Machine Learning-Assisted Approach Towards Finding the Promising Low-Toxic Perovskite Halides for Solar Cell Applications

Humayun Kabir
Student Number 0416112008

THIS THESIS PAPER IS SUBMITTED TO THE DEPARTMENT OF MATERIALS AND METALLURGICAL ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN MATERIALS AND METALLURGICAL ENGINEERING

DEPARTMENT OF MATERIALS AND METALLURGICAL ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY

May, 2023

COPYRIGHT STATEMENT

The following two notes on copyright and the ownership of intellectual property rights:

1. Copyright in the text of this thesis rests with the author. Copies (by any process), whether in full or in extracts, may be made with the author's consent. Further copies (by any method) of documents made by such instructions may not be made without the author's permission.
2. The ownership of any intellectual property rights described in this thesis is vested in the Department of Materials and Metallurgical Engineering of Bangladesh University of Engineering and Technology, subject to any prior agreement to the contrary. It may only be made available for use by third parties with the written permission of the department, which will prescribe the terms and conditions of any such agreement.

CANDIDATE'S DECLARATION

It is hereby declared that

1. This thesis paper or any part of it has not been submitted anywhere else for the award of any degree.

2. The thesis does not contain any material previously published or written by others, except where this is appropriately cited through complete and accurate referencing.

.....

Humayun Kabir

Student ID : 0416112008

Department of MME

BUET, Dhaka -1000

The thesis titled "MACHINE LEARNING-ASSISTED APPROACH TOWARDS FINDING THE PROMISING LOW-TOXIC PEROVSKITE HALIDES FOR SOLAR CELL APPLICATIONS", submitted by Humayun Kabir, Student Number: 0416112008, Session: April 2016, has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Master of Science in Materials and Metallurgical Engineering on 18th May 2023.

BOARD OF EXAMINERS

(i) _____
Dr. Kazi Mohammad Shorowordi
Professor
Dept. of Materials and Metallurgical Engineering
BUET, Dhaka-1000
Chairman
(Supervisor)

(ii) _____
Dr. A. K. M. Bazlur Rashid
Professor & Head
Dept. of Materials and Metallurgical Engineering
BUET, Dhaka-1000
Member
(Ex-Officio)

(iii) _____
Dr. Ahmed Sharif
Professor
Dept. of Materials and Metallurgical Engineering
BUET, Dhaka-1000
Member

(iv) _____
Dr. Hossain Mohammad Mamun Al Rashed
Associate Professor
Dept. of Materials and Metallurgical Engineering
BUET, Dhaka-1000
Member

(v) _____
Dr. Muhammad Shahriar Bashar
Principal Scientific Officer
Institute of Fuel Research & Development,
Bangladesh Council of Scientific & Industrial Research, Dhaka
Member
(External)

DEDICATION

TO MY BELOVED WIFE

ACKNOWLEDGEMENTS

First of all, thanks to the Almighty ALLAH for blessing me with the strength and opportunity to complete this thesis work successfully.

It was a long journey to finish the research work. I express my profound gratitude to the revered mentor, my supervisor, Prof. Dr. Kazi Mohammad Shorowordi, Dept. of MME, for his invaluable advice, guidance, and motivation. This research work would have been impossible without his continuous support and encouragement. I thank him from the bottom of my heart for his sincere cooperation, advice, and appreciation.

I am grateful to the Materials and Metallurgical Engineering department and especially want to thank Prof. Dr. A.K.M. Bazlur Rashid, Head, Department of MME, for providing all kinds of support whenever needed.

I pay my utmost gratitude to Md. Tohidul Islam, Graduate Research Assistant, Department of Materials Design and Innovation (MDI), University at Buffalo, for his constant support with this thesis project's technical aspects. I also want to give thanks to Md. Shafiqul Islam, Research Assistant, Material Research Center, Materials and Metallurgical Engineering (MME), for his unbelievable help.

Finally, I want to express my respect & love to my parents and my beloved wife for their encouragement and endless support throughout my academic journey.

ABSTRACT

Predicting the bandgap of perovskite absorbers properly is vital for an effective perovskite solar cell design. According to the literature, the bandgap of perovskites is the most dominant parameter affecting solar cell performance. The primary objective of this study was to develop a suitable machine-learning (ML) model to predict the bandgap of any halide-based perovskite absorber. First, three ML algorithms, such as linear regression, random forest regression, and gradient boosting regression, were developed based on the elemental properties of each site present in the perovskites. Different error metrics such as root mean squared error (RMSE), coefficient of determination (R^2), and cross-validation scores were calculated to compare their performances. Among the standalone ML models, the gradient-boosting model achieved the best performance. Then, these base models were used to construct the ensemble voting regression model utilizing weighted averages according to the base models' performances. The voting regressor model outperformed the three baseline models, with the lowest RMSE (0.076) and highest fitting accuracy of R^2 (0.95). To reduce the toxicity of lead in $MAPbI_3$, three potential replacements of Ge, Sn, and Si were considered for bandgap prediction and Sn based $MAPb_{1-x}Sn_xI_3$ was found the best. Finally, a suitable composition of $MAPb_{0.75}Sn_{0.25}I_3$ from the predicted dataset and pristine $MAPbI_3$ was selected for SCAPS-1D analysis. It was found that pristine $MAPbI_3$ performed well in a single junction solar cell with the highest power conversion efficiency of 21.65%. However, replacing the Pb site with 25% Sn reduced the performance showing maximum efficiency (PCE) of 18.58%. The $MAPb_{0.75}Sn_{0.25}I_3$ composition needs further attention to enhance its performance. This study gives an effective tool for the experimentalists to make the task of predicting bandgap efficient, making the whole design process smooth and easier.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AM 1.5 SUN	Air mass of 1.5 atmosphere and solar zenith angle of 48.2 degree
CB	Conduction Band
CV	Cross Validation
J_{SC}	Short Circuit Current Density
IV-curve	Current Voltage Curve
DFT	Density Functional Theory
η	Efficiency
ETL	Electron Transport Layer
FF	Fill Factor
GBRT	Gradient Boost Regression Trees
HTS	High-Throughput Screening
HTL	Hole Transport Layer
HOIP	Hybrid Organic–Inorganic Perovskite
ML	Machine Learning
V_{OC}	Open-Circuit Voltage
PSC	Perovskite Solar Cell
PV	Photovoltaics
PCE	Power Conversion Efficiency
SHAP	Shapley Additive Explanations
SQ Limit	Shockley–Queisser Limit

CHAPTER 1: INTRODUCTION

Solar power, one of the most promising sustainable energy sources, is the key to the clean energy of the future. But, its ability to produce electricity efficiently and cost-effectively remains a great challenge. Crystalline silicon leads the solar photovoltaics industry (with over 95% market share) due to its high power conversion efficiency (PCE), long lifetime of more than 25 years, and excellent capability for industrialization [1]. However, the non-renewable silicon resources and high fabrication costs motivate researchers to look for other promising candidates for solar cell applications [2].

Perovskite solar cells (PSCs), especially lead halide perovskites as absorber materials, have recently attained considerable interest from the scientific community [3], [4]. This growing attention is due to their great potential for improved performance and efficiency, flexibility, simple fabrication process [2], [5]–[11], and relatively low cost of material manufacturing. Since the discovery of Methylammonium Lead Iodide ($MAPbI_3$) perovskite as a solar absorber with an initial efficiency of 3.8%[12], great efforts have been devoted to improve the efficiency of this solar absorber material. Pb-based halide perovskites are now the most efficient perovskite solar cells developed as of yet, and the maximum power conversion efficiency (PCE) of 25.6% is attained with single junction α-$FAPbI_3$ perovskites [13].

Despite the exciting progress and extraordinary potential, the commercialization of this emerging technology still faces severe challenges due to the toxicity resulting from Pb in lead halide perovskites and instability against moisture, heat, and irradiation [14]. Hence, searching for a lead-free alternative to the perovskites solar cell is in great demand[15], [16]. In recent years, significant efforts have been made to develop suitable Pb-free halide perovskite and derivative absorber materials to eliminate the toxicity of lead [17]–[20].

Rapidly discovering functional materials remains a significant challenge because traditional trial-and-error methods are usually inefficient, especially when thousands of candidates are treated. Finding efficient, economical, and environment-friendly alternative materials for solar cells using experimental procedures consumes much time and resources.

The PSC design process currently involves careful compositional tuning [21], systematic fabrication of each active layer [22], [23] , and time-consuming stability and aging test.

The high-dimensional perovskite parameter space and lengthy testing times motivate researchers to use machine learning (ML) to reduce the time to develop, characterize, and optimize devices. The ability of ML algorithms to analyze large amounts of data and predict unknown properties with excellent prediction accuracy is making it popular in the scientific community. ML can identify patterns and relationships that might not be apparent to human experts and finally makes predictions on unknown data by a continuous learning process [24].

To find solar materials with a wide range of bandgaps, computational software such as Vienna Ab Initio Simulation Package (VASP) or Materials Studio can also perform powerful functions using DFT. However, the limitations of these software packages include slow calculation speed, high resource consumption, and high requirements for computer hardware configuration, leading to high calculation costs and low work efficiency [25].

Many factors affect the performance of perovskite solar cells. Appropriate bandgap tuning of the perovskite absorber, excellent electrical conductivity, the high carrier mobility of the transport layer, good energy level alignment, and low defect density at the interfaces are some of the most critical factors in evaluating high-performance perovskite solar cells. Among them, the bandgap design of the perovskite layer is crucial, which directly determines the solar spectrum response range for solar cells [26]–[30].

The present research utilizes supervised ML to create a suitable model for predicting the bandgap of halide perovskite materials (both organic and inorganic) from material compositions and their compositional average elemental properties.

The ML model this way can guide the researchers in synthesizing new perovskite material compositions with experimental bandgaps similar to the predicted bandgaps. This research can potentially revolutionize the development of new materials for solar cell applications by enabling researchers to quickly and accurately identify promising perovskite halides with low toxicity and high performance.

CHAPTER 2: LITERATURE REVIEW

2.1 Renewable Energy

Renewable energy refers to energy obtained from natural resources that can be replenished relatively quickly. These resources include sunlight, wind, rain, tides, geothermal heat, and biomass. Unlike finite and non-renewable fossil fuels, renewable energy sources can be continuously replenished, making them a sustainable solution for energy. Renewable energy is suitable for large-scale projects as well as rural and remote areas and developing countries. From 2011 to 2021, renewable energy has grown from 20% to 28% of global electricity supply, especially, the share of power from sun and wind increased from 2% to 10% [31].

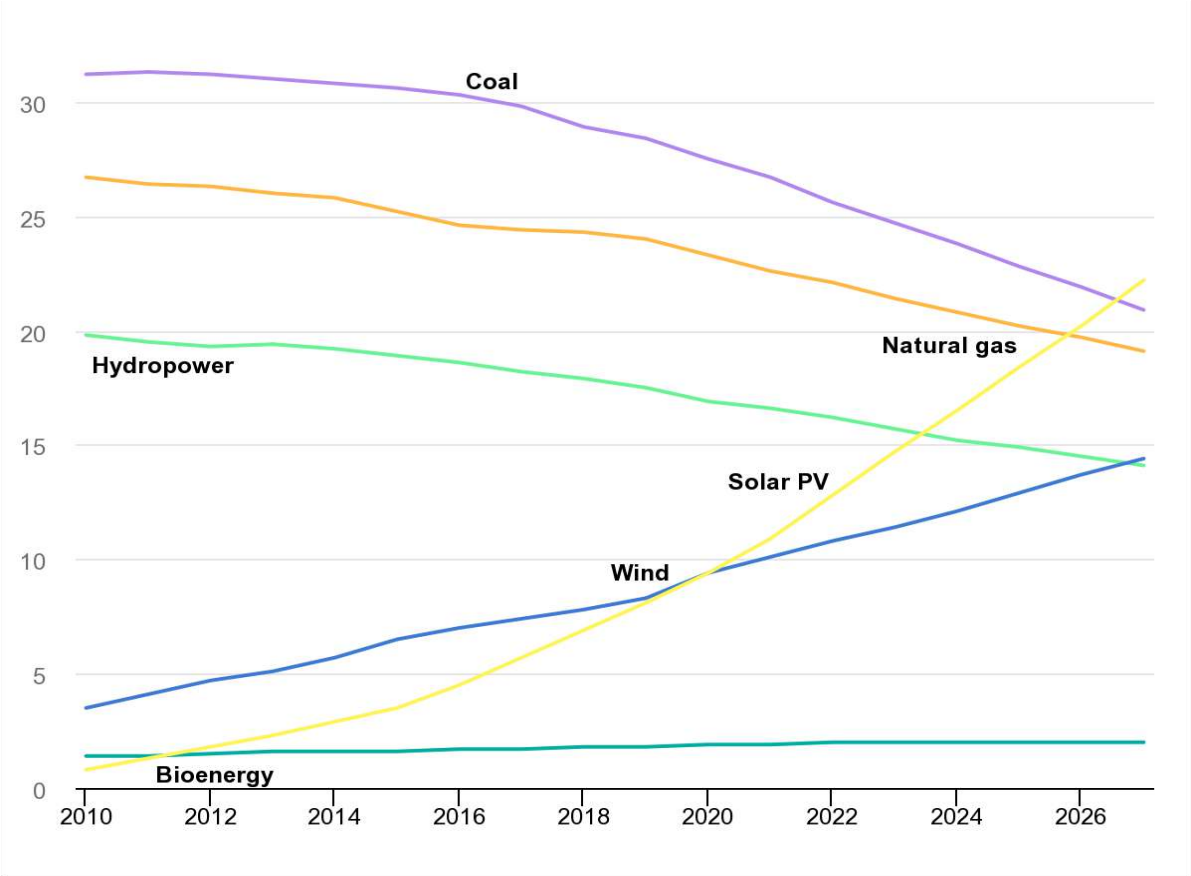


Figure 2.1: Renewable energy sources, especially solar photovoltaic and wind power, are providing an increasing share of power capacity. [32]

2.1.1 Renewable Energy and Climate Change

The world's growing energy need led to the continual use of fossil fuel-based energy sources (Coal, Oil and Gas) which became problematic by creating several challenges. These are: depletion of fossil fuel reserves, greenhouse gas emissions and other environmental concerns, geopolitical and military conflicts, and the continual fuel price fluctuations. These problems will create unsustainable situations which will eventually result in threat to human societies [33]. Hence, a significant climate change has become one of the greatest challenges of the twenty-first century. Its grave impacts may still be avoided if efforts are made to transform current energy systems and if renewable energy sources are used.

Presently, the term "climate change" is of great interest to the world at large. Climate has been changing since the beginning of creation, but what is alarming is the speed of change in recent years and it may be one of the threats facing the earth. The growth rate of carbon dioxide has increased over the past 36 years, averaging about 1.4 ppm per year before 1995 and 2.0 ppm per year thereafter [33].

For more than a decade, the objective of keeping global warming below 2 °C has been a key focus of international climate debate [34]. Since 1850, the global use of fossil fuels has increased to dominate energy supply, leading to a rapid growth in carbon dioxide emissions. Data by the end of 2010 confirmed that consumption of fossil fuels accounted for the majority of global anthropogenic greenhouse gas (GHG) emissions, where concentrations had increased to over 390 ppm (39%) above preindustrial levels [35]. According to the IEA, to achieve net zero emissions by 2050, 90% of global electricity generation will need to be produced from renewable sources

Climate change concerns, coupled with the continuing fall in the costs of some renewable energy equipment, such as wind turbines and solar panels, are driving increased use of renewables. In fact, renewable energy sources may be the most outstanding alternative and the only solution to the growing challenges.

2.1.2 Benefits of Renewable Energy

Fossil fuels are being used more quickly than they are being replenished, whereas renewable energy has vast resources and significant efficiency exist over wide geographical areas. Rapid

deployment of renewable energy and technological diversification of energy sources would result in significant energy security and economic benefits [36].

Some renewable energies, such as solar and wind power have got much cheaper. In some cases, it will be cheaper to transition to these sources as opposed to continuing to use fossil fuels. Moreover, electrification with renewable energy is more efficient and has less energy requirements. It would also reduce environmental pollution such as air pollution caused by the burning of fossil fuels, and improve public health, reduce premature mortalities and save associated health costs [37]. Multiple analyses have found that these benefits can significantly overcome the costs of implementing these strategies [38], [39].

Some studies have shown that a global transition to 100% renewable energy across all sectors– power, heat, transport and industry – is feasible and economically viable. However, the global electricity generation has been shifting from fossil fuels to renewable energy gradually. It is anticipated that the growth of renewable energy (especially solar energy) will significantly reduce the usage of fossil fuels in electricity generation by 2025 (Figure 2.2).

