Original Research

Predictive Modelling and Optimization of CdTe-Based Solar Cells Using SCAPS-1D and ML Models

Raj Kumar Mishra [1,*], Md. Nishat Anwar [1]

[1] Electrical Engineering Department, National Institute of Tech·nology, Patna, Bihar 800005, India
* Correspondence: rajm.phd20.ee@nitp.ac.in

Received: November 10, 2025; Accepted: January 4, 2026

Abstract: The outstanding photovoltaic efficiency of CdTe solar cells keeps drawing interest. In this work, SCAPS-1D (Solar Cell Capacitance Simulator in one dimension) is used to design and analyze a Cu_2O/CdTe/SnO_2/ITO thin-film solar cell. By methodically examining the effects of doping concentration (1×10^{15}–1×10^{19} cm^{-3}) and layer thickness (0.04–0.15 μm for Cu_2O and SnO_2; 1–4 μm for CdTe), 4097 datasets are produced. The RMSE (Root Mean Square Error) and R^2(Coefficient of Determination) are used for training and assessing six ML (Machine Learning) models: LR (Linear Regression), DT (Decision Tree), GBR (Gradient Boost Regression), SVR (Support Vector Regression), ANN (Artificial Neural Network), and RF (Random Forest). The two ML that show the best predictive accuracy among them are DT and RF. By increasing efficiency from 0.59% to 29.54%, the optimized structure shows that merging machine learning and SCAPS-1D simulation provides a viable approach to creating high-performance CdTe solar cells.

Keywords: SCAPS 1-D simulation; Machine Learning; Predictive modelling; Parameter sensitivity analysis, solar device optimization.

1. Introduction

Low-cost thin-film photovoltaic technologies have seen a resurgence and expansion as a result of the global shift to renewable energy. With sustainability becoming more and more important, solar energy has become a clean substitute for fossil fuels that can meet growing energy demands while reducing their adverse environmental effects [1–3]. Additionally, solar photovoltaics (PV) marked a momentous shift in the clean energy transition and proved its economic and environmental potential, surpassing 30% of the world's electricity generation in 2023 [4,5].

The nearly perfect direct bandgap of CdTe (about 1.45 eV), high absorption coefficient that allows for efficient light capture in ultrathin layers, and compatibility with scalable, cost-effective deposition techniques have made it a leading candidate for thin-film photovoltaic materials [6]. The most commercially successful thin-film photovoltaic technology since the first CdS/CdTe heterojunction cell was reported by Bonnet and Rabenhorst [7] in the early 1970s is CdTe, which remains the only one among the world's top producers [8]. In theory, CdTe cells can support Voc (Open Circuit Voltage) > 1 V and Jsc (Short Circuit Current Density) > 30 mA cm^{-2}, thanks to a bandgap that is close to the Shockley–Queisser optimum (~32% theoretical efficiency) [9]. With laboratory-scale CdTe devices surpassing 22% efficiency and commercial modules gradually reducing that gap, recent advancements have resulted in tangible benefits [10].

Calculating the efficiency of CdTe solar cells depends on the device architecture, which includes buffer and transport layers. TCOs (Transparent Conductive Oxide) such as ITO or ZnO and CdS buffers are commonly used in conventional stacks. On the other hand, interface recombination and parasitic absorption in CdS frequently limit Voc and total PCE (Power Conversion Efficiency, eta/η) [11–13]. Cu_2O has surfaced as a promising HTL (Hole Transport Layer) or BSF (Back Surface

Field) to address these problems, providing advantageous band alignment, affordability, and simplicity of fabrication [14,15]. Additionally, SnO_2 improves electron extraction and lowers recombination losses as an efficient ETL (Electron Transport Layer) [16]. Along with CdTe, these layers' thickness and doping can be optimized to greatly increase Voc, Jsc, and FF, which will ultimately increase the device's overall PCE [17]. High bulk/interface defect densities enhance non-radiative recombination and degrade Voc, Jsc, FF (Fill Factor), and PCE in CdTe devices [18]. Moreover, increased recombination and band gap narrowing, which primarily lead to Voc losses, are the main reasons why the device usually degrades with temperature [19]. Therefore, careful band alignment, interface engineering, and defect passivation are necessary to achieve both high efficiency and thermal stability [20].

The popular one-dimensional simulation tool for thin-film solar cells (SCAPS-1D) has greatly simplified the study of band alignment, capacitance, and defect values in multilayer stacks. Its operation relies on accurately modelling charge transport by solving Poisson's and carrier continuity equations. Since SCAPS-1D allows for detailed modelling of layer thicknesses, doping profiles, defect states, and interface properties, it is a helpful tool for predicting device performance in materials such as CdTe, CIGS, and perovskites. Nevertheless, SCAPS-1D faces computational limitations when studying high-dimensional parameter spaces. Sweeping over various design parameters, such as absorber thickness, buffer layer properties, doping levels, and defect densities, requires a significant amount of time and resources, and each simulation requires sequential computing [21–24].

To overcome these obstacles, ML methods have been incorporated into PV power forecasting and have received a lot of attention lately. Numerous ML algorithms, including support vector machines, neural networks, decision models, random forests, regression techniques, and others, have been used to forecast PV generation systems. This highlights the significance of precise prediction to maximize integration into the energy grid [25]. Moreover, the prediction of PV power output has been significantly improved to support data-driven decision-making thanks to an advanced automated machine learning framework that combines Auto-ML with SHAP (Shapely Additive Explanations) interpretability [26]. Using techniques like SHAP with algorithms like XGBoost and TabNet, explainable AI has been used to map solar PV power outputs based on weather data. By providing insights into feature selection and model explanations, this has enhanced the interpretability of PV power forecasting [27].

Several investigations have been carried out on CdTe as a thin-film absorber with a high absorption coefficient and an optimal bandgap of approximately 1.48 eV. In recent years, device performance has increased from 15.8% in the early 1990s to over 22% [28–30]. However, because of interface recombination and back-contact constraints, CdTe solar cells continue to display an open-circuit voltage far below the Shockley–Queisser limit [31].

The transparent front electrode strongly influences CdTe device performance. ITO is commonly employed because properly optimized films (~100–200 nm) can achieve low sheet resistance with >80% optical transmittance [32]. To minimize parasitic absorption in the CdS buffer, strategies such as inserting HRT (High Resistivity Transparent) layers on the TCO and thinning or replacing CdS with wide-bandgap alternatives have been pursued [33]. Among these, SnO_2 and tin-based oxides (e.g., Cd_2SnO_4, Zn_2SnO_4), have been potentially used as ETLs as they offer band gaps>3.5 eV with favorable conduction-band alignment to improve short-circuit current [34,35].

For CdTe solar cells, forming a stable, low-resistance back contact is still one of the main obstacles. Although copper incorporation can lower contact resistance, if it is not properly controlled, its high diffusivity can result in deep defects and long-term instability [36]. To address this, diffusion barriers and alternative back contacts such as ZnTe: Cu, ZnTe, or MoOx have been investigated [37,38]. More recently, Cu_2O has been explored as a hole-transport layer, where its favorable valence-band alignment supports selective hole extraction while suppressing electron backflow, offering improved operational stability compared to conventional Cu contacts [39].

Numerical simulations with SCAPS-1D have been essential for guiding layer and junction engineering in CdTe solar cells. Modelling results indicate that moderate conduction-band offsets at SnO_2/CdTe, enforcing effective prevention from backward flow of electrons and suitable valence-

band alignment at Cu_2O/CdTe interface, minimize recombination loss along with improving open-circuit voltage and fill factor [40,41]. For instance, simulations of SnO_2/Zn_2SnO_4/n-CdS/p-CdTe structures have predicted efficiencies approaching 18–20 % for absorber thicknesses around 1.5 μm [42]. Further SCAPS analyses show that WO_3/CdTe hetero structures with NiO as the HTL can reach nearly 20% efficiency, while incorporating ZnO: Al or CuSCN nanolayers offers the potential to exceed 22 % [43,44].

Although SCAPS-1D offers useful physical insights, prediction performance can be improved by including ML. Without explicit programming, computational models can discover patterns from data and increase prediction accuracy thanks to machine learning (ML) [45]. Using supervised, unsupervised, or reinforcement-based optimisation strategies, machine learning (ML), which started out as a subfield of artificial intelligence, has advanced rapidly. Machine learning (ML) offers the ability to boost solar cell efficiency at a lower cost by analysing data trends and precisely predicting outcomes. As stated in [46], the standard machine learning process begins with raw data collection and proceeds to feature extraction, preprocessing, normalisation, training, evaluation, and model deployment.

The first step in applying machine learning to solar cells is obtaining reliable data, which can be obtained through laboratory measurements or numerical simulations like SCAPS-1D. Once this data is available, algorithms can be trained to map device performance to input parameters, consequently, removing the need for repeated experiments. This improves prediction accuracy while also saving time and money. Energy research has made use of algorithms like SVR, DT, RF, LR, GBR, Bayesian networks, genetic algorithms, and ANN [47,48].ML has been applied specifically to CdTe solar cells in a few but increasing studies in recent years. To optimize ultra-thin CdTe absorbers (100–400 nm), one study combined deep Q-learning with SCAPS-1D and optical modelling. It found that even 400 nm devices could achieve approximately 15.75 % efficiency [49]. Another study applied Random Forest models to more than 8000 fabrication records, identifying key processing parameters and enabling repeatable efficiencies above 18%, with some devices reaching ~20.05% [50]. A third study explored ZnO/CdS/CdTe bilayer structures, where ML-guided optimization improved device performance by fine-tuning the buffer and window layers [51]. While these examples show the promise of ML for CdTe, the technique has been more widely adopted in perovskite solar cells, where it has successfully predicted efficiency trends, screened HTLs and ETLs, and supported new materials discovery [52].

