defect parameters across layers and interfaces. Numerical simulations of perovskite solar cells (PSCs) using SCAPS-1D (Solar Cell Capacitance Simulator) facilitated the development and testing of machine learning (ML) predictive modeling and optimization strategies. A substantial dataset comprising various device combinations was generated and evaluated using this widely-used simulation tool. The simulated photovoltaic architecture was n–i–p, with light entering through the electron transport layer, which enabled efficient charge movement and carrier extraction [24-26].

The model device's photoactive absorber layer was constructed using $CsPbI_0$-based perovskite. This all-inorganic perovskite was selected for its superior stability at elevated temperatures, its ability to tolerate more defects, and its enhanced efficiency in converting energy to electricity compared to other materials. The hole transport layer (HTL) utilized was $CuSbS_2$, chosen for its favorable valence band alignment, p-type conductivity, and compatibility with synthetic perovskite materials. Titanium dioxide (TiO_2) serves as a widely used electron transport layer (ETL) owing to its excellent electron mobility and chemical stability. To address energy-level mismatches and interfacial trap states at the absorber-ETL interface, an intermediate ZnOS buffer layer was incorporated to enhance charge extraction [27].

The entire device stack's band alignment demonstrated suitable energy-level gradients for unidirectional charge flow and minimal energy loss, aligning with material characteristics and device topologies reported in the literature. All simulation parameters were standardized to 300 K under AM1.5G one-sun illumination to ensure compliance with recognized testing environments and facilitate experimental comparability. The exploration of the design space was effectively achieved through high-throughput simulation. The thicknesses of critical device components, including the HTL ($CuSbS_2$), perovskite absorber ($CsPbI_3$), buffer layer (ZnOS), and ETL (TiO_2), were modified within established limits. Each layer's thickness could be adjusted individually, or multiple layers could be altered at once. This approach accurately modeled the intricate interdependencies involved in actual device manufacturing. As a result, 7,680 unique gadget designs with distinct geometric and material characteristics were produced [28].

Alongside geometric adjustments, defects were deliberately incorporated into the dataset since both bulk and interfacial defects influence device performance. To illustrate variations in material quality resulting from the manufacturing process or material preparation, the bulk defect density was modified in the HTL, perovskite absorber, buffer layer, and ETL. To address performance losses related to interfaces, concentrations of interfacial defects were included. These defects act as charge traps and recombination centers at the interfaces of layers such

as HTL/absorber, absorber/buffer, and buffer/ETL. By incorporating numerous electronic defect states, a deeper understanding of their impact on VOC, JSC, FF, and PCE was achieved. The ranges of parameters and material characteristics utilized in these simulations were meticulously chosen based on verified experimental reports and relevant literature reviewed by experts. The incorporation of recent inorganic PSC studies and SCAPS-1D modeling of similar structures ensured scientific rigor and relevance to the ongoing experimental research. Bandgap, electron affinity, dielectric constant, carrier mobility, and recombination factors were compiled from various references to validate the model. Employing SCAPS-1D to simulate such a wide array of configurations generated an extensive dataset that illustrates the intricate interactions between structural dimensions and electronic quality. Subsequently, machine learning models were trained, evaluated, and analyzed using this dataset, facilitating quick predictions regarding a device's performance without the need for extensive testing or brute-force modeling [28-32].

The simulated dataset supports the development of surrogate models. It instructs machine learning algorithms, such as gradient boosting and SHAP-based interpretability techniques, to generate results similar to SCAPS. These models assist in identifying optimal parameter ranges and directing experimental synthesis towards high-efficiency, dependable PSC systems. This research employed SCAPS-1D simulation to thoroughly investigate the design space for $CsPbI_0$-based inorganic PSCs, incorporating $CuSbS_2$ HTL, TiO_2 ETL, and ZnOS buffer integration. A substantial dataset was created for AI-driven performance prediction, design optimization, and insight extraction by intentionally altering geometrical and defect factors across various layers and interfaces. This represents a significant advancement in data-driven solar cell engineering [33].