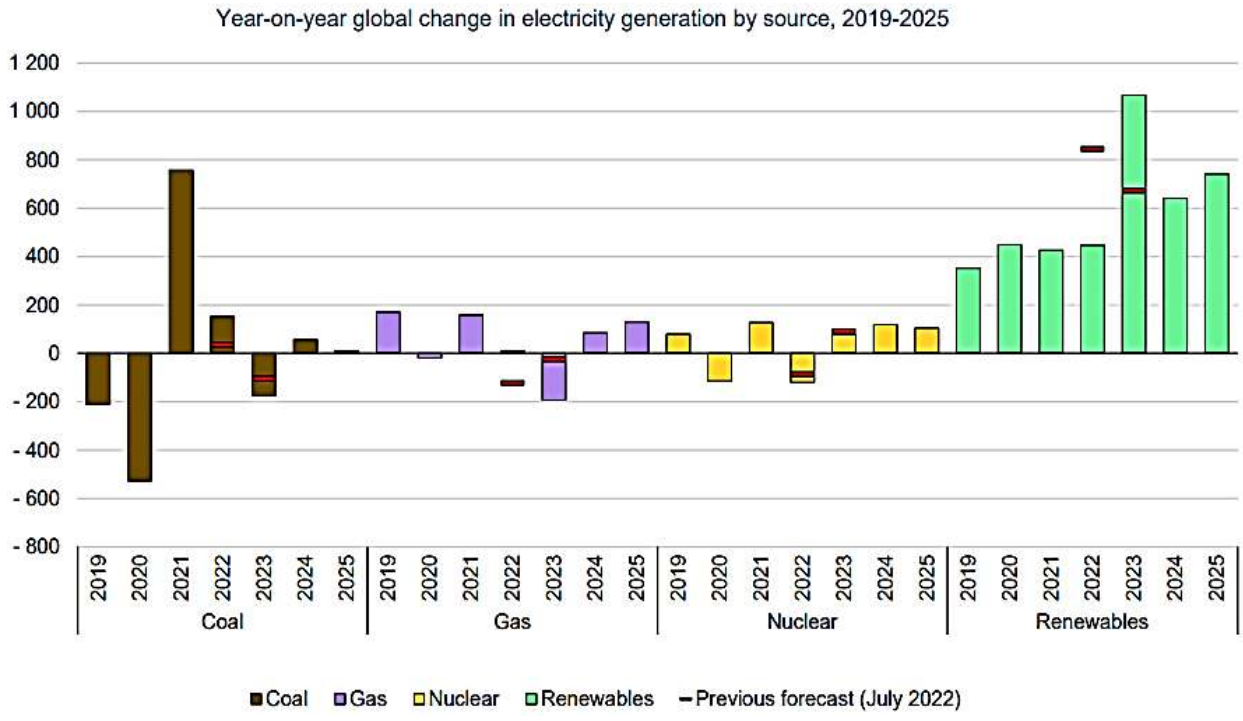


Figure 2.2: Renewable energy growth reducing fossil fuel electricity generation. [39]

In spite of the outstanding advantages of renewable energy sources, certain shortcoming exists such as: the discontinuity of generation due to seasonal variations as most renewable energy resources are climate-dependent. That is why its employment requires complex design, planning and control optimizations. Fortunately, the continuous technological advances in computer hardware and software are allowing researchers to handle these difficulties using computational resources applicable to the renewable energy field [39].

2.2 Solar Energy

Solar energy is the source of nearly all energy on the earth. Humans, like all other animals and plants, rely on the sun for warmth and food. However, people also harness the sun's energy in many other different ways. For example, fossil fuels, plant matter from a past geological age, is used for transportation and electricity generation and is essentially just stored solar energy from millions of years ago. Similarly, biomass converts the sun's energy into a fuel, which can then be used for heat, transport or electricity. Wind energy, used for hundreds of years to provide mechanical energy or for transportation, uses air currents that are created by solar heated air and the rotation of the earth. Today wind turbines convert wind power into electricity as well as its traditional uses. Photovoltaics (often abbreviated as PV) is a simple and elegant method of harnessing the sun's energy. PV devices (solar cells) are unique in that they directly convert the incident solar radiation into electricity, with no noise, pollution or moving parts, making them robust, reliable and long lasting.

Solar energy is harnessed using a range of such as solar heating, photovoltaics (PV), concentrator photovoltaics (CPV), concentrated solar power (CSP), solar architecture and artificial photosynthesis. At present, most of the renewable energy is solar energy [37].

Solar technologies are classified as either passive solar or active solar depending on how they convert, and distribute solar energy. Passive solar techniques include orienting a building to the Sun, selecting materials with favorable thermal mass or light dispersing properties, and designing spaces that naturally circulate air. Active solar technologies include solar thermal energy, using solar collectors for heating, and solar power, converting sunlight into electricity either directly using PV, or indirectly using CSP.

Solar power produces 505 GW electricity annually, which is about 2% of the world's production [37]. The amount of solar energy that can be harnessed for electricity generation is influenced by weather conditions, geographic location (position of the sun) and time of the

day. The global potential of direct solar energy is greater than any other renewable energy resource. It is far beyond the total amount of energy needed in order to support the current century [35]

However, there are, environmental implications of scaling up the usage of solar energy. Especially, the demand for raw materials such as aluminum poses concerns over the carbon footprint that will result from utilizing raw materials needed to implement solar energy [40].

2.3 Properties of Sunlight

Sunlight is a form of electromagnetic radiation and the visible light that we see is a small subset of the electromagnetic spectrum as shown in figure 2.3. Quantum-mechanics explains both the observations of the wave nature and the particle nature of light. In quantum mechanics, a photon, like all other quantum-mechanical particles such as electrons, protons, etc., is most accurately pictured as a wave-packet. A wave packet is defined as a collection of waves that may interact in such a way that the wave-packet may either appear spatially localized, or may alternately appear simply as a wave. In the cases where the wave-packet is spatially localized, it acts as a particle. Therefore, depending on the situation, a photon may appear as either a wave or as a particle, and this concept is called wave-particle duality. A complete physical description of the properties of light requires a quantum-mechanical analysis of light since light is a type of quantum-mechanical particle [41].

The efficiency of a solar cell is sensitive to variations in both the power and the spectrum of the incident light. To facilitate an accurate comparison between solar cells measured at different times and locations, a standard spectrum and power density have been defined for both radiations outside the Earth's atmosphere and at the Earth's surface. The standard spectrum at the Earth's surface is called AM 1.5G (AM for Air Mass and the G stands for global and includes both direct and diffuse radiation) or AM 1.5D (which includes direct radiation only). The intensity of AM1.5D radiation can be approximated by reducing the AM0 spectrum by 28% (18% due to absorption and 10% to scattering). The global spectrum is 10% higher than the direct spectrum. These calculations give approximately 970 W/m2 for AM1.5G. However, the standard AM1.5G spectrum has been normalized to give 1kW/m^2 due to the convenience of the round number and the fact that there are inherent variations in incident solar radiation [42].

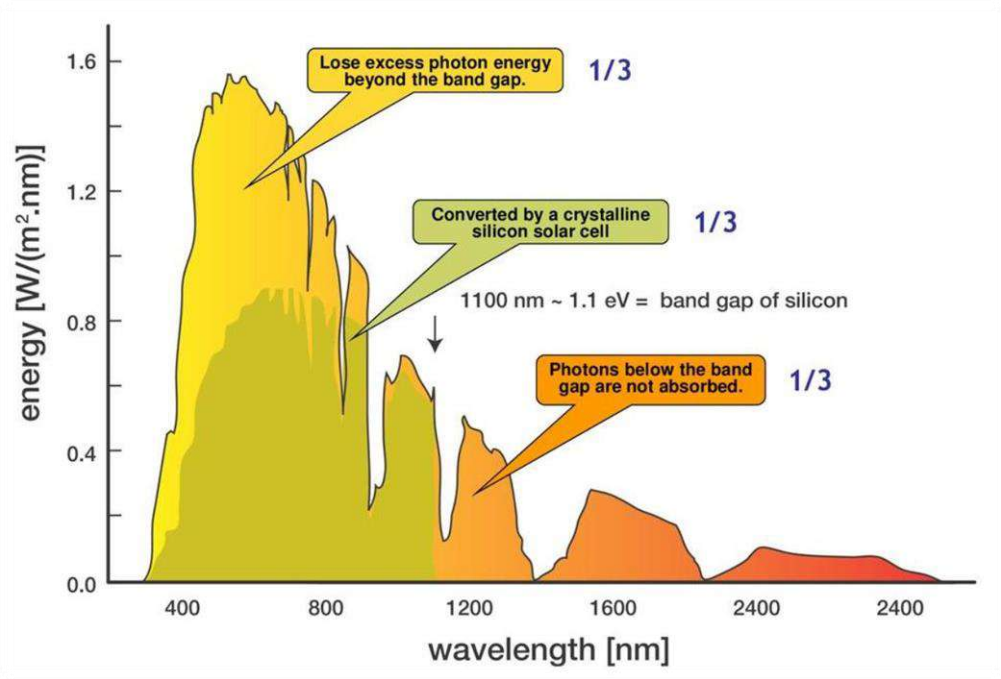


Figure 2.3: Conversion and absorption losses for a crystalline silicon solar cell with a bandgap of 1.1 eV and Solar spectrum of AM1.5G (1000 W/m2). [43]

2.4 Basic Properties of Semiconductor

In most cases, semiconductor is used for solar cell material. The energy conversion consists of the absorption of light (photon) energy producing electron–hole pairs in a semiconductor and charge carrier separation. A variety of materials and processes can potentially satisfy the requirements for photovoltaic energy conversion4, but in practice, nearly all photovoltaic energy conversion uses semiconductor materials in the form of a *p*-*n* junction.

2.4.1 Energy Band and Carrier Concentration

It is very important to learn the basic properties of semiconductor and the principle of conventional *p*–*n* junction solar cell to understand not only the conventional solar cell but also the new type of solar cell.

The electrons of an isolated atom have discrete energy levels. When atoms approach to form crystals, the energy levels split into separate but closely spaced levels because of atomic interaction, which results in a continuous energy band. Between the two bands – the lower one called the *valence band* and the upper one called the *conduction band and* there is an energy gap between these two bands called *band gap*, E_g, which is an important parameter in solar cell. All the energy levels in the valence band are occupied by electrons and those in the

conduction band are empty at a temperature of 0 K. The representation of energy band for semiconductor is shown in (Figure 2.4)**.** E_c and E_v are designated as the bottom of the conduction band and the top of the valence band, respectively.

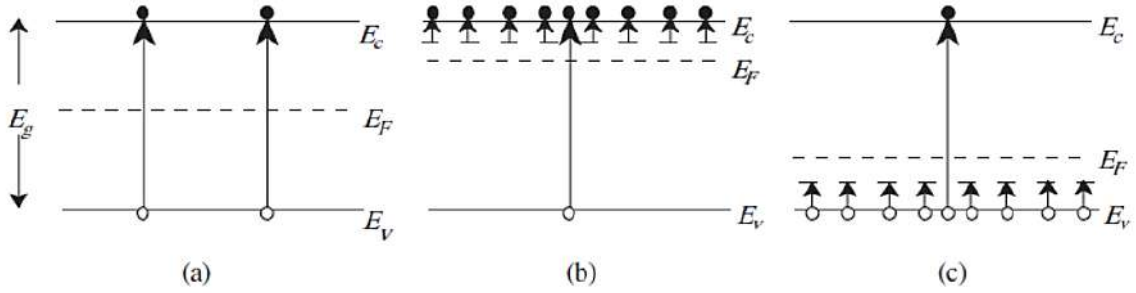


Figure 2.4: Energy band representations of (a) Intrinsic semiconductor (b) n-type semiconductor with donors (c) p-type semiconductor with acceptors [44]

The electrons in conduction band and holes in valence band can contribute to the current flow. On the contrary, in an insulator, the band gap is so large (E_g>5eV) that the conduction band is empty even at room temperature. In a conductor, the conduction band is partially filled with electrons or overlaps the valence band. Consequently, there is no band gap and the resistivity is very small.

When electrons and holes generated from impurities are much smaller than thermally generated electrons and holes, they are called intrinsic semiconductors. Doping impurities in a pure semiconductor crystal, results concentration of electron or hole carriers, which is called extrinsic semiconductors.

When group III atoms such as boron (B) are doped as an impurity into silicon (Si), the boron atom forms covalent bonds with its four neighboring Si atoms. The positively charged conduction hole is created. This is a *p*-type semiconductor where holes are majority carriers. When dopants generate holes and result p-type semiconductors it's called acceptors. Here, the boron atom is called an acceptor

When group V atoms or pentavalent impurities, such as phosphorus (P), are doped as an impurity, the phosphorous atom forms covalent bonds with its four neighboring Si atoms. The fifth electron is bound with the P atom very loosely and therefore ionized even at room temperature. Consequently, it becomes a conduction electron with a negative charge. In this case, Si becomes an *n*-type semiconductor, where electrons are majority carriers. When

dopants generate electrons and results n-type semiconductors it’s called donors. Here, the phosphorous atom is called a donor.

The Fermi level can be controlled by the dopant concentration. When silicon crystal structure is doped, its Fermi level is changed according to type of the dopant. When a crystal structure is not doped Fermi level is approximately between valance and conduction energy levels. In the case of *p*-type semiconductor the Fermi level moves closer to the top of the valence band and for *n*-type semiconductor the Fermi level is near conduction band (Figure 2.3). The fermi level plays a very important role in modulating the overall optical and electrical performance of the material.

The most important parameters of a semiconductor material for solar cell operation are:

- the band gap (E_g) or the minimum amount of energy required for an electron to break free of its bound state;
- the number of free carriers (electrons or holes) available for conduction; and
- the "generation" and recombination of free carriers in response to light incident on the material.

Among the parameters, the solar cell performance is affected dominantly by the bandgap of semiconductors. Islam *et al.* investigated the simulated parameters of the SCAPS-1D software on the performance of Cs-perovskite-based solar devices through a machine learning approach and demonstrated that the energy bandgap, hole mobility, and electron affinity of the absorber perovskites were the most impactful parameters [30].

2.4.2 Built-In Potential in p-n Junction

The important role of p–n junction is the charge separation of light-induced electrons and holes. Figure 2.5 shows the energy band picture and majority carriers for n-type semiconductor and p-type semiconductor. When a p–n junction is formed, the large carrier concentration gradients cause the diffusion of carriers, that is, holes diffuse from p-type semiconductor to n-type semiconductor and electrons diffuse from n-type semiconductor to p-type semiconductor.

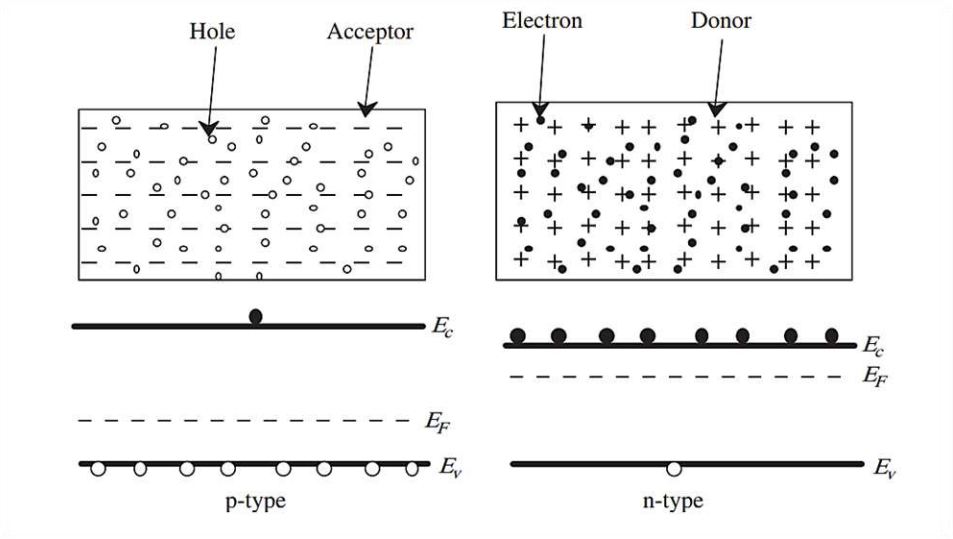


Figure 2.5: Energy band pictures and majority carriers of n- and p-type semiconductors. [44]

Because of the ionized impurity atoms, a layer without mobile charge carriers is formed when the electrons and holes diffuse across the junction. This space charge sets up an electric field, which opposes the diffusion across the junction as shown in Figure 2.6. When the drift current due to the electric field is balanced by the diffusion current because of the carrier concentration gradient for each carrier, the thermal equilibrium is established.