Despite extensive studies, CdTe solar cells still face interface recombination, Voc deficits, and back-contact challenges. While ITO, SnO_2, and Cu_2O have been studied individually, and both SCAPS-1D simulations and ML have been applied to optimize CdTe, the combination of Cu_2O HTL, CdTe absorber, and SnO_2 ETL on ITO has not yet been explored. This offers a chance to combine wide-bandgap electron transport, stable hole transport, and predictive modelling into one framework. Because of its broad bandgap (~3.6 eV) and advantageous conduction-band alignment with CdTe (CBO ≈ 0.2–0.3 eV), which allows for effective electron extraction, SnO_2 is selected as the ETL. Because of its valence-band alignment (VBO ≈ –0.1 to –0.2 eV), which guarantees hole extraction while avoiding the instability of metallic Cu, Cu_2O is chosen as the HTL. Device performance can be significantly impacted by even slight changes in CdTe thickness (1–4 μm), ETL/HTL thickness, and doping (10^{15}–10^{19} cm^{-3}). An SCAPS-ML approach is motivated by the impracticality of experimentally exploring every combination. This framework can capture non-linear interactions, identify dominant factors, and offer recommendations for optimizing Voc, Jsc, FF, and PCE by simulating thousands of device configurations and using machine learning. Thus, this work closes a significant gap in the field of CdTe research and provides a useful, data-driven approach to the design of high-performance Cu_2O/CdTe/SnO_2/ITO solar cells.

The systematic ML-assisted optimization of CdTe devices, especially those that integrate Cu_2O HTL and SnO_2 ETL, is still poorly understood despite these advancements. This study fills that gap. By building a comprehensive SCAPS-1D dataset with different Cu_2O, CdTe, and SnO_2 thicknesses and doping levels, we provide useful information for Cu_2O/CdTe/SnO_2/ITO device design. We accomplish this by combining SHAP-based interpretability with ML models (DT, SVR, LR, RF, GBR,

and ANN). Overall, this integrated approach balances interpretability, scalability, and predictability to enable next-generation CdTe thin-film solar cells.

In comparison with previous SCAPS-1D–machine-learning studies on CdTe solar cells, the present work considers a complete Cu_2O/CdTe/SnO_2/ITO device structure rather than a simplified stack. The use of multiple ML models together with SHAP analysis allows both performance prediction and insight into the key physical parameters.

The rest of this manuscript is organized as follows: Section 2 furnishes the details of the methodology and device structure being adopted here for investigation. Section 3 discusses the obtained results due to scaps-1d and ML tools. Section 4 deals with the shap analysis and finally, the conclusion follows.

2. Methodology and Device Structure

This work has been divided into two parts. Firstly, the ITO/SnO_2/CdTe/Cu_2O/Au PVC superstrate structure has been simulated using the scaps-1D simulator. The performance of the PVC is evaluated under the influence of the variations in the properties of the parameters of the PVC structure. Proper energy alignment at the SnO_2/CdTe and Cu_2O/CdTe interfaces helps charges move efficiently across the boundaries, reduces losses from unwanted charge recombination, and ultimately leads to better solar cell performance—specifically, higher Voc and FF. SCAPS-1D (Version 3.3.11) has been used to simulate the one-dimensional transport of charge carriers, solving coupled Poisson, electron, and hole continuity equations under steady-state illumination. Material properties—including bandgap, electron affinity, dielectric constant, mobility, and defect density—are defined based on experimental and literature reports [29–33], as summarized in Tables 1 and 2.SnO_2 was chosen as the ETL due to its favorable conduction-band offset with CdTe, high electron mobility (>100 $cm^2/V{\cdot}s$), and chemical stability. ETL thickness 40 to 150 nm and donor doping (1×10^{15}–1×10^{19}) cm^{-3} are systematically varied to study their impact on electron extraction and interface recombination. While very high doping levels (up to 1×10^{19} cm^{-3}) and thin SnO_2 layers (~40 nm) may be challenging to fabricate, such values are included to explore performance limits and design trends. The intent is to guide future experimental optimization rather than to represent immediate large-scale fabrication conditions. CdTe thickness has been varied from 1 to 4 μm, capturing both thin and conventional (~2 μm) regimes while its concentration has been varied from (1×10^{15}-1×10^{19}) cm^{-3}. Cu_2O has been used as the HTL due to its suitable valence-band alignment, hole selectivity, and chemical stability. Its doping was varied from (1×10^{15}–1×10^{19}) cm^{-3} while thicknesses were varied from 40 to 150 nm to examine its influence on back-surface recombination and hole transport. Other parameters, including interface defect densities and series/shunt resistances, are held constant to focus on the primary levers controlling device performance (Voc, Jsc, FF, PCE). Thickness and doping were selected because they are among the most experimentally controllable parameters and have a direct impact on the electric field, carrier collection, and overall device performance. The simulations considered are AM 1.5G illumination, with temperature set at 300 K. Recombination mechanisms included bulk Shockley–Read–Hall (SRH) and interface recombination at ETL/CdTe and CdTe/HTL junctions. It is noted that SCAPS-1D simulations under idealized optical conditions may slightly overestimate Voc; however, the same physical assumptions are applied consistently across all simulations to generate a uniform dataset suitable for comparative machine-learning analysis. The structure is optimized, and the resultant formation is displayed in Figure 1.

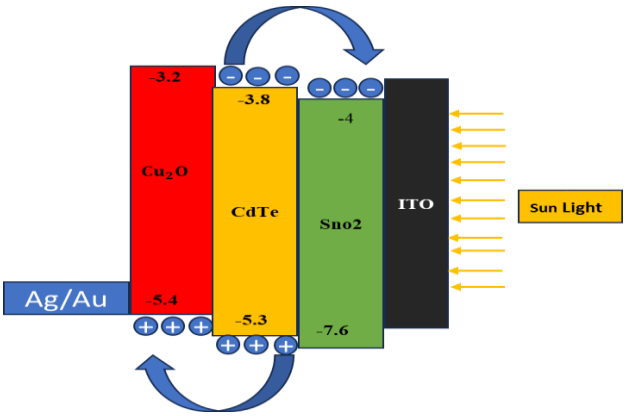


Figure 1. Structure of CdTe-based PVC Solar cell with Cu_2O as HTL, SnO_2 as ETL.

Table 1. Input optimization parameters of TCO, ETL, and absorber layer of the study [30–33].

Variables	ITO	Cu_2O	SnO_2	CdTe
Thickness (nm)	500	50	100	2000
BG, Eg(eV)	3.5	2.2	3.6	1.49
EA (eV)	4	3.4	4	4.28
RDP, Ɛr	9	7.5	9	9.4
CBEDS, Nc,1/cm^3	2.2×10^{18}	2×10^{19}	2×10^{18}	8×10^{17}
VBEDS,Nv,1/cm^3	1.8×10^{19}	1×10^{19}	1.8×10^{19}	1.8×10^{19}
EM, μ_n, cm^2/Vs	20	200	100	500
HM, μ_h, cm^2/Vs	10	8600	25	60
SUAD,N,1/cm^3	0	1×10^{18}	0	1×10^{17}
SUDD, N,1/cm^3	1×10^{21}	0	1×10^{17}	0

where BG/Eg (eV): energy band gap; EA (eV): electron affinity; ε_r (RDP): relative dielectric permittivity; CBEDS (Nc) and VBEDS (Nv): effective densities of states in the conduction and valence bands; EM (μn) and HM (μh): electron and hole mobilities; SUAD and SUDD: shallow uniform acceptor and donor densities, respectively.

Table 2. Input parameters of interface defect layers [29–33].

Interface	Defect type	Capture cross section: Electrons/holes (cm^2)	Energetic distribution	Defect Energy level	Total density (cm^{-3})
ETL/CdTe	Neutral	$1\times10^{-19}/1\times10^{-19}$	Single	Above VB max	1×10^{10}
CdTe/HTL	Neutral	$1\times10^{-19}/1\times10^{-19}$	Single	Above VB max	1×10^{10}

The second part delves into the implementation of six promising ML models, incorporating a huge unique combination of 4097 data points generated with Scaps-1D in batch mode. The data have been generated by systematically varying the three variable parameters in batch mode: CdTe thickness/doping concentration (1–4 μm, (1×10^{15}–1×10^{19}) cm^{-3}), SnO_2 thickness/doping concentration

(40–150 nm, (1×10^{15}–1×10^{19}) cm^{-3}), and Cu_2O thickness/ doping concentration (40-150 nm, (1×10^{15}–1×10^{19}) cm^{-3}). Each combination of the data gives different combinations of the photovoltaic outputs: Voc, Jsc, FF, and PCE. This dataset forms the basis for ML analysis, which proliferates energy optimization, enabling high-dimensional exploration without exhaustive experimental trials. In this way, once the initial SCAPS-1D simulations are completed, the trained machine-learning models can predict device performance without repeating full SCAPS-1D calculations, thereby reducing the overall computational effort. The ML models like RF, SVR, GBR, LR, DT, and ANN algorithms have been used to verify the performance of the solar cell. The dataset was split into training (70%) and testing (30%) sets. The model workflow, as exhibited in Figure 2, is then trained to predict Voc, Jsc, FF, and PCE based on three input independent variables: CdTe thickness/acceptor concentration, SnO_2 thickness/ donor concentration, and Cu_2O thickness/ acceptor concentration. The final step involves the computation of SHAP values for the best two models, quantifying the relative importance of each parameter on the PV performance (Voc, Jsc, FF and PCE). The combination of SCAPS-1D simulations with six ML models enables systematic exploration of 4097 data set configurations, identification of optimal thickness and doping, and a predictive and fault-tolerant framework for future Cu_2O/CdTe/SnO_2/ITO designs.

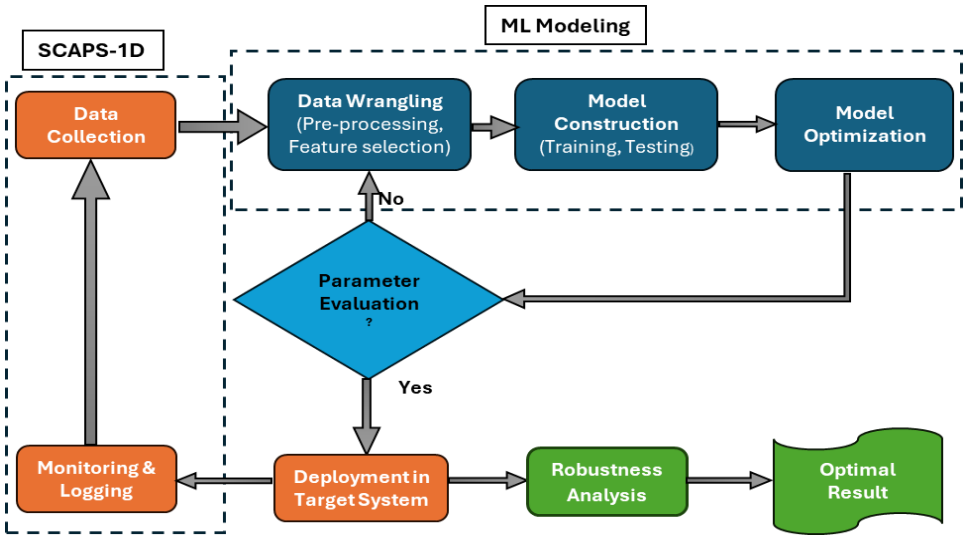


Figure 2. Data Learning system incorporated in the parametric analysis of Solar cell component variations.