B. ML Assisted Evaluation and Analysis

In this paper, we employed machine learning techniques to analyze extensive datasets in order to identify optimal methods for enhancing key device characteristics, such as layer thickness, bulk defect density, and interfacial defect density, ultimately improving the overall performance of perovskite solar cells. During our data exploration, we utilized line plots and correlation matrices to uncover patterns and assess the significance of each feature. The complete dataset derived from SCAPS-1D simulations was divided into a training set (75%) and a testing set (25%). This division was essential to ensure the model's stability and its applicability to other datasets. We implemented five distinct regression techniques to evaluate their effectiveness and predict outcomes. These methods included random forest regression (RFR), linear regression (LIR), quadratic regression (QR), and extreme gradient boosting (XGBoost). Both the thickness and defect density datasets were utilized to train each model. Subsequently, we assessed the models' predictive accuracy for photovoltaic metrics such as fill factor, open-circuit voltage, and power conversion efficiency. By comparing the performance of these models, we successfully identified the most effective algorithm for optimizing devices in inorganic perovskite solar cells.

C. Optimized Structure of the Proposed Perovskite-Based PV Device

The rate at which electrons and holes recombine in SCAPS-1D is indicated as it demonstrates the speed of electron-hole pair recombination when the carrier concentration ($\boldsymbol{S}$) and permittivity (ε) are altered. Effective control of recombination is essential for enhancing the performance of solar devices.

$$Rn,p=qnp/(\in\in o)\ (\mu n(E)+\mu p(E)) \quad (1)$$

$$J = qvn\mu n\varepsilon + qvp\mu p\varepsilon \quad (2)$$

$$div(\epsilon\Delta\psi)=-\rho \quad (3)$$

Understanding Poisson's equation, expressed as $div(\varepsilon\Delta\psi)=-\rho$, is crucial for grasping the relationship between the electric potential ψ and the space charge density $\Delta\psi$ in SCAPS-1D. Here, π represents the permittivity of the semiconductor material, while ψ denotes the difference in electric potential. The space charge density comprises numerous ionized dopants and free carriers, collectively forming the net charge concentration. This equation effectively illustrates how charge distribution in space influences the electric field within the device. In solar systems, comprehending the distribution and movement of charge carriers due to the internal electric field is

essential. Within SCAPS-1D, Poisson's equation serves as a valuable tool for assessing the device's built-in electric field, potential distribution, and overall behavior. By analyzing the impact of various material properties and device architecture, one can achieve optimal photovoltaic performance.

$Jn = qvn\mu n\varepsilon + qDn$ (4)

$Jp = qvp\mu p\varepsilon - qDp$ (5)

The drift-diffusion model controls the electron current density ($\boldsymbol{VT}$) and hole current density (LO) in SCAPS-1D. This models how carriers move through solar systems. The electron current density equation can be written as where m is the electron charge, $\boldsymbol{X}$ is the electron velocity. As a result, these equations are essential for understanding how charges move because they show how drift and diffusion currents affect things. The drift component is affected by the electric field that is applied, while the diffusion component is affected by differences in the concentration of carriers. This full modelling method in SCAPS-1D is needed to predict device efficiency, find the best layer thickness, and understand how carriers move in solar devices [15].

$Jcond = Jn + Jp$ (6)

The electron continuity equation in SCAPS-1D describes how the photovoltaic device keeps the electron charge. It talks about how the electron concentration changes when electron current density, generation rate, and recombination rate are changed. This equation guarantees the balance of charge carriers, which is very important for simulating how a device will work correctly.