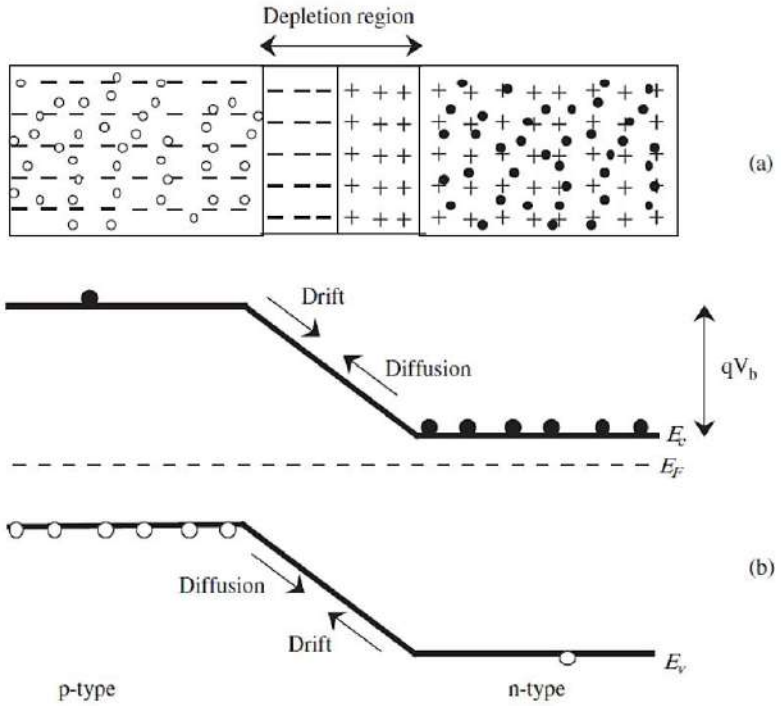


Figure 2.6: (a) Schematic structures of p–n junction and (b) its energy band diagram in thermal equilibrium. [44]

2.5 Solar Cell

A solar cell, also known as a photovoltaic cell, works by converting light energy from the sun into electrical energy. When light particles (photons) strike the surface of the cell, they release electrons from the material they are made of (usually silicon). These electrons are then captured by the cell's electrical field and sent through a circuit, producing a flow of electricity. The flow of electrons generates a direct current (DC) electrical output, which can be converted into alternating current (AC) using an inverter for use in homes and businesses.

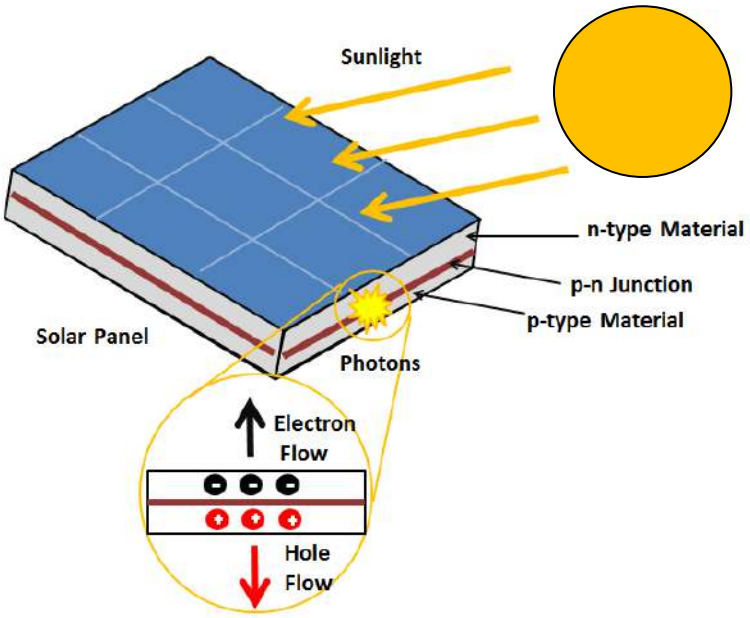


Figure 2.7: How solar cell works [45]

2.5.1 The Photovoltaic Effect

Voltage is generated in a solar cell by a process known as the "photovoltaic effect". The collection of light-generated carriers by the p-n junction causes a movement of electrons to the n-type side and holes to the p-type side of the junction.

When the photons with suitable energy (higher than the bandgap value of PV materials) are incident onto the PV materials, energy from the light can be absorbed and transferred to electrons, which leads to the photoexcitation of electrons. Excitation of electrons causes the generation of excitons, which subsequently undergoes charge separation and forms charge carriers - electrons and holes. The electrons jump onto higher energy levels, namely the conduction band (CB), while leaving an empty vacant hole in the valence band (VB).

Under short circuit conditions, there is no buildup of charge, as the carriers exit the device as light-generated current. However, if the light-generated carriers are prevented from leaving the solar cell, then the collection of light-generated carriers causes an increase in the number of electrons on the n-type side of the p-n junction and a similar increase in holes in the p-type material. This separation of charge creates an electric field at the junction which is in opposition to that already existing at the junction, thereby reducing the net electric field. Since the electric field represents a barrier to the flow of the forward bias diffusion current, the reduction of the electric field increases the diffusion current. A new equilibrium is reached in which a voltage exists across the p-n junction. The current from the solar cell is the difference between I_L (light generated current) and the forward bias current. Under open circuit conditions, the forward bias of the junction increases to a point where the light-generated current is exactly balanced by the forward bias diffusion current, and the net current is zero.

2.5.2 Solar Cell Parameters

2.5.2.1 IV Curve

In the field of solar cell technology, various parameters are used to measure the performance of solar cells. One important parameter is the IV curve, which is the combination of the IV curve of the solar cell diode in the dark with the light-generated current. When illuminated, the light shifts the IV curve down into the fourth quadrant where power can be extracted from the diode. Illuminating a cell adds to the normal dark currents in the diode so that the diode law becomes:

$$I = I_0 \left[\exp\left(\frac{qV}{nkT} \right) - 1 \right] - I_L \tag{2.1}$$

where IL = light generated current.

The equation for the IV curve in the first quadrant is:

$$I = I_L - I_0 \left[\exp\left(\frac{qV}{nkT} \right) - 1 \right] \tag{2.2}$$

The -1 term in the above equation can usually be neglected. The exponential term is usually >> 1 except for voltages below 100 mV. Further, at low voltages, the light generated current I_L dominates the I_0 (...) term so the -1 term is not needed under illumination.

$$I = I_L - I_0 \left[\exp\left(\frac{qV}{nkT} \right) \right] \tag{2.3}$$

Plotting the above equation gives the IV curve below with the relevant points on the curve labeled.

The power curve has a maximum denoted as P_{MP} where the solar cell should be operated to give the maximum power output. It is also denoted as P_{MAX} or maximum power point (MPP) and occurs at a voltage of V_{MP} and a current of I_{MP}.

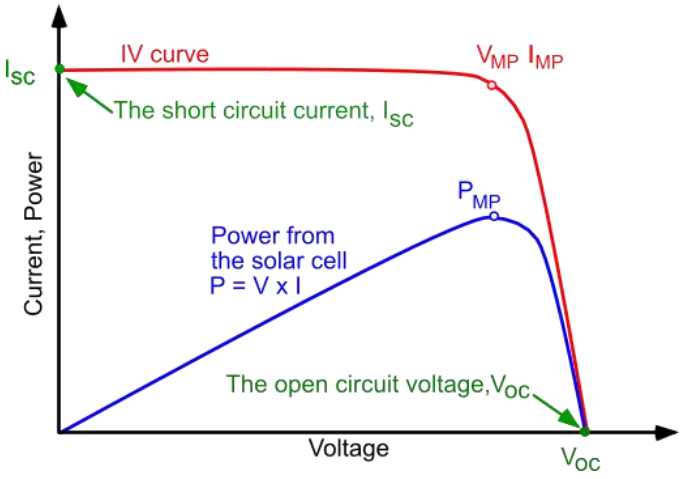


Figure 2.8: Current voltage (IV) curve of a solar cell. [46]

2.5.2.2 Short-circuit Current (I_{SC})

Another important parameter is the short-circuit current (I_{SC}), which is the largest current that can be drawn from the solar cell when the voltage across the solar cell is zero (i.e., when the solar cell is short circuited).

The short-circuit current is due to the generation and collection of light-generated carriers. For an ideal solar cell at most moderate resistive loss mechanisms, the short-circuit current and the light-generated current are identical.

When comparing solar cells of the same material type, the most critical material parameter is the diffusion length and surface passivation. In a cell with perfectly passivated surface and uniform generation, the equation for the short-circuit current density can be approximated as:

$$J_{SC} = qG\,(L_n + L_p) \tag{2.4}$$

where G is the generation rate, and L_n and L_p are the electron and hole diffusion lengths respectively. Although this equation makes several assumptions which are not true for the conditions encountered in most solar cells, the above equation indicates that the short-circuit current depends strongly on the generation rate and the diffusion length.

The short circuit current, I_{SC}, is the short circuit current density, J_{SC}, times the cell area:

$$I_{SC} = J_{SC}A \tag{2.5}$$

2.5.2.3 Open-circuit Voltage (V_{OC})

The open-circuit voltage, V_{OC}, is the maximum voltage available from a solar cell, and this occurs at zero current. The open-circuit voltage corresponds to the amount of forward bias on the solar cell due to the bias of the solar cell junction with the light-generated current. The open-circuit voltage is shown on the IV curve below.

An equation for V_{OC} is found by setting the net current equal to zero in the solar cell equation to give:

$$V_{OC} = \frac{nkT}{q} \ln\left(\frac{I_L}{I_0} + 1\right) \tag{2.6}$$

Silicon solar cells on high quality single crystalline material have open-circuit voltages of up to 764 mV under one sun and AM1.5 conditions, while commercial silicon devices typically have open-circuit voltages around 690 mV [47].

2.5.2.4 Fill Factor (FF)

The fill factor (FF) is a parameter which, in conjunction with V_{oc} and I_{sc}, determines the maximum power from a solar cell. The FF is defined as the ratio of the maximum power from the solar cell to the product of V_{oc} and I_{sc} so that:

$$FF = \frac{P_{MP}}{V_{OC} \times I_{SC}} \tag{2.7}$$

Aspects that play an important role in determining the fill factor includes the recombination rate, the series resistance of the solar cell device, the shunt resistance, all of which are also linked to the grain boundary of the crystalline solar absorber material.

2.5.2.5 Solar Cell Efficiency (η)

The efficiency (η) or power conversion efficiency (PCE) is the most commonly used parameter to compare the performance of one solar cell to another. Efficiency is defined as the ratio of energy output from the solar cell to input energy from the sun. The efficiency depends on the spectrum and intensity of the incident sunlight and the temperature of the solar cell. The conditions under which efficiency is measured must be carefully controlled in order to compare the performance of one device to another.

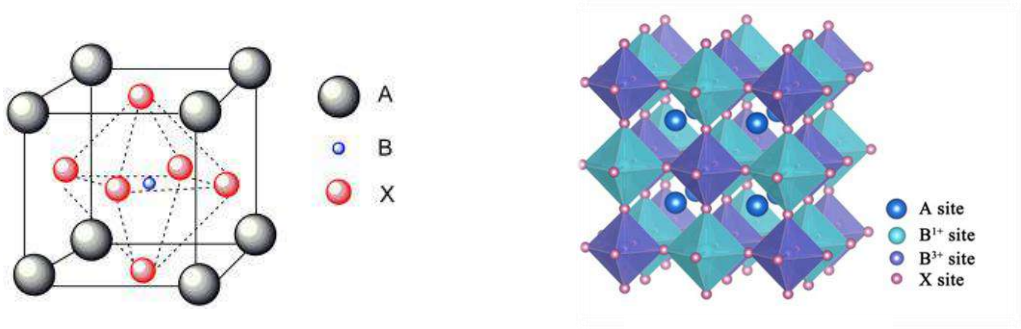
The efficiency of a solar cell is determined as the fraction of incident power which is converted to electricity and is defined as:

$$P_{\max} = V_{OC} I_{SC} FF \tag{2.8}$$

$$\eta = \frac{V_{OC} I_{SC} FF}{P_{in}} \tag{2.9}$$

2.6 Perovskite Solar Material

Perovskite, named after Russian geologist Perovski, is the mineral calcium titanate ($CaTiO_3$), and nowadays any material with the same crystal structure as $CaTiO_3$ belongs to the perovskite family. ABX_3 perovskite is called simple perovskite, while $AA'BB'X_6$ perovskite is known as double perovskite (DP), where A and B are two cations and X is an anion. According to whether A-site cations are organic small molecules or metal ions, the simple perovskite could be classified into inorganic perovskite and hybrid organic-inorganic perovskite (HOIP). Ideally, the larger A anion is in a 12-fold cube-octahedral coordination sphere with twelve X anions, while the smaller B cation occupies the octahedral site surrounded in an octahedron manner by six anions [4].



(a) **(b)**

Figure 2.9: Crystal Structures of (a) simple perovskite (b) double perovskite.[48][49]

Compared to the original perovskite, the double perovskite structure incorporates two different types of cations, B^{1+} and B^{3+}, instead of a single B. With anion X, both B^{1+} and B^{3+} cations form octahedral units.

Perovskite solar cell (PSC) has gained a lot of attention in recent years due to their high conversion efficiency and low cost. These solar cells are made from different hybrid perovskites. The exploration of different chemical compositions and structures for hybrid perovskites is ongoing, and researchers are working to improve the efficiency and stability of perovskite solar cells. Despite their promising potential, perovskite solar cells still face challenges in terms of stability and durability, and further research is needed to address these issues [50].

Their crystallographic stability and probable structure can be deduced by considering a tolerance factor t and an octahedral factor μ [51]. Here, t is defined as the ratio of the distance A-X to the distance B-X in an idealized solid-sphere model ($t = (R\text{A} + R\text{X})/ \{\sqrt{2}(R\text{B} + R\text{X})\}$, where RA, RB and RX are the ionic radii of the corresponding ions) and μ is defined as the ratio RB/RX. For halide perovskites (X = F, Cl, Br, I), generally $0.81 < t < 1.11$ and $0.44 < \mu < 0.90$ [51]. If t lies in the narrower range 0.89–1.0, the cubic structure of Fig. 1a is likely, with lower t values giving less symmetric tetragonal or orthorhombic structures. Despite these constraints, transitions between such structures on heating are common for any given perovskite, with the high-temperature phase generally being cubic. However, it was found that both tolerance factor and octahedral factor are necessary but not sufficient conditions for ABX_3 halide perovskite formability; using these two factors (t and μ) to build up a two-dimensional structure map, this formability can be reliably predicted.

The most-studied perovskite material is methylammonium lead iodide perovskite ($CH_3NH_3PbI_3$) which is used as a light absorber, whereby Pb (II) occupies the octahedral site

while the methylammonium cation is on the cube-octahedral site. The symmetry of a particular perovskite is dependent on the relative ion sizes, and the three-dimensional structure results from the stoichiometry of organic cations and metal halides. It is observed that with increased dimensionality from 2D (A_2BX_4) to 3D (ABX_3), a narrower bandgap is observed, which is beneficial for solar cell applications [52].

2.7 Limitations of Lead-based halide perovskite materials

The properties of Lead-based perovskites are perfect for solar cells. The efficiency of perovskite solar cells has reached 23.7%, which is greater than other solar cells [53]. However, lead-based perovskite solar cells have two main concerns: poor stability and high toxicity. The stability challenge in these perovskites was addressed through the use of 2D perovskites and improved engineering and encapsulation, but the toxicity could not. Hence, using lead is supposed to create hazards in the surrounding environment, causing air, soil and groundwater contamination, ultimately harming the humans (Figure 2.10).