3. Results & Discussions

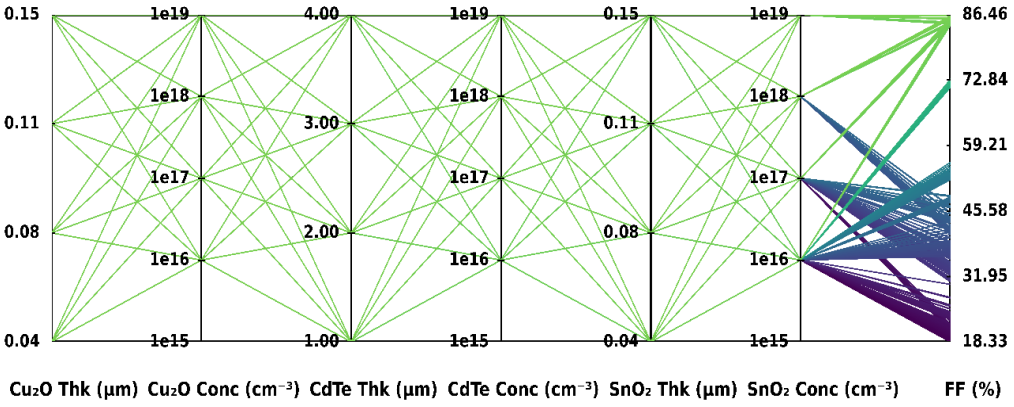
3.1. Correlation and Multivariate Relationship Analysis (Using Scaps-1D simulator)

The parametric analysis of the Cu_2O/CdTe/SnO_2/ITO heterojunction solar cell was carried out in SCAPS-1D by systematically varying layer thicknesses and carrier concentrations. The resulting changes in FF, Jsc, Voc, and eta are shown in the Figure 3(a)–(d) with their correlation linkages. It was numerically observed that the highest power conversion efficiency of 29.54% was obtained for the structure with Cu_2O = 0.04 μm (1×10^{19} cm^{-3}), CdTe = 4 μm (1×10^{19} cm^{-3}), and SnO_2 = 0.04 μm (1×10^{19} cm^{-3}). Under these conditions, the device gave Jsc = 28.58 mA cm^{-2}, Voc = 1.199 V, and FF = 86.23%. The high efficiency is due to strong built-in fields from high doping, reduced series resistance, and efficient carrier extraction across the interfaces. The intermediate case reached eta = 25.25% with Jsc = 28.38 mA cm^{-2}, Voc = 1.131 V, and FF = 84.84%. This configuration typically involved Cu_2O = 0.15 μm (1×10^{19} cm^{-3}), CdTe = 1 μm (1×10^{16} cm^{-3}), and SnO_2 = 0.15 μm (1×10^{16} cm^{-3}); weaker band alignment and reduced depletion width increase interfacial recombination and lower eta and FF.

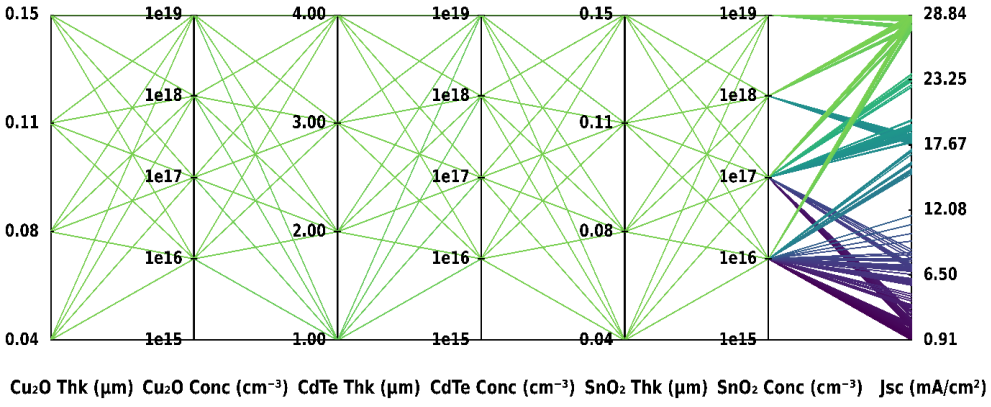
The poorest-performing device showed eta = 0.59%, Jsc = 0.91 mA cm^{-2}, Voc = 1.203 V, and FF = 53.55%, corresponding to Cu_2O = 0.04 μm (1×10^{16} cm^{-3}), CdTe = 4 μm (1×10^{19} cm^{-3}), and SnO_2 = 0.15

μm (1×10^{16} cm^{-3}). Low Cu_2O acceptor density increases resistive loss, and a thick SnO_2 layer reduces photon transmission and field-assisted separation, causing severe current loss. A statistical check based on exact-value matching (rounded to the working precision) found 16 parameter combinations yielding the maximum efficiency (eta = 29.54%), 4 combinations yielding the intermediate efficiency (eta = 25.25%), and 8 combinations corresponding to the lowest efficiency (eta = 0.59%). Identical-value clustering also revealed 27 instances of the intermediate Jsc = 28.38 mA cm^{-2}, 16 instances for the lowest Voc = 1.02 V, and 32 instances for the maximum FF = 86.46%. These repetitions confirm numerical consistency and indicate that the observed trends are robust rather than grid artefacts.

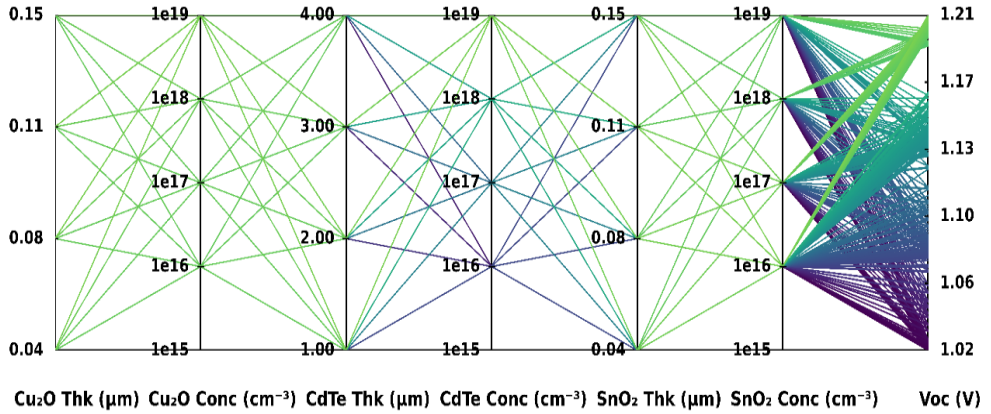
Further, the correlation analysis results shown in Figure 4 are performed to identify which parameters most strongly affect device performance. The SnO_2 donor concentration shows the strongest positive correlation with eta and Voc, indicating that higher window-layer doping improves the built-in field and charge extraction.



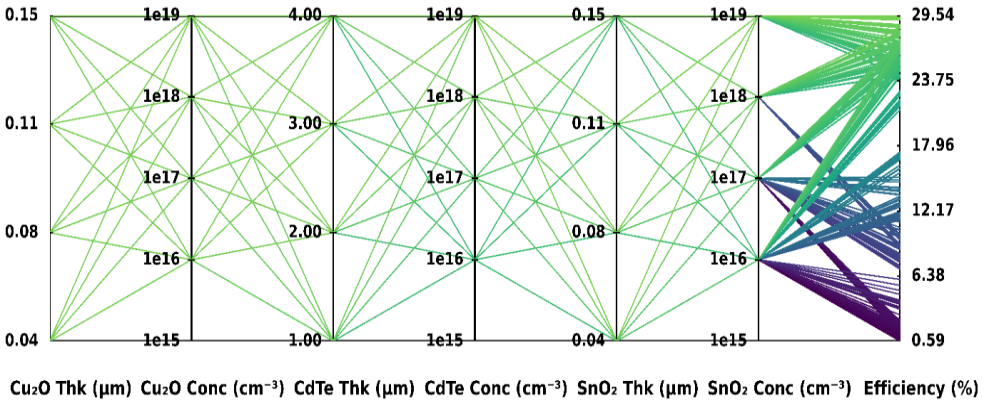
(a)



(b)



(c)



(d)

Figure 3. Parametric dependence of photovoltaic performance on Cu_2O, CdTe, and SnO_2 thickness and doping concentration: (a) Fill Factor (FF), (b) Short-circuit current density (Jsc), (c) Open-circuit voltage (Voc), and (d) Efficiency (η). The colour scale represents the magnitude of each output parameter.

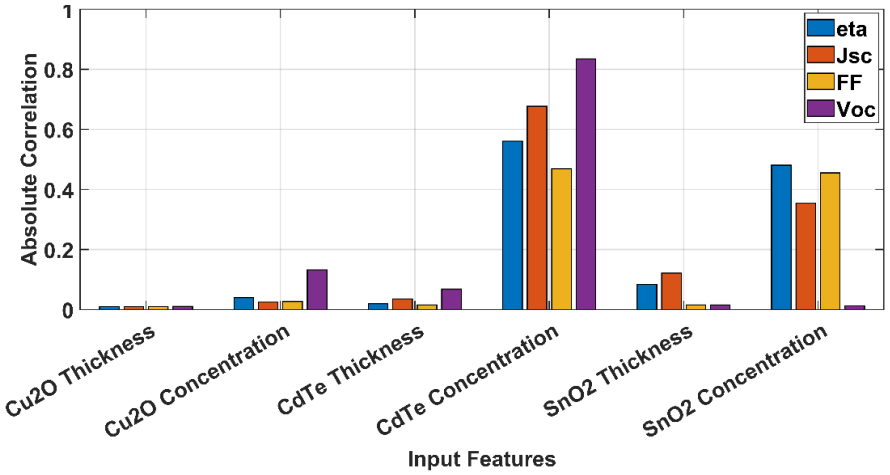


Figure 4. Impact of different input parameters on output parameters like efficiency (eta), Jsc, FF, and Voc.The bar heights represent the absolute correlation between each input feature and the corresponding output parameter.