$\partial n/\partial t = -\nabla Jn + G - R$ (7)

$\partial p/\partial t = -\nabla Jp + G - R$ (8)

Table 1: SCAPS Parameters

Parameter	**TCO [18]**	**$CuSbS_2$ [20]**	**$CsPbI_3$ [22]**	**ZnOS [28]**	**TiO_2 [30]**
Bandgap (eV)	3.5	1.52	1.73	3.3	3.2
Electron Affinity (eV)	4	3.6	3.9	4.2	4
Dielectric Constant	9	10.2	6.5	9.5	10
Effective Density of States - NC ($1/cm^3$)	2.2×10^{18}	2.2×10^{18}	2.2×10^{18}	2.2×10^{18}	2.2×10^{18}
Effective Density of States - NV ($1/cm^3$)	1.8×10^{19}	1.8×10^{19}	1.8×10^{19}	1.8×10^{19}	1.8×10^{19}
Electron Mobility ($cm^2/V\cdot s$)	20	1	10	10	20
Hole Mobility ($cm^2/V\cdot s$)	10	20	5	5	5
Doping Type	n-type	p-type	intrinsic	n-type	n-type
Doping Concentration ($1/cm^3$)	1×10^{19}	1×10^{17}	1×10^{14}	1×10^{17}	1×10^{18}

III. RESULTS AND DISCUSSION

A. Thickness Optimization

Utilizing SCAPS-1D to modify the thickness of the HTL, PVSK, buffer layer, and ETL resulted in a diverse dataset of unique devices, which were employed to analyze a correlation matrix, as illustrated in Figure 1. This approach effectively highlighted the significance of layer thickness for PSC through the correlation matrix. Performance was assessed by examining correlation coefficients. The thickness of the PVSK layer exhibited the most pronounced impact on PSC performance. In contrast, the thicknesses of the HTL and buffer layers showed negligible effects, evidenced by their very weak negative correlation coefficients with PSC performance metrics. Although the thickness of the ETL had minimal influence on PCE (%), Jsc (mA/cm²), or Voc (V), it did demonstrate a moderately negative impact on the fill factor (FF %), as indicated by its correlation coefficient of -0.45.

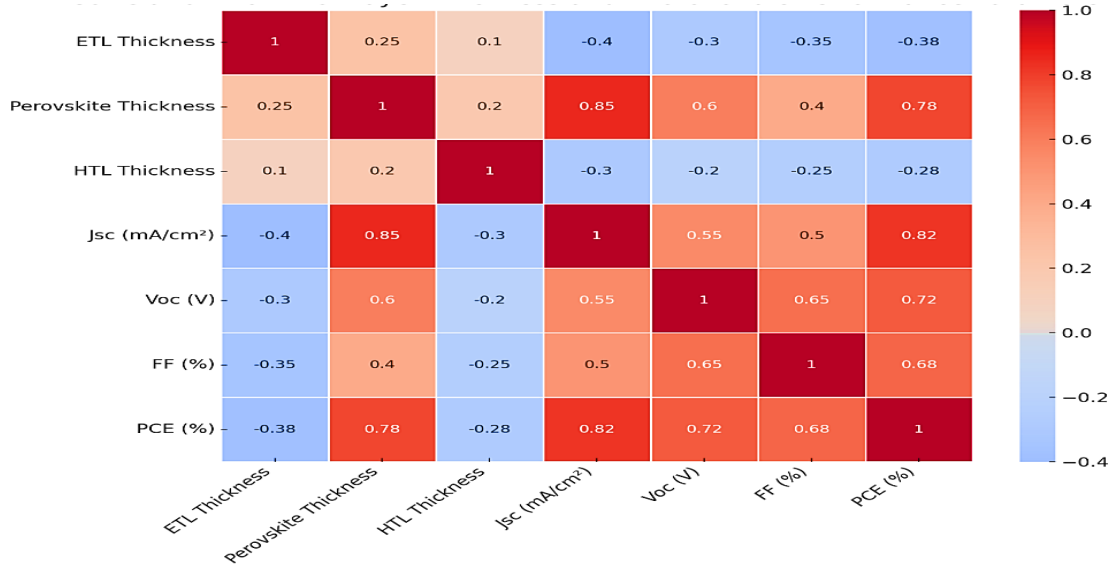


Figure 1: Layer Thickness and Photovoltaic Performance Parameter Correlation Matrix