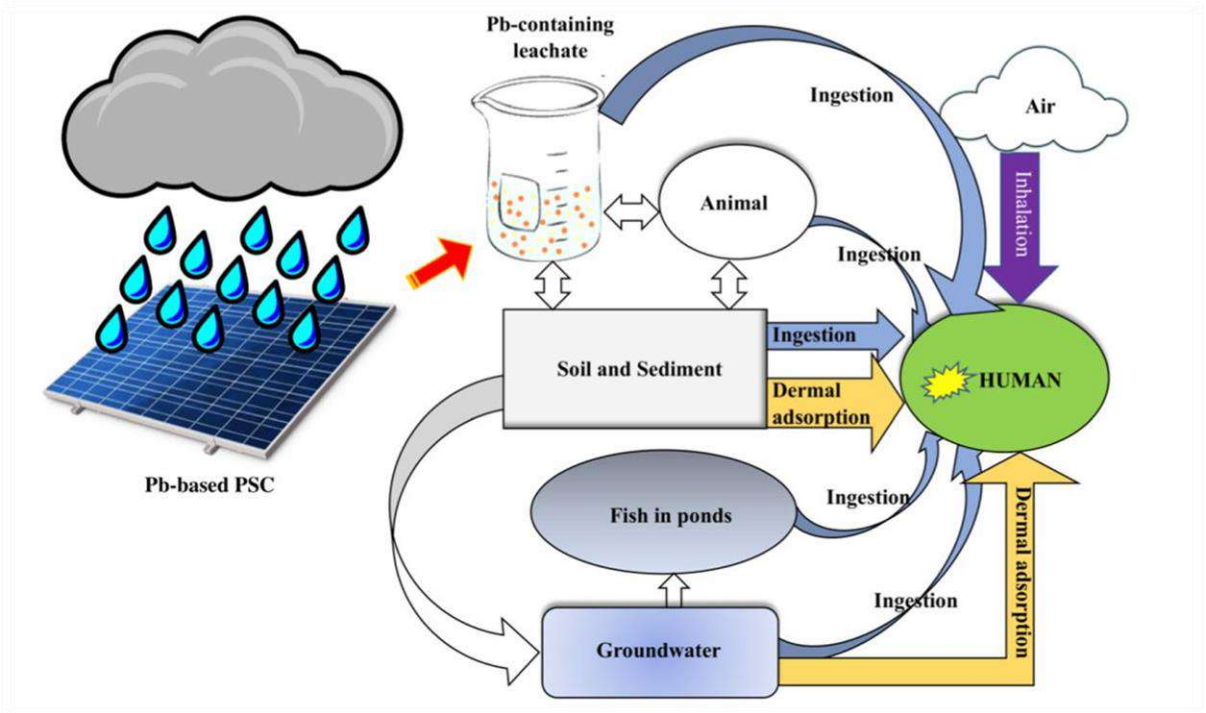


Figure 2.10: Toxicity Caused by Lead used in PSC. [30]

The development of low-toxicity lead-free materials is preferred if its performance is not too compromised. Ideal lead-free candidates as solar cell absorbers should have low toxicity,

narrow direct bandgaps, high optical absorption coefficients, high mobilities, low exciton-binding energies, long charge-carrier lifetimes, and good stability [54].

2.8 Lead-free Perovskite Solar Cells

The toxicity of lead associated with the lifecycle of these PSCs is a serious concern, and it may prove to be a major challenge in the path toward their commercialization [11], [55]–[57]. As current publications have address the toxicity issue of lead [58]–[62], increasing number of researchers have attempt to find suitable replacement of lead in PSC [63], [64]. Thus, significant effort is being devoted toward the development of low-cost, efficient lead-free PSCs. Several low-toxicity cations have been proposed for replacing Pb(II) in halide perovskites, including Ag(I) [65], Bi(III) [66], [67], Sb(III) [68], Ti(IV) [69], [70], Ge(II) [71], and Sn(II) [72], [73] . Among these candidates, halide perovskites based on Sn(II) have shown the highest PCE, and, thus, have attracted the most attention in the PSC field.

Potential materials as solar cell absorbers containing these low-toxicity constituents such as Sn/Ge-based halides, some double perovskites, and some Bi/Sb-based halides with perovskite-like structure are shown graphically in the Figure 2.11 with their bandgaps.

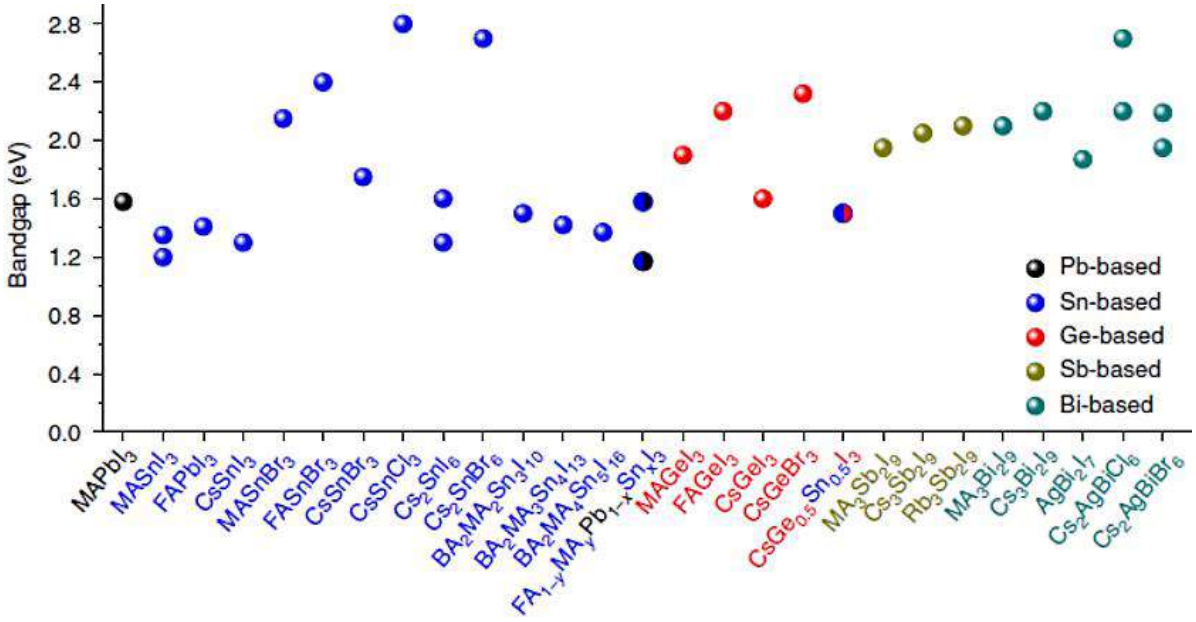


Figure 2.11: Bandgaps of various materials, the suitable materials for solar cells have direct bandgaps of around 1.1 to 2.0 eV. [19], [74]–[78]

2.8.1 Tin-based Perovskites

Tin is the most suitable candidate for substitution of lead for lead-free perovskite solar cell because of its similar chemical nature and valence electronic configuration as that of lead and approximate same ionic radius of Sn^{2+} (115 pm) as that of Pb^{2+} (119 pm). It has lower value of electronegativity Sn^{2+} (1.96) than that of Pb^{2+} (2.33) [79]. Tin-based perovskites have optical bandgap of 1.2–1.6 eV most suitable for their use as light absorbers, large carrier mobilities, and low exciton binding energies of 18 meV [80]–[82]. Tin-based perovskites are represented by the general formula $ASnX_3$ where A can be MA+, Fa+ or Cs+ cation, and X is a halogen anion.

The stability issues of Sn with self-doping had limits the development of Sn-based perovskite [83]. This issue can potentially be managed by adding SnF_2 as inhibitor towards the formation of Sn^{4+} [83]. Additionally, encapsulation also improves the overall device stability which had been proven to be able to maintain 98% of its power conversion efficiency over 100 days. Thus, there are huge potential for the development of other materials as replacement for the highly toxic lead to produce lead-free perovskite solar cells with high-efficiency.

Substituting lead with tin forming perovskite as $CH_3NH_3SnX_3$, led to PCE over 6% [84]. Inorganic and mixed halide tin-based perovskite solar cells were studied with an efficiency of around 6%, made by $CsSnI_3$ and $CH_3NH_3SnI_3–xBr_3$ perovskites, respectively [85], [86]. However, when moving up group 14 elements in the periodic table, the stability of the 2+ oxidation state is much poorer. This disappointing point was shown in the works mentioned above, as none of these Sn-based perovskite solar cells could survive under an ambient environment. Therefore, the poor stability of tin perovskites will largely hinder their future development.

A recent study on developing 2D/3D formamidinium tin iodide perovskite solar cells has shown enhanced PCE up to 9% with good reproducibility [86]. However, decomposition of Sn-perovskite happens rapidly without encapsulation, with only 59% of PCE remaining after only 3 days in the air. This raises the need to explore other substitutes of lead for perovskite solar cells with better stability and less toxicity.

A recent study by Ogomi *et al.* reported a mixed metal, Sn–Pb perovskite which allowed tunability of the band gap of the perovskite absorber by varying the Sn to Pb ratio, indicating

that Sn could be a good choice of metal ion, especially for lower band gap solar cells [87]. However, the same study reported that the neat $CH_3NH_3SnI_3$ perovskite does not exhibit significant photovoltaic properties and that a minimum content of Pb is needed to stabilize Sn in its 2+ oxidation state.

A Sn-based perovskite model with the novel architecture of glass/ZnO:Al/TiO2/ $CH_3NH_3SnCl_3$/CuI/Au has been analyzed by using the solar cell capacitance simulator (SCAPS-ID), with the predicted parameters such as thickness 0–6 μm, defect density of 10^{14} cm^{-3} of light absorber layer, and bandgap 1.3 eV [88], [89]. The model achieved a PCE of 24.82%, V_{OC} of 1.04 V, J^{SC} of 3.50 mA cm-2, and FF of 0.78. The excellent results of this model clearly signify the enormous potential of Sn-based perovskites for their efficient use in solar cells. Since then, extensive work has been carried out on the preparation and characterization of Sn-based perovskite material to examine their structural, optical, and charge transport abilities for efficient use as light absorbers in perovskite solar cells [89].

2.8.2 Germanium-based Halide Perovskite

There Germanium-based perovskites also have similar properties with the lead-based perovskites though not as similar as the Tin analogs [90]. In the $AGeI_3$ perovskite family, $CsGeI_3$ has the narrowest bandgap of around 1.6 eV and is suitable as solar cell absorber. However, the current results in solar cell efficiency show that the Germanium-based perovskites are even more inferior than the Tin-based perovskites.

Making stable and high-performance solar cells using the Germanium-based perovskites has been challenging. More promising are Germanium and Tin alloys, which have narrower bandgaps and better stability than the neat Germanium-based perovskites. All-inorganic high-quality $CsSn_{0.5}Ge_{0.5}I_3$ perovskite films with a bandgap of 1.5 eV can be fabricated by thermal evaporation, and solar cells gave a remarkable PCE of 7.11% [91]. Even when exposed to ambient air, the solar cells can still retain 91% of its initial efficiency.

2.8.3 Antimony-based Perovskite

Antimony is located in the next group neighboring to lead and tin and is widely used in therapeutics acting as a potential non-toxic alternative to lead. A strong lone pair effect can be observed for antimony especially in hybrid complexes, which commonly adopts a (+3)

oxidation state. Due to the difference in oxidation state, a perovskite structure is precluded from the direct formation.

Studies on inorganic Sb_2S_3 and SbSI have been carried out to explore their photovoltaic properties [92], [93]. By using Sb_2S_3 as a photosensitizer in a mesoscopic solar cell structure, nearly 5.2% PCE has been achieved [94]. Sb_2S_3 has achieved 4.17% PCE by adopting a hole transport material-free planar solar cell structure with post surface selenization doping method [95]. Direct substitution of lead in $[CH_3NH_3][PbI_3]$ with antimony gives $Cs_3Sb_2I_9$ with an indirect bandgap of 2.05 eV, but only 0.5% PCE has been achieved in a solar cell device. By mixing sulfur and iodide in the anionic part of $MA_3Sb_2I_9$, a perovskite-like structure has been achieved with a chemical formula of MASbSI. This material has created a new record for hybrid antimony-based solar cells with an efficiency of 3.08%. A bandgap value of 2.03 eV has been experimentally measured, with superior stability manifested in the device study [96].

2.8.4 Bismuth-based Perovskite

As the least toxic heavy metal with desirable chemical properties, bismuth is playing an important role, especially in non-toxic green chemistry. Recently research on bismuth-based hybrids for photovoltaic applications has attracted lots of attention, together with the rapid development of bismuth-based solar cell studies. The potential to be a PV material has been proven by bismuth sulfide [97], bismuth iodide [98], [99], bismuth chalcogenides [100], [101], hybrid bismuth iodide [102], silver bismuth iodide [103], and bismuth-based double perovskites [104]. So far, an efficiency of 4.3% has been achieved by inorganic Ag_3BiI_6, 2.5% has been achieved by $Cs_2AgBiBr_6$ double perovskite, and 1.64% has been achieved by hybrid $[CH_3NH_3][Bi_2I_9]$ [105]. The above-mentioned bismuth-based solar absorbers also demonstrated good air-stability, illustrating the promising potential to develop bismuth-based lead-free photovoltaics.

2.9 Solar Device Structure

A conventional solar device (Figure 2.12) consists of an antireflective coating, glass, a FTO electrode (front contact), an electron transport layer (ETL), a perovskite layer, a hole transport layer (HTL), and another electrode layer (back contact). In a typical PSC, the light-absorbing perovskite is sandwiched between an electron transport layer (ETL) and a hole

transport layer (HTL). When irradiated with sunlight, electron and hole pairs are generated in the perovskite as it absorbs incident light photons with energy larger than the band gap of the perovskite. The electric field between the ETL (n-layer) and the HTL (p-layer) causes the electrons to migrate to the n-layer, while the holes are transported to the p- layer. The ETL extracts electrons from the perovskite layer and prevents electrons in the FTO from recombining with holes.

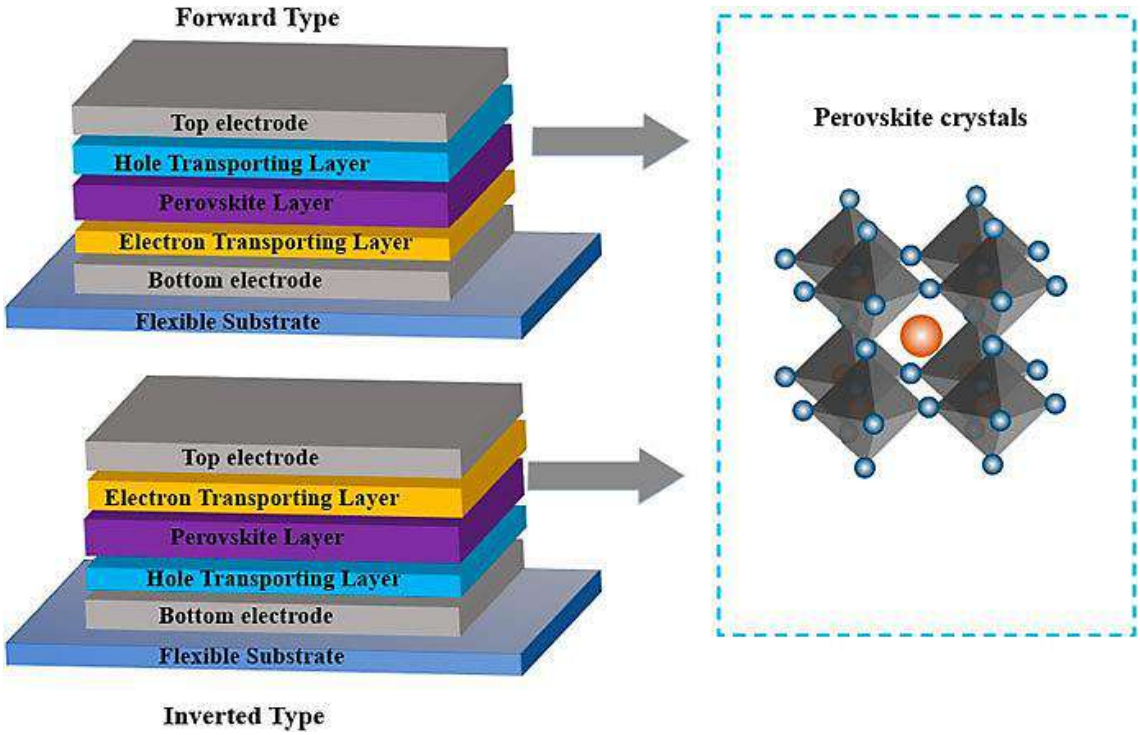


Figure 2.12: Schematic diagram for perovskite solar cell.[106]

2.10 Electron Transport Layer (ETL)

The electron transport layer (ETL) plays a vital function in extracting and transporting photogenerated electrons, modifying the interface, aligning the interfacial energy level, and minimizing the charge recombination in perovskite solar cells.

The electron transport layer for n–i–p architectures are largely consisted of metal oxides such as TiO_2, ZnO, and SnO_2 [107].

Fast electron collection at the electron transport layer (ETL) requires strong electron mobility of the ETL and also the ability to block the passage of holes through it. In addition, the ETL must have a broad bandgap to allow a large part of the electromagnetic spectrum to pass through it and proper band alignment to transport electrons from the perovskite efficiently

[108]–[111]. Taking these characteristics into consideration, numerous studies have found that certain oxides, and in particular, TiO_2, are suitable ETL candidates [112]–[116].