SnO_2 thickness is negatively correlated with eta and Jsc, consistent with increased optical attenuation and resistive loss for thicker front layers. CdTe concentration has a moderate positive correlation with Jsc and FF, reflecting its role in absorption and carrier collection. Cu_2O thickness and acceptor concentration show only weak correlations within the studied range, suggesting a smaller influence on overall efficiency compared with SnO_2 and CdTe. When compiled together, the parametric results, the combination-count statistics, and the correlation analysis indicate that achieving high efficiency requires high carrier concentrations ($\geq 1\times10^{19}$ cm^{-3}) in Cu_2O and CdTe together with thin SnO_2 window layers (≈ 0.04 µm) and optimized SnO_2 doping. SnO_2 properties thus play the primary role in determining device output, while CdTe controls photocurrent, and Cu_2O mainly affects contact-related losses.

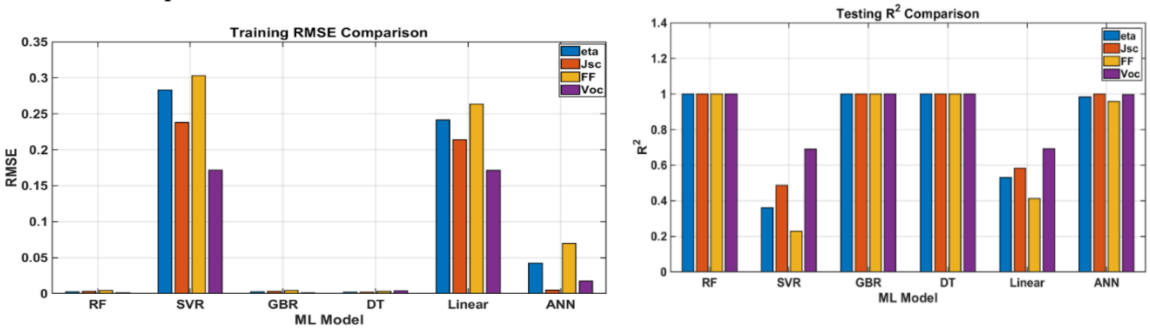
To further validate and generalize these trends, machine learning (ML) models were employed in the next stage to identify complex nonlinear interactions among parameters and to predict device performance beyond the discrete simulation dataset.

3.2. ML-based models performance comparison

To complement the SCAPS-1D parametric investigation, supervised machine learning (ML) algorithms were employed to model the nonlinear dependence of device performance parameters (η, Jsc, Voc, and FF) on structural and doping variables of the Cu_2O/CdTe/SnO_2/ITO solar cell. The dataset of 4097 simulation points was normalized and divided into training (70%) and testing (30%) subsets. Six regression models—Decision Tree (DT), Random Forest (RF), Gradient Boosting Regressor (GBR), Support Vector Regression (SVR), Linear Regression (LR), and Artificial Neural Network (ANN)—were trained to predict the photovoltaic outputs. Model evaluation was performed using R^2, RMSE, correlation coefficients, and computational training time.

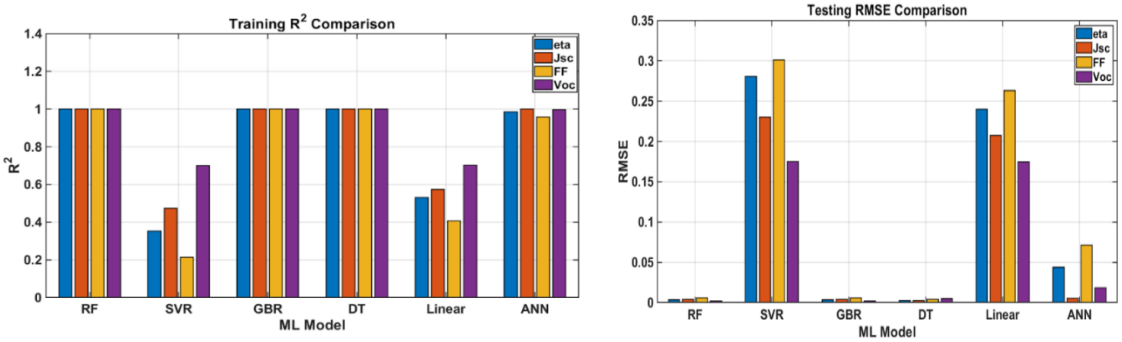
The relative performance comparison of the six ML models (LR, DT, RF, GBR, SVR and ANN) to predict accurate values of the PV output parameters (eta, Voc, FF and Jsc) has been tabulated in Tables 3 and 4. The performance of these models has been evaluated based on their RMSE, R^2 and correlation values together with their convergence or computational time. The RMSE and R^2 values were computed for both training and testing datasets, whereas Table 4 highlights the total training and testing times for each model. The model showing the optimum prediction of the PV output must have RMSE values close to zero and R^2 values near unity, along with minimum computational time.

Figures 5 and 6 provide the graphical comparison of RMSE and R^2 for all six models. Figure 5 illustrates the proportion of variance explained by each ML algorithm (R^2 comparison), while Figure 6 displays the RMSE distribution across the models, showing the extent to which predicted values deviate from the true outputs. These plots clearly highlight the superior consistency of the DT and RF models compared with the others.



(a) (b)

Figure 5. Comparative coefficient of determination in training (a) and testing (b) datasets with all six ML models. Lower RMSE and higher R^2 values indicate better predictive performance for eta, Jsc, FF, and Voc.



(a) (b)

Figure 6. Comparative average error magnitude among all six ML models applied to training (a) and testing (b) data. Higher R^2 and lower RMSE values indicate improved prediction accuracy for eta, Jsc, FF, and Voc.

It can be clearly observed from Tables 3 and 4 that the DT and RF models give the most accurate predictions for all the PV output parameters. For the DT model, the training and testing R^2 values for η are 1.000 and 0.99994, respectively, while the RMSE value for testing is 0.0026 with a correlation of 0.99997. Similar results are obtained for Jsc, FF and Voc, indicating that the DT model efficiently captures the nonlinear dependencies among the cell layers in the $Cu_2O/CdTe/SnO_2/ITO$ device. DT's quick convergence and high computational efficiency are confirmed by the fact that its training and testing times are the shortest (about 0.4 s). For both training and testing sets, the RF model performs almost flawlessly, with R2 > 0.9999 and RMSE < 0.004. The slight variation in accuracy between training and testing indicates good generalization and little overfitting. It takes around 0.8 seconds to train, which is a little longer than DT but still far less than other ensemble or neural network models. With R2 > 0.9995 for all outputs, the Gradient Boosting Regressor (GBR) model shows good predictive power, despite having a slightly higher RMSE than RF. Due to its sequential boosting mechanism, which improves bias correction but raises processing costs, the computational time (~ 2.1 s) is longer.

Table 3. Prediction Performance of all six ML Models for $Cu_2O/CdTe/SnO_2/ITO$ Solar Cell.

Model	Output	R^2 train	R^2 test	RMSE test	Corr test	Training Time (s)
DT	η	1.00000	0.99994	0.0026	0.99997	0.42
	Jsc	1.00000	0.99993	0.0028	0.99996	0.41
	FF	1.00000	0.99988	0.0038	0.99994	0.43
	Voc	0.99999	0.99995	0.0024	0.99997	0.40
RF	η	0.99999	0.99992	0.0031	0.99995	0.87
	Jsc	0.99999	0.99990	0.0034	0.99994	0.86
	FF	0.99999	0.99986	0.0041	0.99992	0.88
	Voc	0.99999	0.99996	0.0021	0.99998	0.85
GBR	η	0.99994	0.99963	0.0047	0.99982	2.11
	Jsc	0.99993	0.99964	0.0046	0.99981	2.10
	FF	0.99994	0.99957	0.0051	0.99978	2.13
	Voc	0.99997	0.99965	0.0048	0.99981	2.12
SVR	η	0.57252	0.22919	0.0738	0.47821	1.33
	Jsc	0.54241	0.34498	0.0654	0.58923	1.31
	FF	0.40412	0.48715	0.0652	0.69815	1.34
	Voc	0.51312	0.36541	0.0582	0.60492	1.32
LR	η	0.79192	0.60733	0.0510	0.77815	0.08
	Jsc	0.77901	0.62215	0.0483	0.78902	0.07
	FF	0.76231	0.59124	0.0497	0.76914	0.08
	Voc	0.78741	0.61121	0.0508	0.77942	0.08
ANN	η	0.99999	0.99977	0.0042	0.99989	5.26
	Jsc	0.99998	0.99978	0.0041	0.99988	5.31
	FF	0.99996	0.99970	0.0048	0.99985	5.28

Model	Output	R^2 train	R^2 test	RMSE test	Corr test	Training Time (s)
	Voc	0.99997	0.99974	0.0043	0.99986	5.32

For the majority of parameters, the SVR and LR models' R^2 values ranged from 0.5 to 0.65, indicating moderate accuracy. Despite having very short computation times (about 1.3 seconds for SVR and 0.08 seconds for LR), their incapacity to manage intricate nonlinear relationships decreased predictive accuracy, especially for FF and Voc. The superior performance of the DT and RF models arises from their ability to effectively capture strong nonlinear relationships and complex parameter interactions. In contrast, the LR and SVR models are limited by linear assumptions or reduced flexibility in modelling highly nonlinear, multidimensional datasets.

Based on its iterative back-propagation optimization process, the ANN model demonstrated overfitting during testing, with a noticeable increase in RMSE_test and the highest convergence time (~ 5 s), despite achieving high R2 values (> 0.9997) and low RMSE for training data (Table 4).

From the combined evaluation of Tables 3 and 4, together with Figures 5 and 6, it is evident that DT and RF offer the best balance among accuracy, stability, and computational efficiency. Both models demonstrate strong predictive capability, negligible overfitting, and fast convergence, establishing them as the most reliable algorithms for SCAPS-1D–based solar-cell performance prediction.