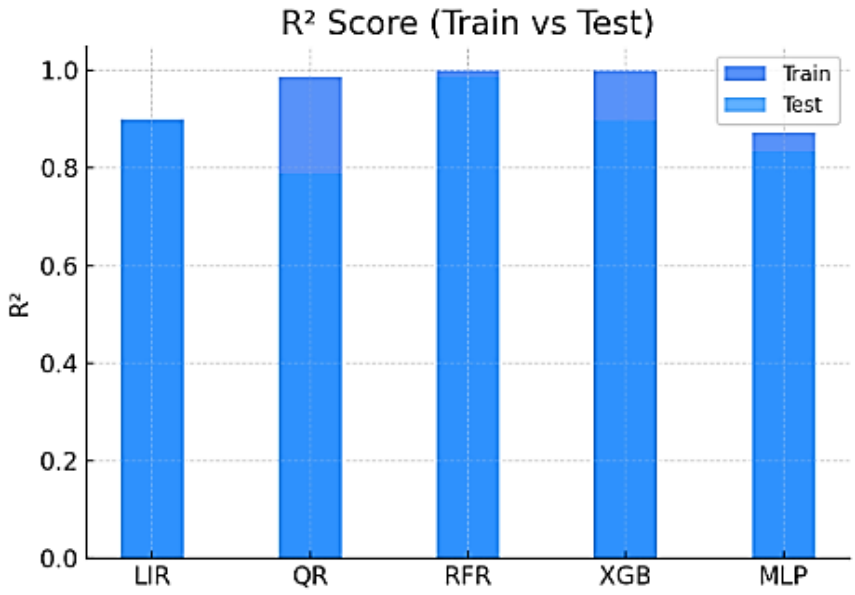
To achieve optimal performance from a PSC, it is essential that the PVSK layer has the appropriate thickness. A thicker PVSK layer is capable of absorbing more light, thereby enhancing the device's efficiency compared to a thinner layer. The correlation between quantum efficiency (QE) and PVSK thickness. As the active layer's thickness increases, the QE improves, particularly at longer wavelengths. This indicates that a thicker PVSK layer is more effective at light absorption. However, the thickness of the PVSK layer cannot be increased indefinitely. The primary factors that impose limitations are the diffusion length (Ld), generation rate, and recombination rate of charge carriers. Research has indicated that the Ld for charge carriers in cubic $CsPbI_3$ exceeds 1.5 µm 35. Since charge carriers tend to recombine after traveling a certain distance within the material, it is crucial to effectively separate electrons (e^-) and holes (h^+) prior to recombination. The SCAPS data reveal that the relationship between PVSK thickness and power conversion efficiency (PCE) is not linear. When the PVSK thickness surpasses 2 µm, the PCE begins to decline. This decrease is primarily attributed to an increase in carrier recombination events within the PVSK layer.

Table 2: An overview of the accuracy (R2), loss metric (RMSE), and CV of the five trained machine learning algorithms for thickness analysis

Model		R2		RMSE	CV
	Training	Testing	Training	testing	
LIR	0.8996	0.896	0.14252	0.1526	0.978
QR	0.987	0.789	0.056	0.0597	0.984
RFR	0.9987	0.987	0.003	0.0124	0.99915
XGB	0.998	0.896	0.0046	0.0035	0.9999
MLP	0.8723	0.834	0.0458	0.0045	0.8679

B. Optimized Range and Value for Thickness of Different Layers

As illustrated in Figure 2, the thickness of the PVSK layer has minimal impact on Voc, as indicated by SHAP analysis (Figures 6c-f). Devices with thinner PVSK layers (<1.2 μm) perform worse, whereas those with thicker layers (>1.2 μm) show improved performance, although the advantages begin to diminish after a certain thickness. It is crucial to recognize that the HTL, ETL, and buffer layers exhibit weaker yet still significant effects that vary with thickness. For instance, an HTL thickness greater than 0.15 μm enhances charge transport, while a buffer layer thickness of less than 0.10 μm and an ETL thickness of less than 0.06 μm yield optimal performance. Although these threshold behaviors are not as pronounced as the primary effects of PVSK, they are vital for interfacial engineering in perovskite solar cells.