Azri et al. (2019) [117] studied various ETL and HTL materials for an n-i-p perovskite ($MAPbI_3$) solar cell using SCAPS simulator and proposed several ETL and HTL materials to enhance the performance. Among the proposed ETL materials, it was found that Zinc oxide (ZnO) and titanium dioxide (TiO_2) are the most suitable ETLs as they provided a higher efficiency than other ETL materials.

For the ETL in PSCs, TiO_2 has been the primary material of choice owing to its non-toxic nature, high chemical stability, transparency to solar radiation, and easy fabrication. In addition, TiO_2 offers a fast injection of photogenerated electrons from the perovskite layer. A recent study by Zhang et al. demonstrated PSCs with dopamine-capped TiO_2 nanoparticles as the ETL, providing long-term stability and enhanced charge-extraction efficiency at the TiO_2 and perovskite interface.

2.11 Hole Transport Layer (HTL)

The hole transport layer (HTL) should first extract holes and block the electrons from the perovskite active layer, guiding or preventing their respective transfer to the electrode. These two properties are respectively intrinsically related to the Highest Energy Occupied Molecular Orbital (HOMO) and its Lowest Energy Unoccupied Molecular Orbital (LUMO) values of the HTL. A HOMO level close enough to the valence band of the active layer will allow for an efficient transfer of the holes. Additionally, the holes' mobility in the HTL must be relatively high to reduce losses during charge transportation to the contact electrode. Also, HTL should have low electron affinity, and direct bandgap with respect to device layers. Finally, this layer should resist external damage factors (temperature, humidity, oxygen, and illumination).

For the hole transport material, poly[bis(4-phenyl)(2,4,6-trimethylphenyl)amine], also known as PTAA, has traditionally been a standard choice [118]. PTAA has a few disadvantages, including the fact that it is readily degradable, has poor mobility, and is very costly as well [119], [120]. Among the proposed materials by Azri et al.[117], Copper (I) thiocyanate (CuSCN) was the appropriate one for the HTL material.

Solution-processed pseudohalide, such as copper thiocyanate (CuSCN), has recently emerged as other effective and robust p-type inorganic semiconductors which can achieve high quality ordered thin films even from room temperature processing [121], [122]. CuSCN is a highly transparent intrinsically p-type semiconductor which has also been shown to work effectively in transparent electronic devices [121], [122] and as an efficient HTL in organic solar cells [123], [124] and hybrid perovskite solar cells [125]–[127]. A CuSCN layer with 600–700 nm thickness was successfully used as a low-cost replacement for spiro-OMeTAD (2,2′,7,7′-tetrakis-(N,Np-di-methoxy-phenylamino) 9,9′-spirobifluorene) in conventional n–i–p planar and mesostructured perovskite solar cells achieving PCE as high as 12.4% [127], [128].

Alternative materials that combine the processing versatility of CuSCN with the p-type character include various metal oxides such as copper oxide (Cu_2O), tin monoxide (SnO), and nickel oxide (NiO) [129]–[131].

Copper oxide (Cu_2O) is a p-type semiconductor with very high hole mobility (200 cm^2/Vs) and low electron affinity (3.2 eV). Due to its well-suited characteristics, Cu_2O has been used as an HTL in PSCs in recent years [132]–[134].

2.12 Back Contact Electrode Layer

Various metals have traditionally been investigated as viable back contact [135]. Work functions of the different back contact metals used in PSCs are listed in the following table.

Table 2.1: Work functions of different metals used as back contact in PSCs [135], [136]

Metal	Work Function (eV)
Ag	4.26
Cr	4.5
Cu	4.65
Au	5.1
Ni	5.15
Pt	5.65

The noble metals Au, Ag, and Pt are commonly used as back contact electrode materials [137]. Jani M *et al.* [138] studied the effects of various back contact metals on the device performance (PCE) of Cs_2TiBr_6 double halide perovskite. Jani M *et al.* found that PCE increases with increasing work function and selected Au (work function of 5.1 eV) as an

optimal back contact metal. Hence, analyzing various studies, Au has been chosen as the metal back-contact in our SCAPS simulation.

2.13 Tandem Solar Cell

Tandem solar cell provides a way to increase the efficiency by splitting the spectrum and use different bandgap absorption layers that is optimized to each section of the spectrum. Multi-junction (tandem) solar cells mitigate losses that come from carrier thermalization by using several semiconductor layers with different bandgaps. In a double-junction tandem, a 'top cell' with a large bandgap absorbs high-energy photons but permits lower energy photons to pass through to be absorbed in the 'bottom cell', which has a small bandgap. In this way, the higher-energy photons generate a high voltage in the top cell while the low-energy photons are absorbed in the bottom cell, raising the obtainable efficiency of the combined tandem cell over that of either a single high- or low-bandgap solar cell.

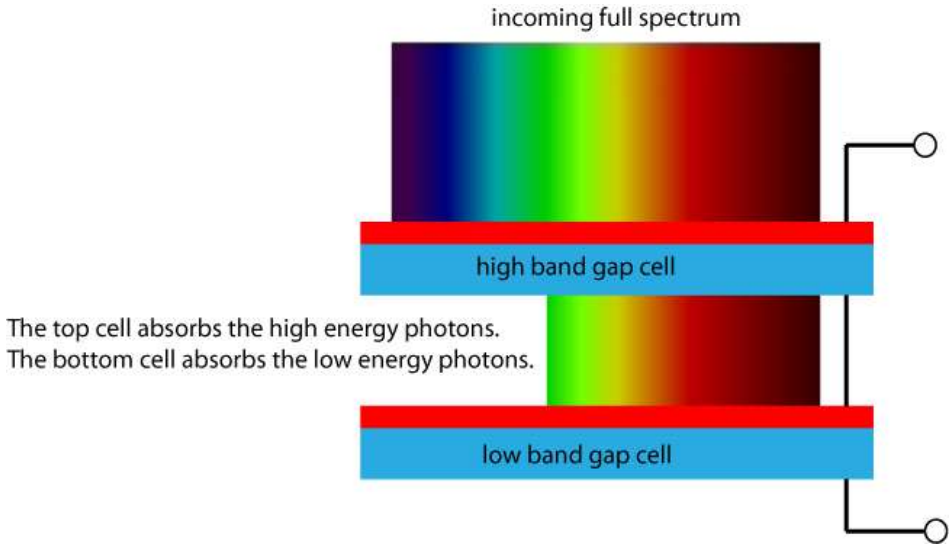


Figure 2.13: Series connected tandem solar cell [46]

Tandem solar cells can either be individual cells or connected in series. Series connected cells are simpler to fabricate but the current is the same though each cell so this constrains the band gaps that can be used. The most common arrangement for tandem cells is to grow them monolithically so that all the cells are grown as layers on the on substrate and tunnel junctions connect the individual cells.

As the number of bandgaps increases the efficiency of the stack also potentially increases. In reality, the semiconductor materials do not exist to allow for arbitrary materials with a specific bandgap and of high quality.

For a tandem solar device configuration, top and bottom cells are traditionally selected based on their spectral efficiency, which is the fraction of a photon's energy of a given wavelength converted into electrical energy. So, a top cell will only improve the efficiency of a bottom cell if it has a higher efficiency resolved at a given wavelength. Due to the good bandgap tunability of perovskite materials, it has become a popular choice for tandem solar cells.

According to Leijtens *et al.*[139], the most promising characteristics for top cells in tandem devices are displayed by the perovskites whose bandgap are in the range of (1.55–1.63) eV [140], [141]. $MAPbI_3$ (1.55-eV bandgap) has been successfully employed as the top cell absorber in different two- and four-terminal tandem devices to achieve a range of PCE values, including 25.20% [142], 21.2% [143], 18.1% [142], and 13.7% [144].

2.14 Machine Learning (ML)

ML is an interdisciplinary subject that combines knowledge of computer science, statistics, mathematics and engineering to form an important branch of artificial intelligence. The most common application of ML is to construct a statistical model used for data analysis and prediction. The main purpose of ML aims at evaluating or predicting the objects after training the model with historical data and specific conditions. With the development of materials genome initiative (MGI), ML keeps playing an important role in materials science with the ability of mapping the relationships and trends through the available data without the physical mechanisms. In addition, the constructed ML model could also be applied again for reverse discovery of high-performance materials. To screen virtual materials with desired properties, ML could be allowed to develop quantitative structure property relationships to predict the properties of virtual materials.

2.14.1 Background

The traditional way to develop materials is usually based on trial and error, continuous synthesis and characterization keep trying until the properties of virtual materials meet the target. The method requires a long-time study on a limited quantity of materials and complicated experimental procedures, which can be a time-consuming and expensive

endeavor. Besides, the discovery of high-performance materials needs a long cycle from experimental design to commercialization. With the development of synthesis and characterization techniques, the corresponding data become more and more complex.

It is a great challenge to figure out the relationship between materials descriptors and properties by traditional experimental methods. To overcome this shortcoming, material simulated methods, including Density Functional Theory (DFT) [145], Monte Carlo simulation and molecular dynamics are employed to explore the relationship between the structural, compositional, and technological descriptors and performance of materials at different scales [146], [147]. However, most computational methods only aim at a specific system, leading to an unbearable amount of computation for complex systems. Some theoretical methods still cannot meet the requirements of quantitative description of material properties. Moreover, computational simulation methods require high computational costs and professional skills.

In recent years, artificial intelligence (AI), known as the 'fourth paradigm of science', has attracted worldwide attentions [148]. Since the 1980s, machine learning (ML) has been the core of AI for the power of reorganizing existing knowledge structures and mining implicit relationships. ML can extract valuable information from existing data, even failed experimental data [149], [150]. For material science, ML has been becoming a powerful tool to assist design and screen various materials. A series of achievements about ML have been made in superconductor, photovoltaic materials and high entropy alloys [112], [113], [116]. Hence, using ML for future work will be helpful in the development of perovskite materials.

2.14.2 Workflow

The workflow of ML in materials generally is divided into four steps: data preparation, feature engineering, model selection, and model evaluation.

2.14.2.1 Data Preparation

Data preparation includes data collection and data manipulation. Materials data could be generally obtained through publicly available materials databases, published papers, experimental data of the same standard, data journals, and density functional theory (DFT) calculations. The latest data can be obtained by searching the publications, but it is time-

consuming and laborious. Data from data journals and databases can be obtained in a short time, but the latest data are generally not available in a timely manner.

The dataset used for ML usually contains dependent and independent variables associated with the materials. Independent variables, also known as features or descriptors, refer to the representative information related to the structure and characteristics of materials, including the chemical composition, atomic or molecular parameters, structural parameters, as well as the technological conditions for synthesis process.

Data can be collected from available databases or published papers. The database contains many types of data, which could be generated from experiments, simulations and ML. A large amount of data could be collected from database, while the reproducibility of the data is uncertain, which might not ensure the quality of data. To guarantee the quantity and quality, data should be collected from the authoritative databases if the data are available. Using autonomous workflows to generate data could be a very convenient and fast way, but the quality of data obtained in this way may be inferior to the data from the database. If the database or autonomous workflows does not obtain the needed data, the dataset could also be generated through lab-scale calculations. Generally, ML models constructed by the calculated data have relatively good evaluation metrics. However, the calculations of complex materials may take up too many calculation resources and take a long time.

The origin data collected from computational simulations or experimental measurements are often presented with incompleteness, noise, and inconsistency [114]. Thus, data manipulation should be performed to ensure consistency and integrity. Specifically, individual data need to be reformatted into a single tabular form, imputed missing values, eliminated erroneous or incomparable data points, and normalized and rescaled the data. Data standardization can improve model accuracy and convergence speed. The results of several ML algorithms can vary with whether any standardization or scaled.

2.14.2.2 Feature Engineering

Feature engineering, including feature extraction, feature selection, feature construction and feature learning, is the process of using domain knowledge regarding the data to create features that allow machine learning algorithms to function. Feature engineering is fundamental to the application of machine learning and is both difficult and expensive.

Features are generally derived from known properties of the constituent elements, such as atomic radius and electronegativity. The quantity of features should be less than that of dataset samples for effectively training and avoiding overfitting. Therefore, feature selection should reduce the dimension of input space as much as possible without losing important information. In particular, redundant and high self-correlation features should be removed to guarantee the efficiency and accuracy of models [115], [135], [136], [138]. Reasonable material features should meet the following three conditions of perfect representation of material properties, sensitive to target properties, and easy to obtain [151].

2.14.2.3 Model Selection

ML algorithms could be briefly divided into two categories: supervised learning and unsupervised learning. Supervised learning is the process of using a set of samples with known labels to adjust the parameters of the models and achieve the required performance, which be further divided into regression and classification. With the target property being a continuous value, the process is called regression. If the target is a discrete value, the process of searching the prediction function is called classification. Generally, the best model is obtained by comparing multiple algorithms. The criteria of algorithm selection are mainly based on the results of cross validation and independent test. The commonly used evaluation metrics include mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), determination coefficient (R2), correlation coefficient (R) for regression; confusion matrix, precision, recall, receiver operating characteristic curve (ROC), and area under ROC curve (AUC) for classification.

It is important for researchers to construct optimal models that map existing data without overfitting or underfitting [152]. After the model selection, it is generally necessary to optimize the internal hyper-parameters of the model algorithm to balance overfitting and underfitting [153]. Overfitting refers to that the model focuses too much on each individual data point in the training set, the unknown data cannot be well predicted. While underfitting means that the model is too simple to capture the general information of the training set, resulting in a large deviation. Models insensitive to hyper-parameters usually do not require repeated adjustments to achieve satisfying results [154]. For the hyper parametric sensitive model, the performance of the model will be closely related to the selection. It is necessary to adjust different hyper-parameters to achieve the robust ML model.

Decision Tree model, popular in ML for its intelligibility and simplicity, was also explored for fitting the data [155], [156]. The model is easy to visualize and interpret, and provides an elegant path for identifying the most significant features and relations between multiple variables – even if the outputs are not continuous, and manifest as clusters of data. The model therefore gives exceptionally accurate predictions [157], [158]. The Decision Tree model can be trained using the random forest regression, gradient boosting regression etc [159]. Random forest optimizes the prediction model with an attractive anti-overfitting feature. The gradient boosting works by building an ensemble of decision trees in a sequential manner. The first decision tree is trained on the input data and output variable. The subsequent decision trees are then trained on the residual errors of the previous tree.

Another commonly used supervised ML algorithm is linear regression. The goal of linear regression is to minimize the difference between the predicted values and the actual values in the dataset. It works by finding a line that best fits the data points in a scatter plot. The line is defined by an equation of the form $y = mx + b$, where y is the predicted value, x is the input value, m is the slope of the line, and b is the y-intercept. The goal of linear regression is to find the values of m and b that minimize the difference between the predicted values and the actual values in the dataset.

A voting classifier is a machine learning estimator that trains various base models or estimators and predicts on the basis of aggregating the findings of each base estimator. The aggregating criteria can be combined decision of voting for each estimator output. There are two types of voting: hard voting and soft voting. Hard voting entails picking the prediction with the highest number of votes, whereas soft voting entails combining the probabilities of each prediction in each model and picking the prediction with the highest total probability.

Shapely Additive exPlanations [160], shortly known as the SHAP method, is used to determine the relative importance and impact of each feature on the final predicted target values. The SHAP plot eliminates the necessity of feature importance and partial dependence plots by intuitively representing the impact of each feature values along with their positive or negative effect on the final output values.

2.14.2.4 Model Evaluation

The core of ML is to realize accurate prediction of unknown samples based on known information. There are inevitably some statistical errors in the calculation, which should rationally be checked and evaluated in the process model evaluation for the subsequent model application. There are three commonly used model evaluation methods: independent test, cross validation, and bootstrapping.

In general, the generalization error of the model can be evaluated by testing, but the goal of the model is the prediction of unknown samples. Therefore, a testing set is needed to test the generalization ability. The error obtained with the testing set can be taken as an approximation of the generalization error. The smaller error of independent test usually indicates the stronger generalization of the model available. It is worth noting that the independent testing set should be mutually exclusive to the training set [161].

Cross validation (CV) could be used to evaluate the reliability of the ML models. In the CV, the input data is divided into k mutually exclusive subsets of the similar size, each subset is generated by 'stratified samples'. Then, the union of k-1 subsets is used as the training set with the remaining one used as the testing set. After k times of training and testing, all test results are averaged to represent the final ML performance. The stability and fidelity of the evaluation results of the CV method depend to a large extent on the value of k. Hence, the CV method is usually called k-fold cross validation. In the k-fold CV, k is a specified number, the commonly used values are 5, 10, and 20.