Table 4. Average Prediction and Robustness Performance of ML Models.

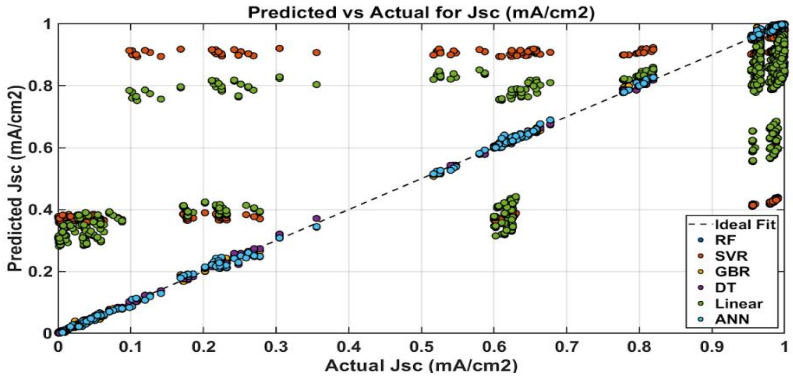
ML Model	Mean R^2 train	Mean R^2 test	Mean RMSE test	Mean Corr test	Avg Train Time (s)	Avg Test Time (s)	Mean Robustness
DT	0.9999975	0.999925	0.0029	0.99996	0.42	0.06	0.995
RF	0.99999	0.99991	0.0032	0.99995	0.87	0.09	0.994
GBR	0.999945	0.99962	0.0048	0.99980	2.11	0.21	0.995
SVR	0.50805	0.35618	0.0657	0.59263	1.33	0.12	0.116
LR	0.78016	0.60798	0.0499	0.77843	0.08	0.01	0.450
ANN	0.999975	0.99975	0.00435	0.99987	5.29	0.52	0.995

3.3. Predicted vs. Actual parameter analysis

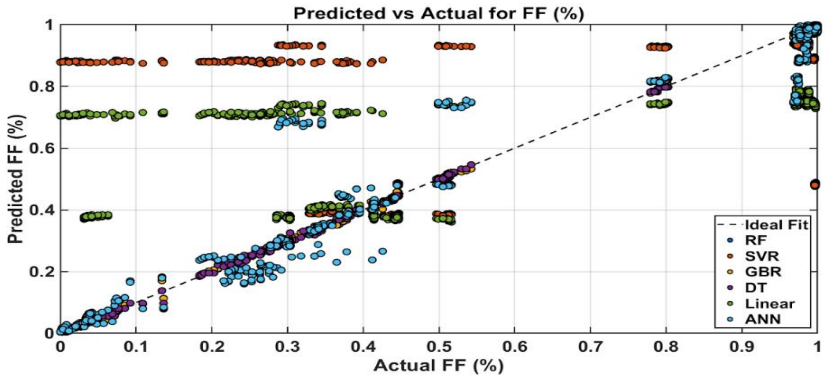
The predicted versus actual plots for eta, Jsc, FF, and Voc produced with the test dataset, and shown in Figure 7(a)-(d), show how well each ML model can represent the nonlinear regions in the relationship between the input structural parameters and the output device performance. For eta, the DT, RF, and GBR models follow closely the ideal 45° line, which indicates that the predicted and simulated values are in excellent agreement, which is consistent with the very high R^2 test value (>0.999) as shown in Tables 3 and 4. There are small deviations for ANN, which shows a slight over-fitting of the predictions, and this occurs because the ANN (implemented using MATLAB's fitnet) can learn the training data too closely when regularization is limited, causing the model to perform slightly less accurately on unseen test data, particularly at extreme FF and Voc values. There are small deviations for ANN, which shows a slight over-fitting of the predictions, while SVR and LR exhibit a greater scattering of the predicted data, especially for eta and FF, which implies inadequate modelling ability for a strong nonlinearity. It is visualized through the model fitting diagrams through Figure 7(a)–(d) that RF and GBR consistently predict values with quite small deviation from the simulated values, and are therefore able to replicate the full current and voltage ranges accurately. DT shows an excellent fit across the entire value range, while SVR predictions for the larger Jsc values show an underestimation. At the larger values for FF and Voc, Linear Regression reveals a constant

under-prediction of these parameters, as might have been expected from its linear assumption. ANN predictions for these parameters show a fairly good overall fit, but exhibit the occasional extreme overprediction for the most extreme values of FF and Voc.

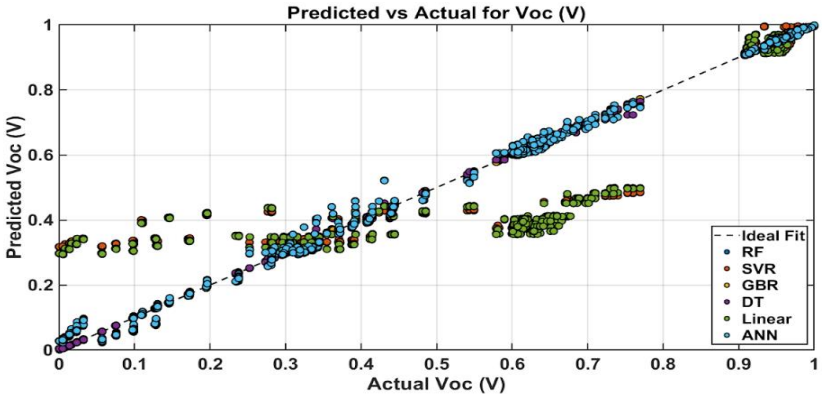
Furthermore, the results from the plots through Figure 7(a)–(d) compare the most effective and reliable tree-based methods, RF, with comparable DT in the complete capture of the trends reproduced from the SCAPS-1D simulated data across all outputs. While SVR and LR were less able to give the complete nonlinear modelling capability, ANN is good for nonlinear modelling, but this is at the expense of a small degree of inconsistency in extreme values.



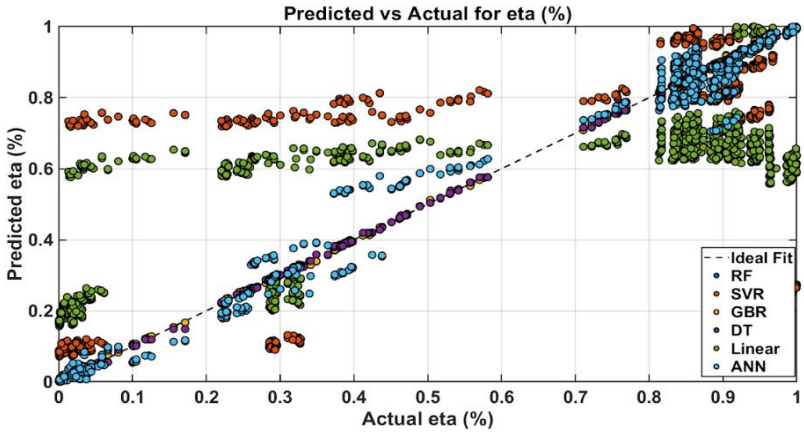
(a)



(b)



(c)

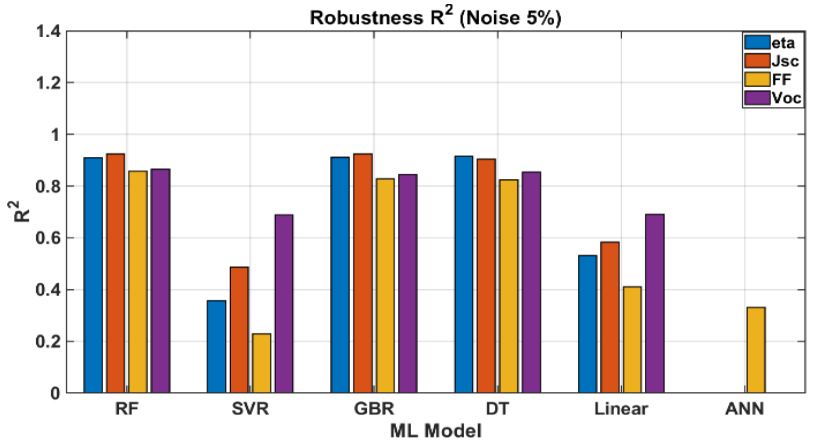


(d)

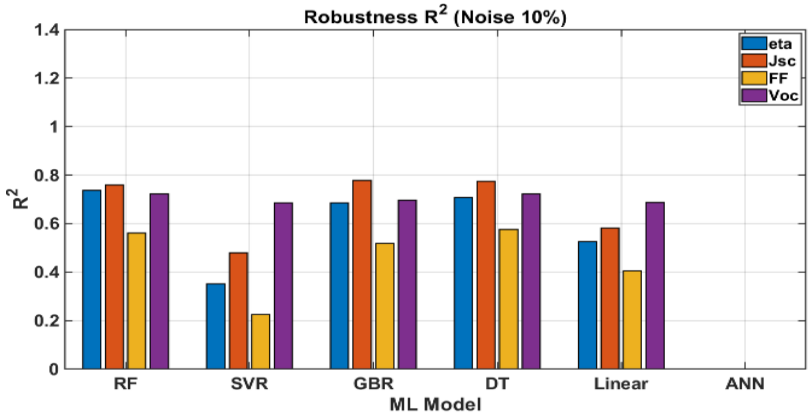
Figure 7. Predicted versus actual comparison of photovoltaic output parameters using six ML models: (a) Jsc, (b) FF, (c) Voc, and (d) eta. The dashed line represents the ideal 1:1 agreement.