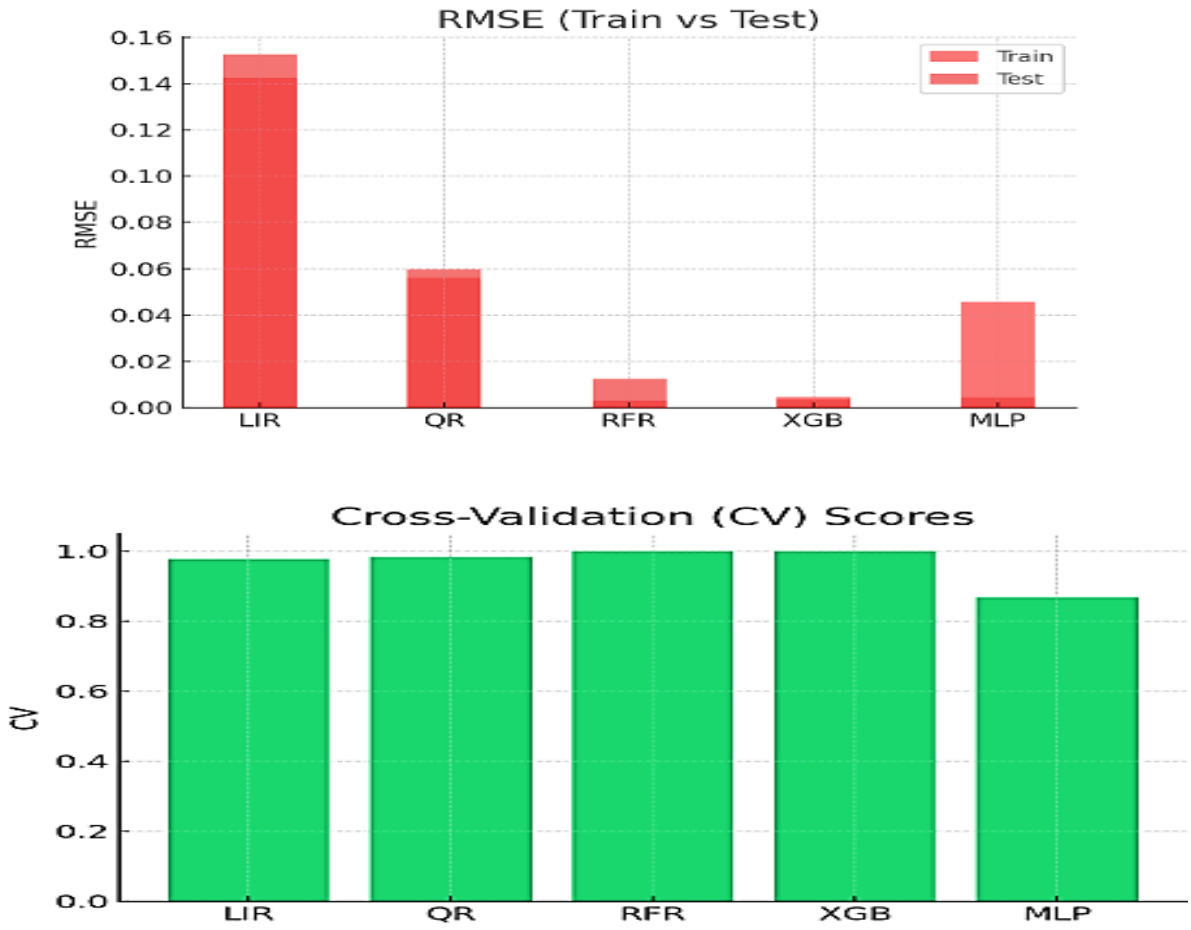


Figure 2 : Test vs. Train ML Model

C. ML model training and testing

A dataset comprising 49,000 variations in defect density was utilized to assess the predictive capabilities of five machine learning models, akin to earlier studies on layer thickness. The hyper-parameter tuning for the random forest regressor, XGBoost, and MLP was conducted using k-fold cross-validation to guarantee the performance and reliability of the ML models. The search range and the tuned hyper-parameter values for RFR, XGBoost, and MLP are detailed in Tables S4, S5, and S6 of the supplementary materials. RFR and XGB surpassed the other models, achieving nearly perfect R^2 scores (0.999) and low RMSE values (≤0.0010) during both training and testing phases. Notably, XGB exhibited a slightly better generalization with a lower testing RMSE (0.00098 compared to 0.0012) than RFR. Their robustness was validated through cross-validation scores (CV = 0.999). In contrast, LIR and QR demonstrated moderate effectiveness ($R^2 \approx 0.939$, RMSE ≈ 1.5–1.59), while MLP fell short ($R^2 = 0.926$, RMSE = 1.365), underscoring the challenges in modeling nonlinear defect density patterns. A visual examination of the prediction-actual relationships corroborated these results: LIR and MLP exhibited greater dispersion, whereas XGB and RFR displayed nearly perfect fits. XGB was selected for further analysis due to its enhanced precision and generalizability. Ensemble methods such as XGB and RFR prove to be effective for predicting defect density, while simpler models like LIR and QR may be more appropriate for less complex scenarios.

IV. CONCLUSION