2.14.3 Applications of Machine Learning

The final aim of ML is to predict the target variables of unknown samples based on the trained model. The three major scenarios of model application are high-throughput screening (HTS), inverse design, and the development of online prediction programs. HTS uses the constructed model to predict the target variables of a huge number of virtual samples in order to filter out samples with high performance potential to guide experimental synthesis. The inverse design can be used to obtain the features of designed samples via the inverse projection method, which is an effective way to realize the material from properties to composition. The prediction of designed samples helps screen out candidates with breakthrough performance and improves computational efficiency. The development of an online prediction program makes it possible to quickly achieve the prediction of target

properties by simply inputting the necessary information, such as a chemical formula, on the input page, which facilitates the extension of model application and effectively realizes model sharing.

Based on many existing experimental and computational data, ML technology has gradually played an important role in the perovskite's discovery [162]. Up to now, there are about 1000 perovskites that have been developed through experiments. There is still a huge space for stable perovskites to be excavated.

2.15 Perovskites Solar Cell Research by Machine Learning

The main target of researchers in PV research is to continuously improve the performance and reduce the costs of solar cells, making them a more accessible, environment-friendly, and sustainable energy source for the world. Solar cell research is focused on several key factors, including: material development, process optimization, improved energy conversion, increased durability, improved flexibility, increased cost-effectiveness.

Based on many existing experimental and computational data, ML technology has gradually played an important role in perovskite discovery. As ML has promising potential in solar cell applications, researchers are exploring ML to predict and modulate perovskite absorber material's optical and electrical properties [50], [163]–[169] utilizing either the theoretical [163], [170] or experimental values of the perovskite material to build the ML model. Up to now, many scholars have brought ML method into the field of computational materials.

Ç. Odabaşı and R. Yıldırım *et al.* built a database containing 404 organo-lead halide perovskites cells stability data, using the apriori algorithm in ML to analyze the influence of cell manufacturing materials, manufacturing processes, and storage conditions on cell stability [171].

Guo *et al.* proposed a machine learning framework to investigate thermodynamic stability and band gap of lead-free halide double perovskites [172]

Lu *et al*. used the ML algorithm to assist DFT calculations and found six orthorhombic lead-free hybrid organic-inorganic perovskites (HOIPs) with suitable bandgap for solar cells and room temperature thermal stability from 5158 undiscovered HOIPs [170].

Weng *et al*. guided the design of new oxide perovskite catalysts and improved the activity of oxygen evolution reaction (OER) by the Symbolic regression (SR) algorithm in ML [173]. G. Pilania predicted the bandgaps of the double perovskites by using the KRR algorithm and the DFT method and the RMSE was around 0.37eV [166].

Jino *et al.* [163] investigated how the Gradient Boost Regression Trees (GBRT) ML method can be used for designing Pb-free perovskites. They developed a dataset containing the electronic structures of candidate halide double perovskite. Using the dataset, the GBRT ML model was implemented to predict the values of heat formation and bandgap.

2.16 SCAPS- 1D Software

SCAPS-1D (Solar Cell Capacitance Simulator One dimension), devised by the University of Gent's ELIS department [174], [175], is used to simulate single-junction perovskite solar cells' numerical simulations. SCAPS-1D software can numerically solve one-dimensional equations that govern semiconductor materials and calculates energy bands, current-voltage (I-V), and capacitor-voltage characteristic curves of solar cells. It analyses the physics of the model and it explains the recombination profiles, electric field distribution, carrier transport mechanism and individual current densities [176].

SCAPS is originally developed for cell structures of the $CuInSe_2$ and the CdTe family. Several extensions however have improved its capabilities so that it is also applicable to crystalline solar cells (Si and GaAs family) and amorphous cells (a-Si and micromorphous Si) [177]. This software can predict device characteristics such as current density–voltage curve, quantum efficiency, energy bands, and other specific responses of the planar thin-film structure under illumination. To examine the performance of configured solar cells, different types of HTLs and perovskites are simulated through a comparative survey. Then, the effects of thickness and other parameters on PCE can be studied. To improve the real-life application of this simulation, the adverse effect of the charge carrier recombination at interfaces on the efficiency of cells are taken into account. The most recent version of SCAPS includes:

- up to 7 semiconductor layers
- recombination mechanisms: band-to-band (direct), Auger, SRH-type
- defect levels: in bulk or interface; their charge state and recombination are accounted for

- illumination: a variety of standard and other spectra included (AM0, AM1.5D, AM1.5G, AM1.5Gedition2, monochromatic, white etc.)
- illumination: from either the p-side or the n-side; spectrum cut-off and attenuation
- working point for calculations: voltage, frequency, temperature
- the program calculates energy bands, concentrations and currents at a given working point, J-V characteristics, spectral response (also with bias light or voltage)
- batch calculations possible; presentation of results and settings as a function of batch parameters
- a script language facility to run SCAPS from a 'script file'; all internal variables can be accessed and plotted via the script.
- a built-in curve fitting facility and a panel for the interpretation of admittance measurements

This software has been extensively applied to extract basic electrical characteristics and determine the PCE of PSCs. Despite the considerable success of SCAPS-1D in predicting PSC performance, the accuracy, with which different transport layers are modeled, can be compromised by using built-in absorption spectrums in SCAPS-1D. This is because device simulators like SCAPS-1D employ simplified combination of power laws to derive the absorption spectrum as a function of photon energy resulting into unrealistic outputs, creating discrepancies between simulation and experimental results.

2.16.1 Governing equations and critical SCAPS output parameters

SCAPS solves three systems of equations for the carrier: transport, Poisson, and continuity. The following is the carrier continuity equation [44]:

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\nabla . J_p - R(x) + G(x) \tag{2.10}$$

$$\frac{\partial n}{\partial t} = \frac{1}{q}\nabla . J_n - R(x) + G(x) \tag{2.11}$$

The electron-hole current densities, respectively, are denoted by J_n and J_p, the recombination and generation rates are denoted by *R(x), G(x)* respectively, and the position-dependent electron and hole concentrations, respectively, are denoted by *p(x)* and *n(x)*.

The drift-diffusions of electron-hole pairs are described by the following equations:

$$J_n = \ qn(x)\mu_n \frac{dv(x)}{dx} + qD_n \frac{dn(x)}{dx} \tag{2.12}$$

$$J_p = \ -qn(x)\mu_p \frac{dv(x)}{dx} + qD_p \frac{dn(x)}{dx} \tag{2.13}$$

where μ_p and μ_n denote the electron and hole mobility, respectively, and D_n and D_p denote the electron and hole diffusion constants, respectively. SCAPS-1D solves the Poisson equation and continuity equations for electrons and holes together by taking the appropriate boundary conditions at the interfaces and contact into consideration. The fill factor *(FF)* aof the device is defined as follows:

$$FF = \frac{J_{Mp}, V_{Mp}}{J_{SC}, V_{OC}} \tag{2.14}$$

where V_{Mp} and J_{Mp} represent the voltage and current density at the maximum power points. The short-circuit current density and the open-circuit voltage are denoted by J_{sc} and V_{oc} respectively.

The power conversion efficiency (η) is calculated using the following equation:

$$\eta = \frac{V_{OC}.J_{SC}.FF}{P_{in}} \tag{2.15}$$

where P_{in} is the input power obtained from incident light.

CHAPTER 3: METHODOLOGY

3.1 Data Collection and Manipulation

There are three elemental sites (A, B, X) in the formula of simple perovskite structure ABX_3 where, A is a monovalent cation (+1), B is a divalent cation (+2) and X is a monovalent halide anion (-1). The oxidation state of the cations or anions may change depending on the type of perovskites. The halide double perovskite structure corresponds to a general chemical formula of $A_2BB'X_6$, where A is a monovalent cation, B is a monovalent cation, B' is a trivalent cation and X is a monovalent halide anion. In case of double perovskites, mixed cation or anion was considered as a single elemental site and a weighted average for each of the sites has been used in those cases.

Different types of perovskites containing, MA^+ or $[CH_3NH_3]^+$, FA^+ or $[CH(NH_2)_2]^+$, Cs^+, Ag^+, K^+, GA^+, Rb^+ cations in the A site, and Pb^{2+}, Sn^{2+}, Ge^{2+}, Bi^{3+}, Sb^{3+}, Cu^{2+}, Na^+, Li^+, In^{3+}, Ti^{3+} cations in the B site and halogens in the X site, are selected for the training dataset.

Twenty properties of each elemental site of the three sites (A, B, X) in perovskites give sixty (60) descriptors for the training dataset. The input variables also known as ‘features’ used to train the ML models are shown in the Table 3.1.

The symbol 'x' in the name of elemental features represents elemental sites, where its value 1, 2 and 3 indicate A site, B site and X site, respectively.

The variables that could potentially affect the bandgap of perovskites were selected as input variables or features. Some of these input parameters were utilized in previous studies [178], [179]. But there are some new input parameters that have been utilized in this model.

Table 3.1: Elemental properties (features) used in the study

SL.	Name of the features	Feature names used in the dataset	Ref.
1	Atomic mass	atomic_mass_x	[180]
2	Molar volume	molar_volume_x	[181]
3	Atomic radius (Å)	atomic_radius_x	[180]
4	Ionic radius	ionic_radius_x	Given in the next table 3.2
5	Covalent radius (Å)	covalent_radius_x	[180]
6	Van der Waals radius (Å)	van_der_waals_radius_x	[181]
7	Electron negativity (Pauling scale)	electron_negativity_x	[180]
8	Electron affinity (kJ/mol)	electron_affinity_x	[180]
9	Electrical resistivity (Ohm m)	electrical_resistivity_x	[181]
10	Thermal conductivity ($Wm^{-1}K^{-1}$)	thermal_cond_x	[181]
11	Boiling point (K)	boiling_point_x	[181]
12	Melting point (K)	melting_point_x	[181]
13	Critical temperature (K)	Tc_x	[181]
14	Enthalpy of fusion (kJ/mol)	enthalpy_of_fusion_x	[181]
15	First ionization energy(kJ/mol)	first_ionization_energy_x	[181]
16	s electrons	s_x	[181]
17	p electrons	p_x	[181]
18	d electrons	d_x	[181]
19	f electrons	f_x	[181]
20	Valance	valance_x	[181]

The independent or input variables for training dataset were obtained through publicly available materials databases, namely- RSC Periodic Table, Web Elements, and Materials Project Database.

Ionic radii of each of the A, B, X site was determined based on the oxidation state of each ion in the perovskite and it was calculated by their weight in the composition. The ionic radii of the cations, anions and organic molecules used for training the machine learning (ML) models have been shown in Table 3.2.

Table 3.2: Ionic radii of the cations, anions and organic molecules

Ion	Effective ionic radii (r_{eff}) [pm]	Ref.
Formamidinium, $[CH(NH_2)_2]^+$	253	[182]
Methylammonium, $[CH_3NH_3]^+$	217	[182]
Potassium, K^+	164	[183]
Rubidium, Rb^+	172	[183]
Caesium, Cs^+	188	[183]
Strontium, Sr^{2+}	118	[183]
Barium, Ba^{2+}	135	[183]
Cadmium, Cd^{2+}	78	[183]
Germanium, Ge^{2+}	73	[183]
Tin, Sn^{2+}	115	[79]
Lead, Pb^{2+}	119	[183]
Tin, Sn^{4+}	69	[183]
Antimony, Sb^+	76	[183]
Bismuth, Bi^{3+}	103	[183]
Tellurium, Te^{4+}	97	[183]
Fluoride, F^-	129	[184]
Chloride, Cl^-	181	[184]
Bromide, Br^-	196	[184]
Iodide, I^-	220	[184]
Silicon, Si^{2+}	113	[185]

The target property that the models will predict is the bandgap energy of perovskites. We also searched for the literatures reporting the experimental bandgap of the perovskites. The experimental bandgap value of various perovskites was collected from research articles published in various peer reviewed journals including ACS, Elsevier, Wiley, RSC databases, Nature Group, and Science.

All these input features along with the target property are assembled into an Excel spreadsheet. Then, the data points were cleaned by removing the duplicate data points with same material composition and bandgap values. The data with the same chemical composition but different bandgap values (output data) are kept in the dataset due to the different conditions and environments of the experiments. Finally, 99 data points were obtained for training the ML models after data manipulation. The perovskites used for training are all experimentally synthesized and majority of them are formable according to

perovskites tolerance factor and octahedral factor. The perovskites used in the training dataset with their experimental bandgaps are attached in Appendix.

For various reasons, many real-world datasets contain missing values, often encoded as blanks, NaNs or other placeholders. Such datasets however are incompatible with scikit-learn estimators which assume that all values in an array are numerical, and that all have and hold meaning. A basic strategy to use incomplete datasets is to discard entire rows and/or columns containing missing values. To ensure consistency and integrity, data was reformatted into a single tabular form, missing values were imputed, erroneous or incomparable data points were eliminated. ML algorithms used in this study required data to be normalized and rescaled. However, eliminating redundant or irrelevant features is an effective way to improve the performance of the predictive ability of the model.

3.2 Building Predictive Machine Learning Models

In the present work, four supervised ML models namely- Linear Regression, Random Forest Regression, Gradient Boosting Regression and Voting Regression, were built for the training and testing phases. These models were written and built in Jupyter Notebook using python programming language with the simple and efficient data analysis tool named Sci-kit Learn. The Jupyter Notebook is an original web-based application for creating and sharing computational documents. However, the basic outline is described in this section briefly.

At first, various important packages required for ML modeling are imported in the Jupyter Notebook. Then, the Data Frame including the independent input parameters as listed in Table 3.1 and the bandgap which is the dependent variables (also called ‘target variables’) is created using Pandas Data Frame in scikit-learn. Now, the columns that are not useful input parameters, are dropped from the Data Frame and the remaining columns are used for training variables and assigned to ‘X’ representing input training variables. Similarly, the target feature (bandgap) is extracted and assigned to ‘y’. First 99 rows in Data Frame were selected for training and testing purpose. The features were scaled for better performance. Remaining 113 datapoints are selected for predicting their bandgaps.

Hence, the Data Frame was split into two single Data Frames: one containing the features, and the other containing the target variables. After splitting the raw Data Frame into features

and target, 90% of the dataset was utilized to train the model, and the rest (10%) was used as the test set. The test set remained unseen – that is, independent from the model during its execution for predicting the output.

Then the ML models were built according to the codes attached in Appendix. After building the models, their performance is evaluated in terms of R^2, RMSE (root mean squared error) and cross-validation score. The relative importance of the features is also analyzed with SHAPley Additive exPlanations tool (SHAP) in scikit-learn.

3.3 Selecting Suitable Compositions for ML Predictions

As $MAPbI_3$ is one of the most promising perovskites among halide-based perovskites, it has been taken as a prediction data point. Now, as we want to make the bandgap prediction faster and hence to contribute in reducing the toxicity of perovskites, the Pb site (B site) in $MAPbI_3$ is replaced with Sn, Ge, and Si for bandgap prediction. All of these elements have similar electronic configurations as Pb, and they are group 14 elements. So, they might give similar results as the Pb site produced in $MAPbI_3$ perovskites. All the elemental features used in training the ML models are needed to be given as the input features for obtaining the predicted bandgap of these perovskites. After analyzing the predicted bandgaps of these perovskites, suitable perovskite with an optimum bandgap will be simulated with SCAPS.

3.4 SCAPS Analysis

The present work includes a numerical simulation of single junction solar cell using SCAPS-1D to obtain its J-V parameters like open circuit voltage (V_{oc}), Short circuit current (J_{sc}), fill factor (FF) and power conversion efficiency (PCE). A suitable absorber with an optimum bandgap obtained from the predicted dataset will be used in this step for SCAPS analysis. Also, the predicted bandgap of pristine $MAPbI_3$ will be used in the analysis for a comparative study. The simulation will be carried out under the standard *AM 1.5G* solar spectrum with an incident power density of 1000 W/m^2 at room temperature (300 K). A conventional perovskite solar cell consists of an antireflective coating, glass, an FTO electrode, an electron transport layer (ETL), a perovskite layer, a hole transport layer (HTL), and an electrode layer

[186]. In the present research, the device layers constitute (Glass/FTO (Fluorine doped Tin Oxide)/ETL/MA-based halide Perovskite/HTL/Au.

3.5 Overview of Methodology

The summary of the whole working process is shown in a flowchart below.