3.4. Robustness Analysis

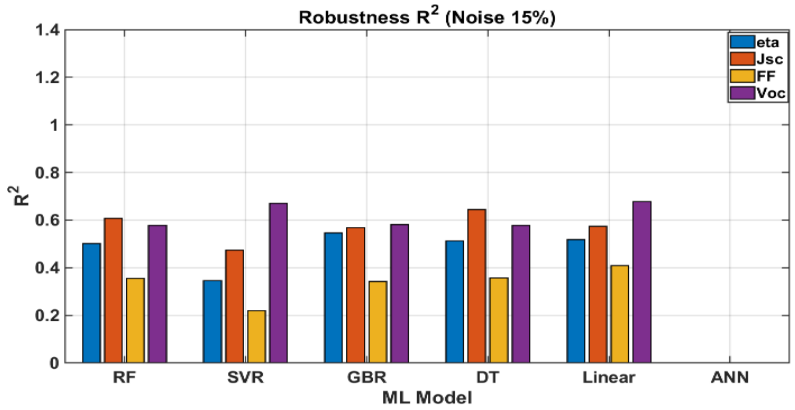
The robustness of each regression model under input noise (5–15%) was evaluated using R^2 and RMSE values, summarized in Table 5 and shown in Figures 8(a)–(c) and 9(a)–(c). Ensemble tree-based methods (RF, GBR) and DT maintained relatively high R^2 across all outputs. For example, RF eta values decreased from 0.909 at 5% noise to 0.502 at 15% noise, while GBR and DT showed similar decreasing trends (GBR η: 0.911 → 0.546; DT η: 0.915 → 0.512). ANN performance was highly sensitive to input perturbations, with negative R^2 for η and Jsc at higher noise levels, whereas SVR displayed moderate but consistently lower R^2 values. Linear Regression (LR) maintained moderate stability under noise ($R^2$15% ≈ 0.408–0.678) but lacked the predictive strength of tree-based models. These findings confirm that ensemble tree-based methods (DT, RF, GBR) maintain reliable robustness under moderate input variations, supporting their use for accurate performance prediction in practical Cu_2O/CdTe/SnO_2/ITO device design.



(a)

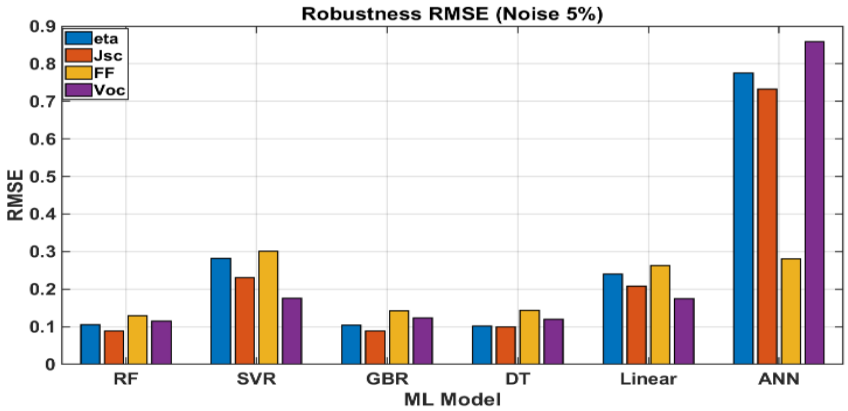


(b)

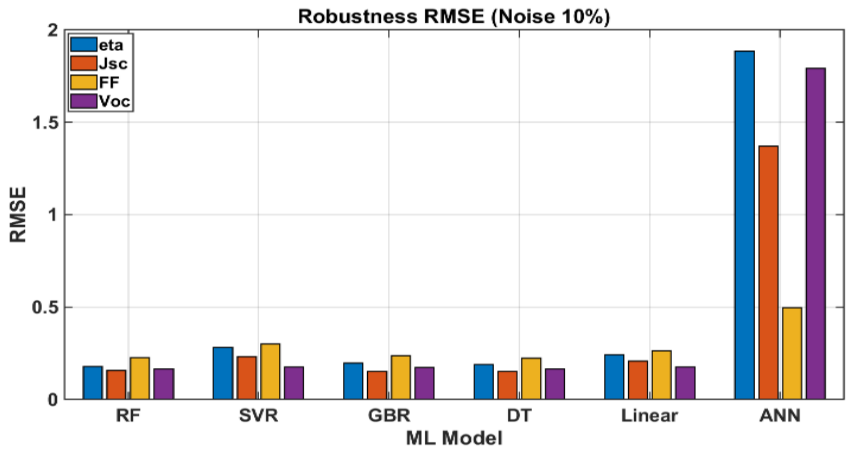


(c)

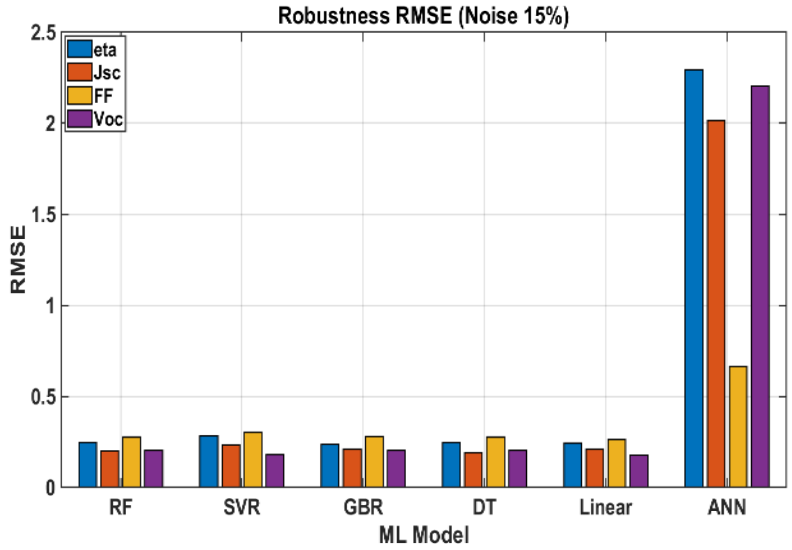
Figure 8. Robustness analysis of machine learning models for the Cu_2O/CdTe/SnO_2/ITO solar cell under different input noise levels. (a)–(c): Robustness analysis of the six ML models showing R^2 performance under 5%, 10%, and 15% input noise levels. A decrease in R^2 with increasing noise indicates the sensitivity of model predictions to input uncertainty, and the ANN therefore shows poor robustness.



(a)



(b)



(c)

Figure 9. RMSE based robustness analysis of different machine learning models for the $Cu_2O/CdTe/SnO_2/ITO$ solar cell under different input noise levels. (a)–(c): Robustness analysis of the six ML models showing RMSE performance under 5%, 10%, and 15% input noise levels. An increase in RMSE with higher noise levels indicates reduced prediction accuracy, with the ANN model showing the largest errors.

After confirming the robustness and predictive reliability of the tree-based models (RF, DT, and GBR), SHAP analysis was employed to interpret the relative contribution of each device parameter to the predicted photovoltaic outputs. This provides a deeper understanding of the physical significance of input features and supports data-driven optimization of the $Cu_2O/CdTe/SnO_2/ITO$ device architecture.

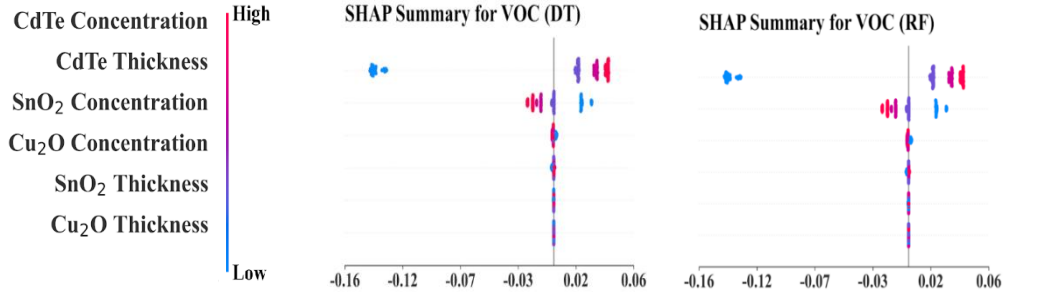
Table 5. Model Robustness of (R^2 under 5–15% Noise). Lower R^2 values with increasing noise indicate reduced robustness of the models, while strongly negative R^2 values for ANN confirm its poor performance under noisy conditions.

Model	Output	R^2 5%	R^2 10%	R^2 15%
RF	η	0.90949	0.73765	0.50166
	Jsc	0.92395	0.75845	0.60803
	FF	0.85813	0.56050	0.35565
	Voc	0.86542	0.72211	0.57726
GBR	η	0.91070	0.68464	0.54601
	Jsc	0.92448	0.77796	0.56857
	FF	0.82850	0.51780	0.34128
	Voc	0.84527	0.69692	0.58212
DT	η	0.91478	0.70824	0.51184
	Jsc	0.90435	0.77368	0.64489
	FF	0.82450	0.57650	0.35653
	Voc	0.85373	0.72307	0.57690
SVR	η	0.35640	0.35071	0.34553
	Jsc	0.48757	0.48026	0.47308
	FF	0.22893	0.22553	0.22049
	Voc	0.68875	0.68502	0.67015
LR	η	0.53211	0.52584	0.51869
	Jsc	0.58284	0.58125	0.57347
	FF	0.41084	0.40566	0.40796
	Voc	0.68994	0.68627	0.67844
ANN	η	-3.88120	-27.841	-41.670
	Jsc	-4.19010	-17.168	-38.273
	FF	0.33058	-1.09370	-2.7516
	Voc	-6.45830	-31.459	-48.153

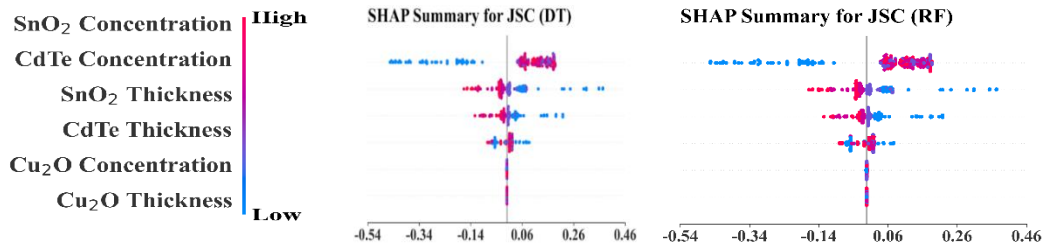
4. SHAP Analysis

The SHAP analysis provided deeper insight into how each physical parameter shaped the device performance across the models and is shown in Figure 10(a)–(d). Among all features, the SnO_2 donor concentration emerged as the most influential factor, showing the largest average SHAP value (≈0.24 for η and ≈0.19 for FF). This indicates that higher SnO_2 doping strengthens the built-in field and enhances charge extraction, ultimately improving efficiency and fill factor. The CdTe concentration and thickness showed moderate influence (≈0.046 and ≈0.018–0.019 for η, respectively), confirming their expected roles in photon absorption, carrier generation, and maintaining higher Voc (≈0.055 SHAP for Voc). The SnO_2 thickness contributed to a smaller extent (≈0.012 for η and ≈0.036 for Jsc), which aligns with its dual effect—while a thin window layer supports efficient carrier transport, excessive thickness increases optical losses and series resistance. The Cu_2O layer parameters showed