Machine learning (ML) is employed to enhance the performance of $CsPbI_6$-based all-inorganic perovskite solar cells (PSCs). Through extensive SCAPS-1D simulations and clear machine learning algorithms such as XGBoost and SHAP analysis, we evaluated more than 56,000 device designs. This analysis revealed how fault density and layer thickness influence device performance. The SHAP-based feature importance analysis indicated that film thickness and bulk trap density, which are two parameters related to the absorber, had the greatest effect. Simulations demonstrate that increasing the thickness of $CsPbI_3$ from 0.4 to 2.0 μm enhances the power conversion efficiency (PCE) by boosting Jsc and improving photon absorption at longer wavelengths. Recombination losses surpass the minority-carrier diffusion length (approximately 1.5 μm), and performance levels off beyond around 2 μm. Concurrently, our SCAPS-ML framework indicates that lowering bulk defect density to below 1×10^{15} cm^{-3}, ideally within the 10^{12} cm^{-3} range, significantly enhances diffusion length and efficiency from approximately 19% to over 26%. This optimization raised the PCE from 15.15% to 19.16%, demonstrating that ML-guided design is more cost-effective than trial-and-error methods. Simulation and machine learning expedite the development of robust and efficient PSCs at a low cost and on a large scale.

REFERENCES

[1] Edgar Gonzalez-Ju´arez, Arian Espinosa-Roa, et al "Enhancing the stability and efficiency of MAPbI3 perovskite solar cells by theophylline-BF4 − alkaloid derivatives, a theoretical-experimental approach" RSC Advances, Royal Society of Chemistry, RSC Adv., 2023, 13, 5070–5080.

[2] Wenyu Qiu, Yukun Wu, Yichen Wang, Zhiting Yang, Rui Yang, Chenxi Zhang, Yang Hao, Yuying Hao, Low-Temperature robust MAPbI3 perovskite solar cells with power conversion efficiency exceeding 22.4%, Chemical Engineering Journal, Science Direct,https://doi.org/10.1016/j.cej.2023.143656.