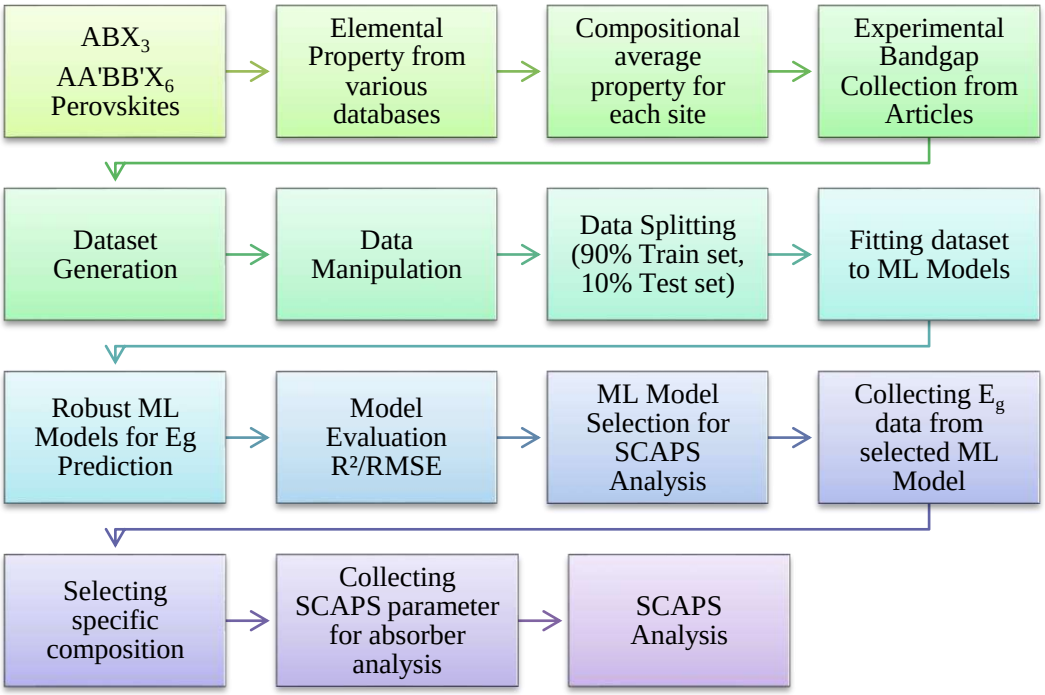


Figure 3.1: Flowchart of methodology

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Performances of ML Models

The developed ML models can predict bandgap values of any simple, double (ABX_3 or $AA`BB`X_6$), or mixed ionic perovskite materials with significant accuracy. Different error metrics such as Root mean square (RMSE), R^2, and cross-validation (CV) score were calculated by *sci-kit learn* to evaluate the reliability and performance of the ML models. R^2 is a performance analyzing metric used in statistics and ML that represents how well the regression model fits the dataset. It determines the amount of variance in the ML models caused by the input variables, whereas the mean squared error (RMSE) measures how close a regression line is to a set of data points. As a rule of thumb, for perfect prediction, it is required to have the R^2 value as close to 1.0 and the RMSE value as close to 0.0.

At first, the *linear regression model* was evaluated by comparing the model's predicted bandgap with the actual bandgap in the dataset. It is found that the linear regression model shows a significantly higher R^2 value of 0.86 (toward R^2=1) and a higher root mean squared error of 0.13. Figure 4.1 shows the comparison of the predicted bandgaps created by the *linear regression model* with the actual bandgaps of various perovskites in the dataset.

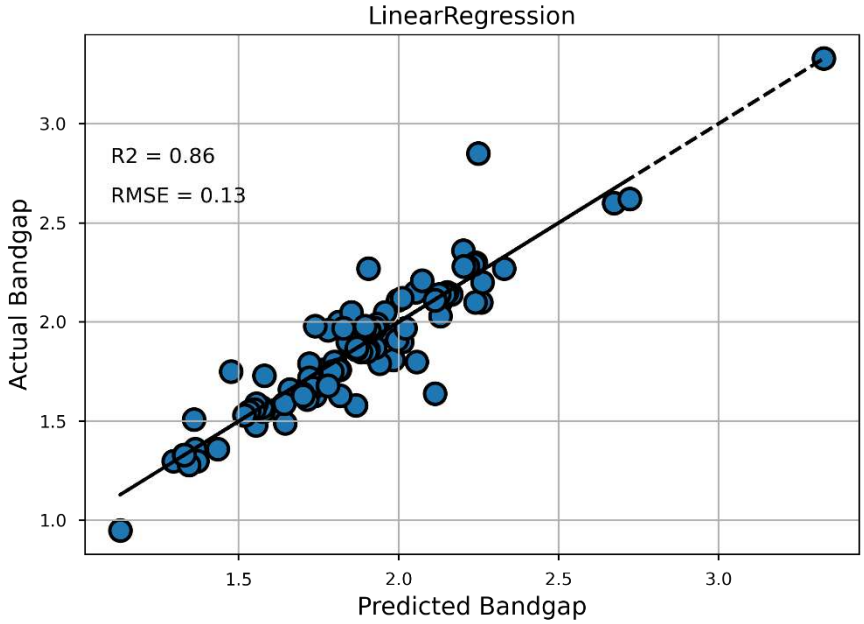


Figure 4.1: Evaluation of Linear Regression Model by R^2 & RMSE.

Next, the performance of the *random forest regression model* is evaluated in the same way as the linear regression model. However, it was found that the R^2 value increased from 0.86 to 0.90, and the RMSE value decreased from 0.13 to 0.11 than the linear regression model (Figure 4.2). Both of the error metrics show significant improvement according to our expectations. Tree-based models, such as the *random forest regression*, typically perform better in complex situations where many input parameters are utilized for training and fitting the ML models.

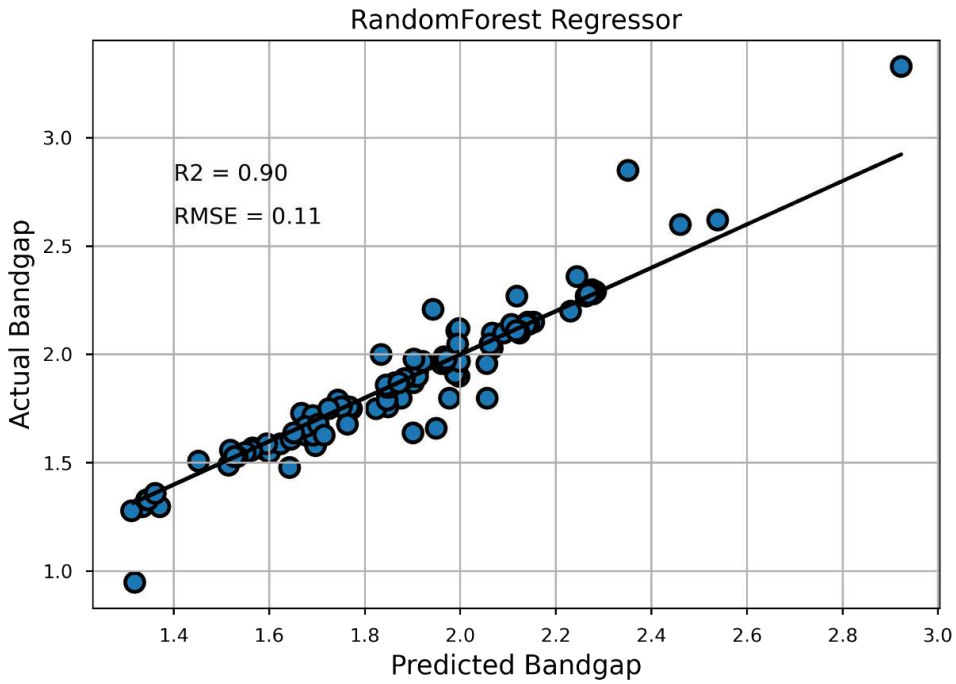


Figure 4.2: Evaluation of Random Forest Regression Model by R^2 & RMSE.

In the same way, the performance is evaluated for the *gradient-boosting algorithm* by R^2 and RMSE. Here, we can observe that the R^2 and RMSE value remain the same as *the random forest model* as can be seen from Figure 4.3. We can assume from this observation that these two models are performing equally. But it is to be noted here that the performance of any ML model is not only determined by the R^2 and RMSE value. There are some other factors that need to be considered during analysis. It will be discussed later in this section.

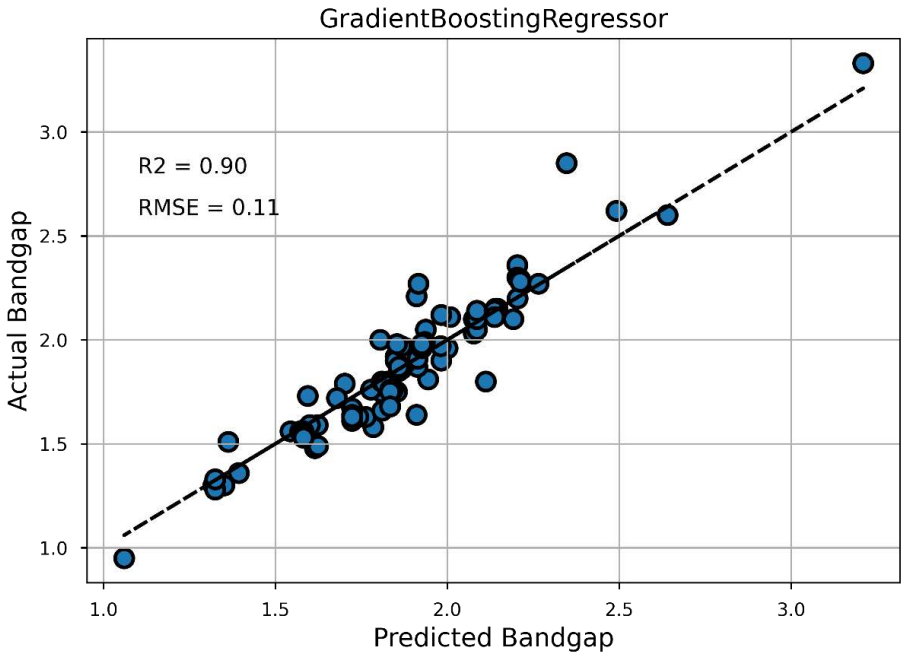


Figure 4.3: Evaluation of Gradient Boosting Regression Model by R^2 & RMSE.

Finally, the *voting regression model* was used to improve model performance, ideally achieving better performance than any single model used in the ensemble. After analyzing and building the voting regression, it was found that the R^2 value significantly improved than any other model (Figure 4.4).

However, the R^2 and RMSE value indicates that the voting regressor model performs the best. But there is a limitation of the voting regressor model. It averages the performance made by all the models individually. There is a chance to make significant errors for any particular composition if any of the models ensembled in the voting regressor make a bad prediction. Voting regression only serves to benefit when the single machine learning model performs at similar levels. A voting estimator built from models with contrasting levels of efficiency may perform erratically.

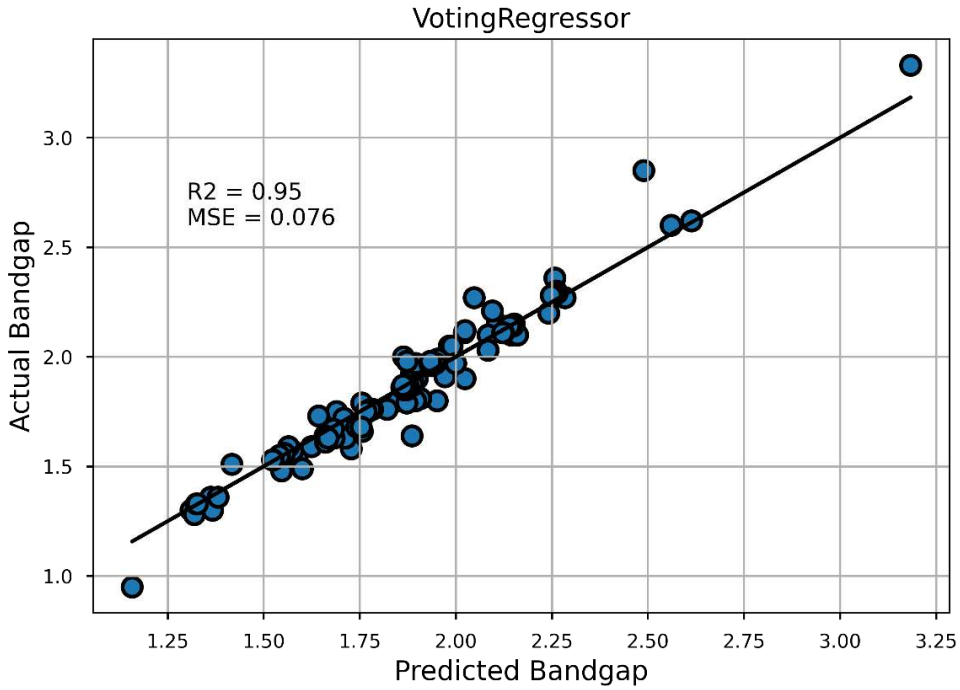


Figure 4.4: Evaluation of Voting Regression Model by R^2 & RMSE.

Overfitting is suspected if the model performs excellent on the training set, but performs badly on the test dataset, i.e., the test accuracy is significantly lower [50], [169], [187]. Cross-validation was done with sci-kit learn to ensure the authenticity of the prediction.

Typically, the cross-validation score is less than the training score because the model fits on training data, and validation data is unseen by the model. Overfitting happens when a model learns the detail in the training data to the extent that it negatively impacts the performance of the model on new data. From the Table 4.1, it is clear that overfitting was found for each of the models developed, as cross-validation scores were significantly less than the model's R^2 score. The linear regression model performs the worst though it has an eye-catching R^2.

Table 4.1: Error metrics and cross validation score summary for the models developed

Models	Fitting Accuracy R^2	RMSE	Cross Validation Score
Linear Regression	0.859750	0.13	-2.201832
Random Forest Regression	0.903741	0.11	0.320438
Gradient Boosting Regression	0.896337	0.11	0.365321
Voting Regression	0.949635	0.08	-

Table 4.1 shows the summary of the Root mean square (RMSE) and R^2 value obtained from the ML algorithms. In our case, cross-validation score of linear regression model was poor. Hence, voting regression combining these three models cannot perform well for all the predicted composition. Some predictions might not be very accurate because of the small number of data points (training sample) used, but these ML models will definitely perform better with higher accuracy if more data are added continuously into the dataset.

Figure 4.5 shows the combined output of four ML models for the first 20 perovskite samples used in the training dataset.

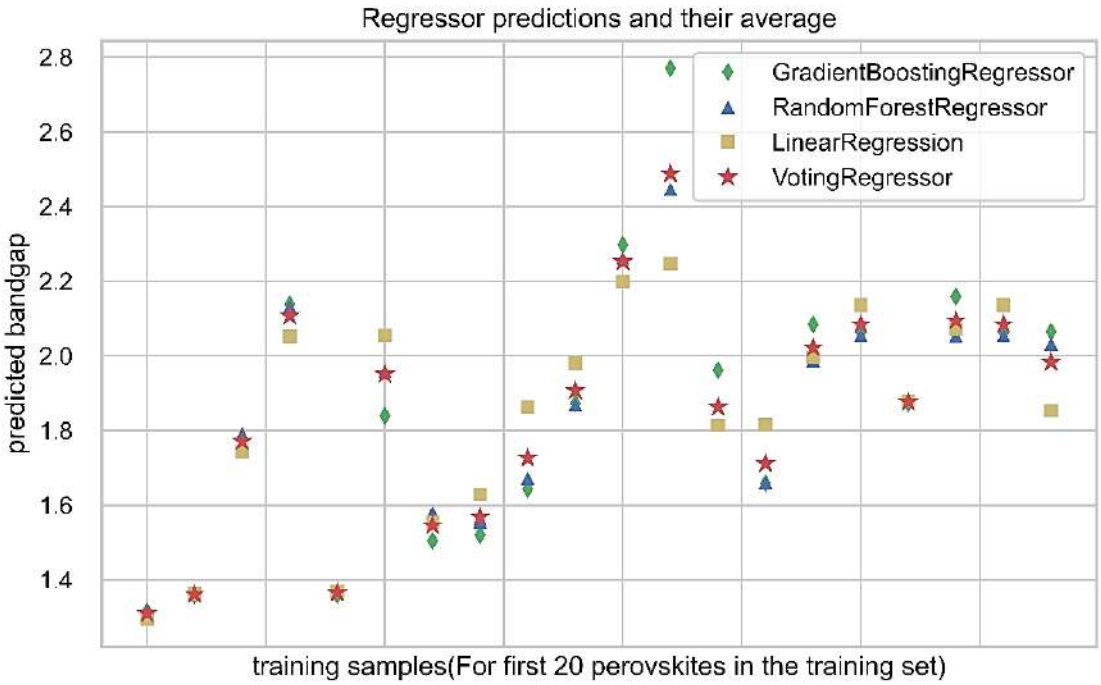


Figure 4.5: Combined output of four ML models for the first 20 perovskites.