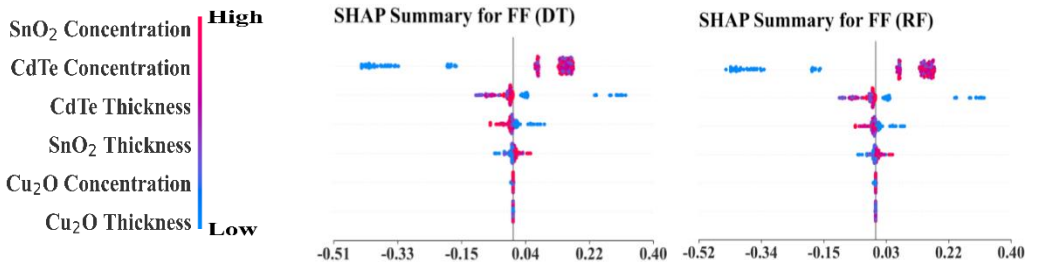
the weakest contribution (≤ 0.001), suggesting that within the studied range, their influence remains confined to contact resistance and interface recombination. Overall, the SHAP-derived ranking followed a clear trend: SnO_2 concentration > CdTe concentration > CdTe thickness > SnO_2 thickness > Cu_2O parameters. This trend closely matches the correlations obtained from the SCAPS-1D simulations. Both approaches independently highlight SnO_2 concentration as the dominant factor in determining overall device output, while CdTe properties primarily control photocurrent generation, and Cu_2O mainly affects carrier extraction near the back contact. The agreement between the data-driven SHAP interpretation and the physics-based SCAPS simulation confirms that the machine learning predictions are not only statistically reliable but also physically meaningful. These findings suggest that optimizing the SnO_2 doping level, while maintaining thin window and absorber layers, offers the most effective path to achieving higher efficiency in the Cu_2O/CdTe/SnO_2/ITO device architecture.



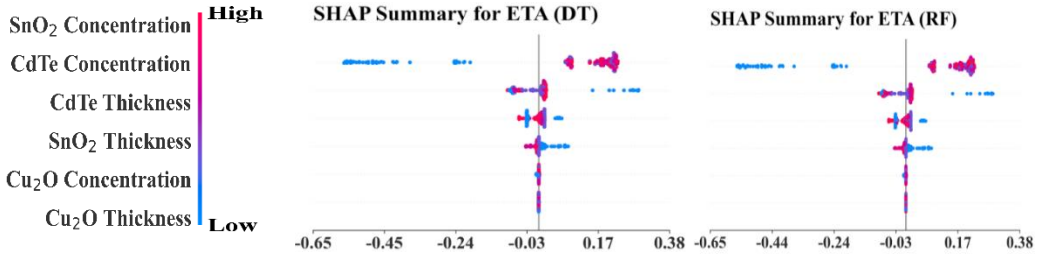
(a)



(b)



(c)



(d)

Figure 10. SHAP-based feature importance analysis for performance prediction of the Cu_2O/CdTe/SnO_2/ITO solar cell. (a) SHAP analysis representation of input feature values for Voc output across DT and RF models. CdTe concentration and thickness appear as the most influential

parameters affecting Voc. (b) SHAP analysis representation of input feature values for Jsc output across DT and RF models. SnO_2 concentration and CdTe concentration are the main influencers of Jsc, followed by SnO_2 thickness, while Cu_2O parameters show a weaker influence. (c) SHAP analysis representation of input feature values for FF output across DT and RF models. SnO_2 concentration emerges as the most influential parameter for FF. (d) SHAP analysis representation of input feature values for ETA output across DT and RF models. SnO_2 concentration is the strongest influencer of efficiency (ETA), followed by CdTe concentration and CdTe thickness, while Cu_2O parameters show minimal influence.

5. Conclusion

This research develops an integrated approach that employs artificial intelligence to enhance the modelling, optimization, and fabrication guidance of photovoltaic devices. The $Cu_2O/CdTe/SnO_2/ITO$ thin-film solar cell is designed and analyzed using the SCAPS-1D simulator under one-sun AM 1.5G illumination. The effects of varying the thickness (0.04–0.15 μm) and doping concentration (1×10^{15}–1×10^{19} cm^{-3}) of the Cu_2O and SnO_2 layers, along with CdTe layer thickness (1–4 μm) and doping concentration (1×10^{15}–1×10^{19} cm^{-3}), are systematically investigated, generating a dataset of 4097 points. This dataset trains and validates multiple machine learning (ML) algorithms, including Linear Regression, Decision Tree (DT), Gradient Boosting Regressor, Support Vector Regressor, Artificial Neural Network, and Random Forest (RF).

The RF and DT models provide the highest accuracy with an excellent correlation of values (R^2 > 0.98) and lowest root mean square error (RMSE) when compared to SCAPS-1D, while also displaying a robust validation capability. Thus, combining the ML (machine learning) model with SCAPS-1D results in an excellent reduction in both the execution time and the time for optimization calculations. SHAP (Shapley Additive exPlanations) analysis has shown that the SnO_2 donor concentration and the thickness of the CdTe absorber layer are the factors having the strongest influence on the efficiency of the solar cell devices. This indicates that optimizing the SnO_2 doping level and CdTe absorber thickness should be prioritized to achieve higher efficiency, while variations in Cu_2O parameters play a comparatively smaller role within the studied range.The domination of the Cu_2O parameters is slight in comparison. Optimization of these parameters results in an increase in the power conversion efficiency (PCE) as much as 0.59 % to 29.54%. This validates the efficacy of the hybrid framework employed here. The proposed framework not only leads to acceleration of simulation-led design possibilities, but also leads to predictive possibilities which could be useful in the fabrication of experimentally optimized $Cu_2O/CdTe/SnO_2/ITO$ solar cells. In further work, an extension of the combined framework to other possible heterojunction type structures may be completed, while adoption of deep learning as a means of looking at feature interactions might be beneficial, and experiments completed as a method of leading to more direct bridges between simulation and fabrication. In this way, a path would be formed for next-generation devices, which would have a much more efficient photovoltaic effect with a CdTe base.

Acknowledgements: The authors are grateful to Dr Marc Burgelman at the University of Gent, Belgium, for providing the open-source SCAPS-1D simulator.

Availability of Data and Materials: Data will be made available on request.

Funding: This research received no external funding.

Author Contributions: RKM and MNA conceived and designed the experiments; RKM performed the experiments; RKM and MNA analyzed the data; RKM contributed analysis tools; RKM wrote the paper; MNA supervised the study. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Conflict of Interest: The authors declare no conflict of interest.

References

1. Mishra, R. K.; Anwar, M.N.; Hassan, M.A. Performance evaluation of CdTe-based heterojunction solar cell with IGZO-based window layer and electron transport layer. *Chalcogenide Letters* 2025, 22(10), 871–882. https://doi.org/10.15251/CL.2025.2210.871
2. Nijsse, F.J.M.M.; Mercure, J.F.; Ameli, N.; et al. The momentum of the solar energy transition. *Nature Communications* 2023, 14, 6542. https://doi.org/10.1038/s41467-023-41971-7
3. Pourasl, H.H.; Barenji, R.V.; Khojastehnezhad, V.M. Solar energy status in the world: A comprehensive review. *Energy Reports* 2023, 10, 3474–3493. https://doi.org/10.1016/j.egyr.2023.10.022
4. Global Electricity Review 2023. Available online: https://ember-energy.org/latest-insights/global-electricity-review-2023/ (accessed on 10 September 2025).
5. International Energy Agency (IEA), Renewables 2023: Analysis and Forecast to 2028. Available online: https://www.iea.org/reports/renewables-2023 (accessed on 10 October 2025).
6. Sivasankar, S.M.; Amorim, C.D.O.; Cunha, A.F. Progress in thin-film photovoltaics: A review of key strategies to enhance the efficiency of CIGS, CdTe, and CZTSSe solar cells. *Journal of Composites Science* 2025, 9(3), 143. https://doi.org/10.3390/jcs9030143
7. Bonnet, D.; Rabenhorst, H. New results on the development of a thin film p-CdTe–n-CdS heterojunction solar cell. In: Proceedings of the 9th Photovoltaic Specialists Conference; 2–4 May 1972; Silver Spring, MD, USA. pp. 129–131.
8. Lameirinhas, R.A.; Torres, J.P.N.; Cunha, J.P.M. A photovoltaic technology review: History, fundamentals and applications. *Energies* 2022, 15(5), 1823. https://doi.org/10.3390/en15051823
9. Rühle, S. Tabulated values of the Shockley–Queisser limit for single junction solar cells. *Solar Energy* 2016, 130, 139–147. https://doi.org/10.1016/j.solener.2016.02.015
10. Scarpulla, M.A.; McCandless, B.; Phillips, A.B. CdTe-based thin film photovoltaics: Recent advances, current challenges and future prospects. *Solar Energy Materials and Solar Cells* 2023, 255, 112289. https://doi.org/10.1016/j.solmat.2023.112289
11. Emon, E.I.; Islam, A.M.; Sobayel, M.K.; et al. A comprehensive photovoltaic study on tungsten disulfide (WS_2) buffer layer based CdTe solar cell. *Heliyon* 2023, 9(3), e14438. https://doi.org/10.1016/j.heliyon.2023.e14438
12. Bittau, F.; Potamialis, C.; Togay, M.; et al. Analysis and optimisation of the glass/TCO/MZO stack for thin film CdTe solar cells. *Solar Energy Materials and Solar Cells* 2018, 187, 15–22. https://doi.org/10.1016/j.solmat.2018.07.019
13. Li, X.; Xiao, D.; Wu, L.; et al. CdTe thin film solar cells with copper iodide as a back contact buffer layer. *Solar Energy* 2019, 185, 324–332. https://doi.org/10.1016/j.solener.2019.04.082
14. Tinedert, I.E.; Saadoune, A.; Bouchama, I. et al. Numerical modelling and optimization of CdS/CdTe solar cell with incorporation of Cu_2O HT-EBL layer. *Optical Materials* 2020, 106, 109970. https://doi.org/10.1016/j.optmat.2020.109970
15. Hossain, M.M.; Jahan, N.; Hossain, R.U. Simulation and optimization of a highly efficient ZnO/Cu_2O/CdS/CdTe solar cell using SCAPS-1D. In: Proceedings of the 2022 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE); 24–26 February 2022; Gazipur, Bangladesh. pp. 1–5. https://ieeexplore.ieee.org/document/9836410
16. Rani, A.; Dutta, M.; Shekhar, C.; et al. Study on the numerical modeling of CdTe/CdS/SnO_2 based thin film solar cells using SCAPS–1D. *Energy & Environment Advances* 2024, 1(1), 15–24. https://doi.org/10.69626/eea.2024.0015
17. Ndiouga Ndiaye, E.H.; Seck, S.M.; Fall, M.; et al. Optimization of the electrical characteristics of cadmium telluride based thin layer solar cell according to thickness and gap energy. *Research Journal of Applied Sciences, Engineering and Technology* 2020, 4, 129–135. https://doi.org/10.19026/rjaset.17.5821
18. Johora, F.T.; Khan, M.A.; Asadl, A.A.; et al. Effect of defect density variation in CdTe solar cell using Cu as a hole transport layer and back contact. *International Journal of Electrical Engineering and Applied Sciences* 2024, 7(2), 77–86. https://doi.org/10.54554/ijeeas.2024.7.02.009.
19. Singh, P.; Ravindra, N.M. Temperature dependence of solar cell performance–an analysis. *Solar Energy Materials and Solar Cells* 2012, 101, 36–45. https://doi.org/10.1016/j.solmat.2012.02.019
20. Ablekim, T.; Colegrove, E.; Metzger, W.K. Interface engineering for 25% CdTe solar cells. *ACS Applied Energy Materials* 2018, 1(10), 5135–5139. https://doi.org/10.1021/acsaem.8b01173