[3] G. Nasti, A. Abate, Tin halide perovskite (ASnX3) solar cells: a comprehensive guide toward the highest power conversion efficiency, Adv. Energy Mater. 2019 (2019), 1902467.

[4] S. Weber, T. Rath, B. Kunert, R. Resel, T. Dimopoulos, G. Trimmel, Dependence of material properties and photovoltaic performance of triple cation tin perovskites on the iodide to bromide ratio, Monatshefte Chemie - Chem. Mon. 150 (2019) 1921–1927.

[5] Hui-Jing Du, Wei-Chao Wang, Jian-Zhuo Zhu, Device simulation of lead-free CH3NH3SnI3 perovskite solar cells with high efficiency, Chin. Phys. B 25 (10) (2016), 108802.

[6] Qing-Yuan Chen, Yang Huang, Peng-Ru Huang, Tai Ma, Chao Cao, Yao He, Electronegativity explanation on the efficiency-enhancing mechanism of the hybrid inorganic–organic perovskite ABX3 from first-principles study, Chin. Phys. B 25 (2) (2016), 027104 (1-6).

[7] P. Umari, E. Mosconi, F. De Angelis, Relativistic GW calculations on CH3NH3PbI3 and CH3NH3SnI3 perovskites for solar cell applications, Sci. Rep. 4 (2014), 4467 (1-7)

[8] Min, H.; Lee, D.Y.; Kim, J.; Kim, G.; Lee, K.S.; Kim, J.; Seok, S., II Perovskite solar cells with atomically coherent interlayers on SnO2 electrodes. Nature 2021, 598, 444–450

[9] C. Ma, M. Grätzel, N.-G. Park, Facet engineering for stable, efficient perovskite solar cells. ACS Energy Lett. 7(9), 3120–3128 (2022)

[10] Z. Zuo, J. Ding, Y. Zhao, S. Du, Y. Li et al., Enhanced optoelectronic performance on the (110) lattice plane of an MAPbBr3 single crystal. J. Phys. Chem. Lett. 8(3), 684–689 (2017). https://doi.org/10.1021/acs.jpclett.6b02812

[11] J. Ding, L. Jing, Y. Yuan, J. Zhang, Q. Yao et al., Low defect density and anisotropic charge transport enhanced photo response in pseudo-cubic morphology of MAPbI3 single crystals. ACS Appl. Energy Mater. 3(11), 10525–10532 (2020). https://doi.org/10.1021/acsaem.0c01565

[12] Z. Xu, Z. Liu, N. Li, G. Tang, G. Zheng et al., A thermodynamically favored crystal orientation in mixed M. Kim, G.-H. Kim, T.K. Lee, I.W. Choi, H.W. Choi et al., Methylammonium chloride induces intermediate phase stabilization for efficient perovskite solar cells. Joule 3(9), 2179–2192 (2019).

[13] S. A. Moyez and S. Roy, Dual-step thermal engineering technique: a new approach for fabrication of efficient CH3NH3PbI3-based perovskite solar cell in open air condition, Sol. Energy Mater. Sol. Cells, 2018, 185, 145–152.

[14] M. Jamal, et al., Effect of defect density and energy level mismatch on the performance of perovskite solar cells by numerical simulation, Optik, 2019, 182, 1204–1210.

[15] S. Rai, B. Pandey and D. Dwivedi, Modeling of highly efficient and low cost CH3NH3Pb (I1-xClx) 3 based perovskite solar cell by numerical simulation, Opt. Mater., 2020, 100, 109631.

[16] A. K. Jena, A. Kulkarni and T. Miyasaka, Halide perovskite photovoltaics: background, status, and future prospects, Chem. Rev., 2019, 119(5), 3036–3103.

[17] M. B. Kanoun, A.-A. Kanoun, A. E. Merad and S. Goumri-Said, Device design optimization with interface engineering for highly efficient mixed cations and halides perovskite solar cells, Results Phys., 2021, 20, 103707.