21. Burgelman, M.; Nollet, P.; Degrave, S. Modelling polycrystalline semiconductor solar cells. *Thin Solid Films* 2000, 361–362, 527–532. https://doi.org/10.1016/S0040-6090(99)00825-1
22. SCAPS-1D Manual. University of Gent, Belgium. Available online: https://scaps.elis.ugent.be (accessed on 21 May 2025).
23. University of Toledo, SCAPS-1D Lecture Notes. 2016. Available online: https://www.utoledo.edu/med/depts/radther/pdf/SCAPS-1D%20Lecture%204%20-%20handout.pdf (accessed on 21 May2025).
24. Park, H.J.; Son, H.; Jeong, B. S. SCAPS-1D simulation for device optimization to improve efficiency in lead-free $CsSnI_3$ perovskite solar cells. *Inorganics* 2024, 12(4), 123. https://doi.org/10.3390/inorganics12040123
25. Scott, C.; Ahsan, M.; Albarbar, A. Machine learning for forecasting a photovoltaic (PV) generation system. *Energy* 2023, 278, 127807. https://doi.org/10.1016/j.energy.2023.127807
26. Bakht, M.P.; Mohd, M.N.H.; Ibrahim, B.S.K.K.; et al. Advanced automated machine learning framework for photovoltaic power output prediction using environmental parameters and SHAP interpretability. *Results in Engineering* 2025, 25, 103838. https://doi.org/10.1016/j.rineng.2024.103838
27. Gomes, E.; Esteves, A.; Morais, H.; et al. Leveraging explainable artificial intelligence in solar photovoltaic mappings: Model explanations and feature selection. *Energies* 2025, 18(5), 1282. https://doi.org/10.3390/en18051282
28. Britt, J.; Ferekides, C. Thin-film CdS/CdTe solar cell with 15.8% efficiency. *Applied Physics Letters* 1993, 62(22), 2851–2852. https://doi.org/10.1063/1.109629
29. Wu, X.; Dhere, R.G.; Yan, Y.; et al. High-efficiency polycrystalline CdTe thin-film solar cells. In: Proceedings of the 17th European Photovoltaic Solar Energy Conference; 22-26 October 2001; Munich, Germany. pp. 995–1000.
30. Green, M.A.; Emery, K.; Hishikawa, Y.; et al. Solar cell efficiency tables (Version 39). *Progress in Photovoltaics: Research and Applications* 2012, 20(1), 12–20. https://doi.org/10.1002/pip.2163
31. Hall, R.S.; Lamb, D.; Irvine, S.J.C. Back contact materials used in thin film CdTe solar cells – a review. *Energy Science & Engineering* 2021, 9(5), 606–632. https://doi.org/10.1002/ese3.843
32. Bosio, A.; Rosa, G.; Romeo, N. Past, present and future of the thin film CdTe/CdS solar cells. *Solar Energy* 2018, 175, 31–43. https://doi.org/10.1016/j.solener.2018.01.018
33. Romeo, A.; Artegiani, E. CdTe-based thin film solar cells: Past, present and future. *Energies* 2021, 14(6), 1684. https://doi.org/10.3390/en14061684
34. McCandless, B.E.; Dobson, K.D. Processing options for CdTe thin film solar cells. *Solar Energy* 2004, 77(6), 839–856. https://doi.org/10.1016/j.solener.2004.04.012
35. Wu, X.; Sheldon, P.; Coutts, T.J.; et al. Application of Cd/sub2/Sno/Seb4/ transparent conducting oxides CdS/CdTe thin-film devices. In: Proceedings of the Conference Record of the Twenty Sixth IEEE Photovoltaic Specialists Conference 1997; 29 September–03 October 1997; Anaheim, CA, USA. pp. 347–350.
36. Hall, R.S.; Lamb, D.; Irvine, S.J.C. Back contact materials used in thin film CdTe solar cells–a review. *Energy Science & Engineering* 2021, 9(5), 606–632. https://doi.org/10.1002/ese3.843
37. Chen, B.; Liu, J.; Cai, Z. et al. The effects of ZnTe:Cu back contact on the performance of CdTe nanocrystal solar cells with inverted structure. *Nanomaterials* 2019, 9(4), 626. https://doi.org/10.3390/nano9040626
38. Lin, H.; Irfan; Xia, W.; et al. MoOx back contact for CdS/CdTe thin film solar cells: Preparation, device characteristics, and stability. *Solar Energy Materials and Solar Cells* 2012, 99, 349–355. https://doi.org/10.1016/j.solmat.2012.01.001
39. Lin, C.; Liu, G.; Xi, X.; et al. The investigation of the influence of a Cu_2O buffer layer on hole transport layers in $MAPbI_3$-based perovskite solar cells. *Materials* 2022, 15(22), 8142. https://doi.org/10.3390/ma15228142
40. Kanevce, A.; Reese, M.O.; Barnes, T.M.; et al. The roles of carrier concentration and interface, bulk, and grain-boundary recombination for 25% efficient CdTe solar cells. *Journal of Applied Physics* 2017, *121(21)*, 214506. https://doi.org/10.1063/1.4984320
41. Türck, J.; Nonnenmacher, H.J.; Connor, P.M.L.; et al. Copper(I) oxide (Cu_2O) based back contact for pin CdTe solar cells. *Progress in Photovoltaics: Research and Applications* 2016, 24(9), 1229–1236. https://doi.org/10.1002/pip.2782
42. Rassol, R.A.; Hasan, R.F.; Ahmed, S.M. Numerical analysis of SnO_2/Zn_2SnO_4/n-CdS/p-CdTe solar cell using the SCAPS-1D simulation software. *Iraqi Journal of Science* 2021, 62(2), 505–516. https://doi.org/10.24996/ijs.2021.62.2.17

43. Medina, J.C.Z.; Andrés, E.R.; Ruíz, C.M.; et al. Numerical simulation and performance optimization of a solar cell based on WO_3/CdTe heterostructure using NiO as HTL layer by SCAPS-1D. *Coatings* 2023, 13(8), 1436. https://doi.org/10.3390/coatings13081436
44. Montoya De Los Santos, I.; Pérez-Orozco, A.A.; Liña-Martínez, D.A.; et al. Towards a CdTe solar cell efficiency promotion: The role of ZnO:Al and CuSCN nanolayers. *Nanomaterials* 2023, 13(8), 1335. https://doi.org/10.3390/nano13081335
45. Samuel, A.L. Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development* 1959, *3(3)*, 210–229. https://doi.org/10.1147/rd.33.0210
46. Mellit, A.; Kalogirou, S.A. Machine learning and deep learning for photovoltaic applications. *Artificial Intelligence for Smart Photovoltaic Technologies* 2022, 1, 1–20. https://doi.org/10.1063/9780735424999_001
47. Benti, N.E.; Chaka, M.D.; Semie, A.G. Forecasting renewable energy generation with machine learning and deep learning: Current advances and future prospects. *Sustainability* 2023, 15(9), 7087. https://doi.org/10.3390/su15097087
48. Bansal, M.; Goyal, A.; Choudhary, A. A comparative analysis of K-Nearest Neighbour, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning. *Decision Analytics Journal* 2022, 3, 100071, https://doi.org/10.1016/j.dajour.2022.100071
49. Cokduygulular, E.; Cetinkaya, C. Machine learning-assisted novel photovoltaic optimization for tailored ultra-thin CdTe-based solar cells. *Solar RRL* 2025, 9(18), 100071. https://doi.org/10.1002/solr.202500455
50. Murugan, P.; Munshi, A. Machine learning-driven process optimization for CdTe PV using large dataset for repeatable PCE over 18%. In: Proceedings of the 2025 IEEE 53rd Photovoltaic Specialists Conference (PVSC 2025); 8–13 June 2025; Montreal, QC, Canada. pp. 1433–1435.
51. Krishnaiah, V.V.J.R; Prakash, V.N.V.S; Chandra, G.R.; et al. Optimizing ZnO/CdS/CdTe bilayer structures for enhanced CdTe solar cell efficiency: A machine learning approach. *MRS Advances* 2024, 9(9), 640–645. https://doi.org/10.1557/s43580-024-00772-w
52. Hui, Z.; Wang, M.; Yin, X.; et al. Machine learning for perovskite solar cell design. *Computational Materials Science* 2023, 226, 112215. https://doi.org/10.1016/j.commatsci.2023.112215