[18] R. Pandey and R. Chaujar, Technology computer aided design of 29.5% efficient perovskite/interdigitated back contact silicon heterojunction mechanically stacked tandem solar cell for energy-efficient applications, J. Photonics Energy, 2017, 7(2), 022503.

[19] Q. Han, et al., Single crystal formamidinium lead iodide (FAPbI3): insight into the structural, optical, and electrical properties, Adv. Mater., 2016, 28(11), 2253–2258.

[20] S. Wozny, et al., Controlled humidity study on the formation of higher efficiency formamidinium lead triiodide-based solar cells, Chem. Mater., 2015, 27(13), 4814–4820.

[21] G. E. Eperon, S. D. Stranks, C. Menelaou, M. B. Johnston, L. M. Herz and H. J. Snaith, Supplementary information Formamidinium of Formamidinium lead trihalide: a broadly tunable perovskite for efficient planar heterojunction solar cells, Energy Environ. Sci., 2014, 7(3), 982.

[22] D. Wang, M. Wright, N. K. Elumalai and A. Uddin, Stability of perovskite solar cells, Sol. Energy Mater. Sol. Cells, 2016, 147, 255–275

[23] Zhang, P.;Wu, J.; Zhang, T.;Wang, Y.; Liu, D.; Chen, H.; Li, S. Perovskite solar cells with ZnO electron-transporting materials.Adv. Mater. 2018, 30, 1703737.

[24] Son, D.Y.; Im, J.H.; Kim, H.S.; Park, N.G. 11% efficient perovskite solar cell based on ZnO nanorods: An effective charge collection system. J. Phys. Chem. C 2014, 118, 16567–16573. [CrossRef]

[25] Tseng, Z.L.; Chiang, C.H.; Wu, C.G. Surface engineering of ZnO thin film for high efficiency planar perovskite solar cells. Sci. Rep. 2015, 5, 13211.

[26] Liu, G.; Zhong, Y.; Mao, H.; Yang, J.; Dai, R.; Hu, X.; Chen, Y. Highly efficient and stable ZnO-based MA-free perovskite solar cells via overcoming interfacial mismatch and deprotonation reaction. Chem. Eng. J. 2022, 431, 134235.

[27] Wu, X.; Zhang, J.; Qin, M.; Liu, K.; Lv, Z.; Qin, Z.; Lu, X. ZnO electron transporting layer engineering realized over 20% efficiency and over 1.28 V open-circuit voltage in all-inorganic perovskite solar cells. EcoMat 2022, 4, e12192.

[28] 29. Ong, S.P.; Richards, W.D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V.L.; Persson, K.A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. Comput. Mater. Sci. 2013, 68, 314–319.

[29] Ward, L.; Dunn, A.; Faghaninia, A.; Zimmermann, N.E.R.; Bajaj, S.; Wang, Q.; Montoya, J.; Chen, J.; Bystrom, K.; Dylla, M.; et al. Matminer: An Open Source Toolkit for Materials Data Mining. Comput. Mater. Sci. 2018, 152, 60–69.

[30] Barupal, D.K.; Fiehn, O. Generating the Blood Exposome Database Using a Comprehensive Text Mining and Database Fusion Approach. Environ. Health Perspect. 2019, 127, 2825–2830.

[31] McKinney,W. Data Structures for Statistical Computing in Python. In Proceedings of the Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; Volume 445, pp. 56–61.

[32] Jin, Z.; Shang, J.; Zhu, Q.; Ling, C.; Xie, W.; Qiang, B. RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis. In Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Berlin/Heidelberg, Germany, 2020; Volume 12343, pp. 503–515; ISBN 9783030620073.

[33] Bendib, T.; Bencherif, H.; Abdi, M.A.; Meddour, F.; Dehimi, L.; Chahdi, M. Combined Optical-Electrical Modeling of Perovskite Solar Cell with an Optimized Design. Opt. Mater. 2020, 109, 110259.