4.2 Data Validation & ML Model Selection:

Based on the fitting accuracy and cross validation score, we have selected the voting regression model as the final algorithm due to its better ability to predict the band gaps of new materials. Though, it will not predict accurately for all the compositions, it can predict closely to the actual bandgap for most of the perovskite compositions. After the ML models have been developed, the models are utilized to predict the bandgap for $MAPbI_3$, $MASnI_3$, $MAGeI_3$, and $MASiI_3$. The predicted bandgaps of these perovskites are shown in Table 4.2.

Table 4.2: Predicted band gap of perovskites

Composition	Linear Regression	Random Forest Regression	Gradient Boosting Regression	Voting Regression
$MAPbI_3$	1.575163	1.6001	1.600194808	1.587283
$MASnI_3$	1.365736	1.3364	1.351341499	1.339446
$MAGeI_3$	1.765154	1.8218	1.775320000	1.851335
$MASiI_3$	1.842056	1.7141	1.688991000	1.7834267

In the voting regression model, we can see that $MAGeI_3$ and $MASiI_3$ have higher predicted bandgap whereas $MASnI_3$ has a lower predicted bandgap than that of $MAPbI_3$. The bandgap of a typical good light absorber lies between 1.3 and 1.7 eV [188]–[190]. So, $MASnI_3$ is a potential candidate as a solar cell absorber material. Hence, the Pb site in $MAPbI_3$ is varied by increasing proportion of Sn, and their bandgaps are predicted. The predicted bandgap dataset of $MAPb_{1-x}Sn_xI_3$ created by four ML models are shown in Table 4.3.

Table 4.3: Bandgap dataset for $MAPb_{1-x}Sn_xI_3$ compositions predicted with different ML Models

Composition	index	LR Predicted	RF Predicted	GB Predicted	Voting Predicted
$MAPbI_3$	60	1.575163	1.6001	1.600194808	1.587283
$MAPb_{0.98}Sn_{0.02}I_3$	61	1.570906	1.5986	1.568185132	1.560173
$MAPb_{0.96}Sn_{0.04}I_3$	62	1.566817	1.5958	1.568185132	1.557996
$MAPb_{0.94}Sn_{0.06}I_3$	63	1.562468	1.5958	1.525178996	1.549477
$MAPb_{0.92}Sn_{0.08}I_3$	64	-0.244740	1.6811	1.553973661	1.031885
$MAPb_{0.9}Sn_{0.1}I_3$	65	1.554106	1.5958	1.485522593	1.540762
$MAPb_{0.88}Sn_{0.12}I_3$	66	1.549986	1.5958	1.485522593	1.539369
$MAPb_{0.86}Sn_{0.14}I_3$	67	1.545698	1.5958	1.485522593	1.537996
$MAPb_{0.84}Sn_{0.16}I_3$	68	1.541640	1.5958	1.485522593	1.536541
$MAPb_{0.82}Sn_{0.18}I_3$	69	1.537413	1.5958	1.485522593	1.535086

$MAPb_{0.8}Sn_{0.2}I_3$	70	1.533247	1.5933	1.485522593	1.533733
$MAPb_{0.78}Sn_{0.22}I_3$	71	1.529051	1.5933	1.485522593	1.532421
$MAPb_{0.76}Sn_{0.24}I_3$	72	1.524840	1.5933	1.485522593	1.530997
$MAPb_{0.74}Sn_{0.26}I_3$	73	1.520537	1.5933	1.485522593	1.529552
$MAPb_{0.72}Sn_{0.28}I_3$	74	1.516432	1.5933	1.485522593	1.528169
$MAPb_{0.7}Sn_{0.3}I_3$	75	1.512266	1.6019	1.485522593	1.528265
$MAPb_{0.68}Sn_{0.32}I_3$	76	1.508055	1.5994	1.485522593	1.527687
$MAPb_{0.66}Sn_{0.34}I_3$	77	1.503844	1.5994	1.485522593	1.524791
$MAPb_{0.64}Sn_{0.36}I_3$	78	1.649916	1.5060	1.427605953	1.550524
$MAPb_{0.62}Sn_{0.38}I_3$	79	1.495619	1.5994	1.485522593	1.521963
$MAPb_{0.60}Sn_{0.4}I_3$	80	1.491301	1.6128	1.485522593	1.528405
$MAPb_{0.58}Sn_{0.42}I_3$	81	1.487074	1.6128	1.485522593	1.526991
$MAPb_{0.56}Sn_{0.44}I_3$	82	1.482939	1.6173	1.485522593	1.530385
$MAPb_{0.54}Sn_{0.46}I_3$	83	1.478743	1.6271	1.485522593	1.531878
$MAPb_{0.52}Sn_{0.48}I_3$	84	1.474669	1.6207	1.485522593	1.530444
$MAPb_{0.5}Sn_{0.5}I_3$	85	1.470412	1.4864	1.468747199	1.488196
$MAPb_{0.48}Sn_{0.52}I_3$	86	1.466261	1.4541	1.449902826	1.469316
$MAPb_{0.46}Sn_{0.54}I_3$	87	1.462080	1.4272	1.435081907	1.464589
$MAPb_{0.44}Sn_{0.56}I_3$	88	1.457793	1.4129	1.435081907	1.459696
$MAPb_{0.42}Sn_{0.58}I_3$	89	1.453627	1.4129	1.435081907	1.458312
$MAPb_{0.4}Sn_{0.6}I_3$	90	1.449461	1.4157	1.435081907	1.456949
$MAPb_{0.38}Sn_{0.62}I_3$	91	1.545500	1.4434	1.433363037	1.495631
$MAPb_{0.36}Sn_{0.64}I_3$	92	1.441069	1.4184	1.435081907	1.457304
$MAPb_{0.34}Sn_{0.66}I_3$	93	1.436888	1.4219	1.435081907	1.455920
$MAPb_{0.32}Sn_{0.68}I_3$	94	1.432738	1.4219	1.435081907	1.454486
$MAPb_{0.3}Sn_{0.7}I_3$	95	1.428541	1.4219	1.435081907	1.453133
$MAPb_{0.28}Sn_{0.72}I_3$	96	1.424361	1.4219	1.435081907	1.453026
$MAPb_{0.26}Sn_{0.74}I_3$	97	1.420149	1.4152	1.433363037	1.449430
$MAPb_{0.24}Sn_{0.76}I_3$	98	1.415938	1.4152	1.433363037	1.447975
$MAPb_{0.22}Sn_{0.78}I_3$	99	1.411696	1.416	1.433363037	1.449561
$MAPb_{0.2}Sn_{0.8}I_3$	100	1.407561	1.3968	1.412995359	1.414213
$MAPb_{0.18}Sn_{0.82}I_3$	101	1.40341	1.3968	1.412995359	1.41289
$MAPb_{0.16}Sn_{0.84}I_3$	102	1.399214	1.3968	1.412995359	1.411436
$MAPb_{0.14}Sn_{0.86}I_3$	103	1.395064	1.3968	1.412995359	1.410032
$MAPb_{0.12}Sn_{0.88}I_3$	104	1.390898	1.3968	1.412995359	1.408587
$MAPb_{0.1}Sn_{0.9}I_3$	105	1.38658	1.3646	1.370410166	1.368894
$MAPb_{0.08}Sn_{0.92}I_3$	106	1.382445	1.3646	1.370410166	1.36745
$MAPb_{0.06}Sn_{0.94}I_3$	107	1.378187	1.3389	1.351341499	1.346845
$MAPb_{0.04}Sn_{0.96}I_3$	108	1.374037	1.3389	1.351341499	1.345441
$MAPb_{0.02}Sn_{0.98}I_3$	109	1.369948	1.3364	1.351341499	1.340891

As it can be seen from Table 4.3, linear regression performed poor in predicting the bandgap of $MAPb_{0.92}Sn_{0.08}I_3$ giving a garbage value, voting regression also predicted the bandgap with a large error for this composition. Except some of these anomalies for some compositions, voting regression can combine the overall performance of all three models in a single model. In cases of garbage values, we can utilize the gradient boosting regression, as it has shown a higher value of cross-validation score among the base ML models.

We have also predicted the bandgap for another composition where A site (MA), B site (Pb), and X site (I) in $MAPbI_3$ are replaced with FA (Formamidinium), Sn, and mixed proportion of I and Br, respectively. Hence, the predicted bandgaps are obtained for $FASnI_{3-x}Br_x$, where x increases by 0.05 in each subsequent composition. A part of the predicted dataset of $FASnI_{3-x}Br_x$ created by four ML models is shown in Table 4.4.

Table 4.4: Predicted dataset for 30 $FASnI_{3-x}Br_x$ compositions formed with different ML Models

Composition	Linear	Random Forest	Gradient Boosting	Voting
$FASnI_3$	1.364998	1.36090	1.393089000	1.361873
$FASnI_{2.95}Br_{0.05}$	1.363859	1.42400	1.477812320	1.443853
$FASnI_{2.9}Br_{0.1}$	1.469473	1.42400	1.477812320	1.479072
$FASnI_{2.85}Br_{0.15}$	1.446150	1.42400	1.477812320	1.471288
$FASnI_{2.8}Br_{0.2}$	1.422873	1.42620	1.477812320	1.463616
$FASnI_{2.75}Br_{0.25}$	1.464716	1.43500	1.477812320	1.491332
$FASnI_{2.7}Br_{0.3}$	1.441477	1.45320	1.486493772	1.490636
$FASnI_{2.65}Br_{0.35}$	3693705	1.45320	1.486493772	1363713
$FASnI_{2.6}Br_{0.4}$	1.523669	1.45320	1.486493772	1.520702
$FASnI_{2.55}Br_{0.45}$	1.500418	1.45320	1.486493772	1.512390
$FASnI_{2.5}Br_{0.5}$	1.477141	1.48300	1.507558271	1.534183
$FASnI_{2.45}Br_{0.55}$	1.582686	1.49930	1.509515629	1.616195
$FASnI_{2.4}Br_{0.6}$	1.559294	1.50150	1.509515629	1.608403
$FASnI_{2.35}Br_{0.65}$	1.472418	1.50150	1.509515629	1.579442
$FASnI_{2.3}Br_{0.7}$	1.578032	1.50150	1.509515629	1.614674
$FASnI_{2.25}Br_{0.75}$	1.554579	1.54010	1.509515629	1.617833
$FASnI_{2.2}Br_{0.8}$	1.531332	1.85826	1.781149751	1.811776
$FASnI_{2.15}Br_{0.85}$	1.636931	1.86736	1.781149751	1.849542
$FASnI_{2.10}Br_{0.9}$	1.613615	1.86736	1.781149751	1.841760
$FASnI_{2.05}Br_{0.95}$	1.60100	1.84266	1.808338906	1.764528
$FASnI_2Br$	1.695837	1.88436	1.781149751	1.869914
$FASnI_{1.95}Br_{1.05}$	1.672591	1.88786	1.781150000	1.832161
$FASnI_{1.9}Br_{1.1}$	1.585493	1.88956	1.781150000	1.781092

$FASnI_{1.85}Br_{1.15}$	1.691069	1.88956	1.781150000	1.816317
$FASnI_{1.8}Br_{1.2}$	1.667815	1.88956	1.781150000	1.808520
$FASnI_{1.75}Br_{1.25}$	1.644621	1.88240	1.776044000	1.783476
$FASnI_{1.7}Br_{1.3}$	1.75012	1.82510	1.788550000	1.801763
$FASnI_{1.65}Br_{1.35}$	1.726774	1.82510	1.788550000	1.793960
$FASnI_{1.6}Br_{1.4}$	1.703558	1.82510	1.788550000	1.786145
$FASnI_{1.55}Br_{1.45}$	1.809164	1.82510	1.788550000	1.821373

We have compared some of the predicted bandgaps with the experimental bandgap value as shown in Table 4.5.

Table 4.5: Comparison between the predicted and experimental bandgap values

Composition	Bandgap by GB Regression	Bandgap Predicted by Voting Regression	Experimental Composition	Experimental Bandgap
$MAPbI_3$	1.60	1.58	$MAPbI_3$	1.55[191]
$MAPb_{0.74}Sn_{0.26}I_3$	1.48	1.52	$MAPb_{0.75}Sn_{0.25}I_3$	1.38 [192]
$MAPb_{0.5}Sn_{0.5}I_3$	1.46	1.48	$MAPb_{0.5}Sn_{0.5}I_3$	1.30 [192]
$MAPb_{0.24}Sn_{0.76}I_3$	1.43	1.44	$MAPb_{0.25}Sn_{0.75}I_3$	1.27 [192]
$MASnI_3$	1.35	1.33	$MASnI_3$	1.37 [192]
$MAGeI_3$	1.78	1.85	$MAGeI_3$	1.90 [90]

It is clear from the Table 4.5 that the predicted value of bandgap is not 100% accurate but it gives a very close value to the experimental values. We have also investigated the prediction accuracy of the known perovskites used in the training dataset and found better performance. It happened due to overfitting in the models. However, overfitting can be solved by adding more data in the training set and reducing the less critical features by analyzing feature importance.

However, there are many promising candidates among the predicted $MAPb_{1-x}Sn_xI_3$ compositions, that can be utilized as an absorber in a solar cell. But based on the performance in previous experimental works, we have chosen the bandgap of composition $MAPb_{0.75}Sn_{0.25}I_3$ for SCAPS analysis [197].

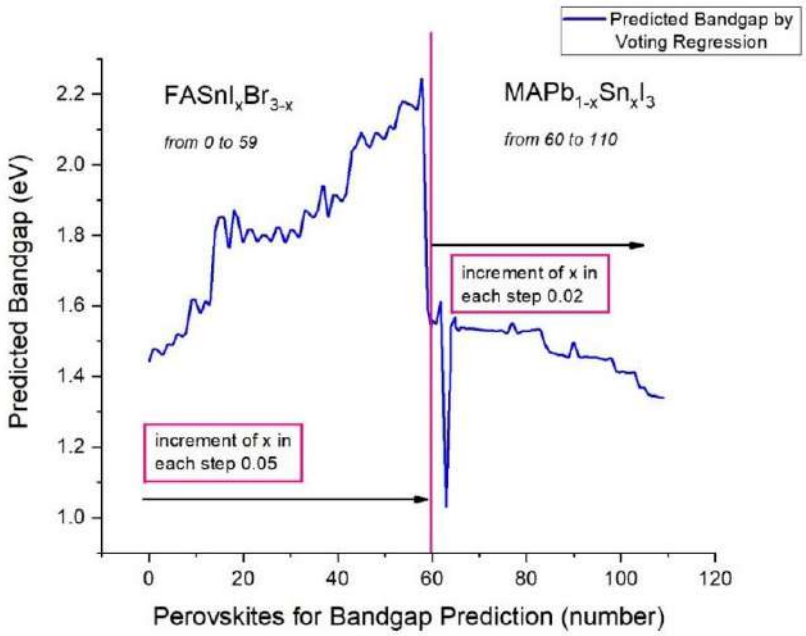


Figure 4.6: Variation of predicted bandgaps with $FASnI_xBr_{3-x}$ and $MAPb_{1-x}Sn_xI_3$ compositions by voting regression

Figure 4.6 shows the final plot of predicted bandgaps for $FASnI_xBr_{3-x}$ and $MAPb_{1-x}Sn_xI_3$, where x denotes the fraction of doping. It is clear that the bandgap is increasing as x is increasing in $FASnI_xBr_{3-x}$ whereas the value of bandgap is decreasing in $MAPb_{1-x}Sn_xI_3$ as x is increasing. There is a sharp peak (at around 1 eV) in the 2nd portion of the graph. This significant reduction in bandgap is due to prediction error in linear regression model leading an error in voting regression also.

4.3 Feature Importance Analysis (SHAP)

To understand the impact of the features in predicting the bandgap, the *feature importance* from scikit-learn has been utilized. The feature importance in Figure 4.7 is evaluated using generalized feature importance in scikit-learn and it shows the importance of the first 20 features in descending order.

From Figure 4.7, it is clear that the electron affinity of the halide materials plays the most crucial role in affecting the target property (bandgap), then the electron negativity of the halide site, and so on. The B and X site mainly affects the bandgap, as seen in Figure 4.7.

However, this tells us the importance of the features on the target variable (bandgap) but does not indicate whether its influence is positive or negative.

To understand the impact of features with a clear picture, we have performed Shapley Additive Analysis.

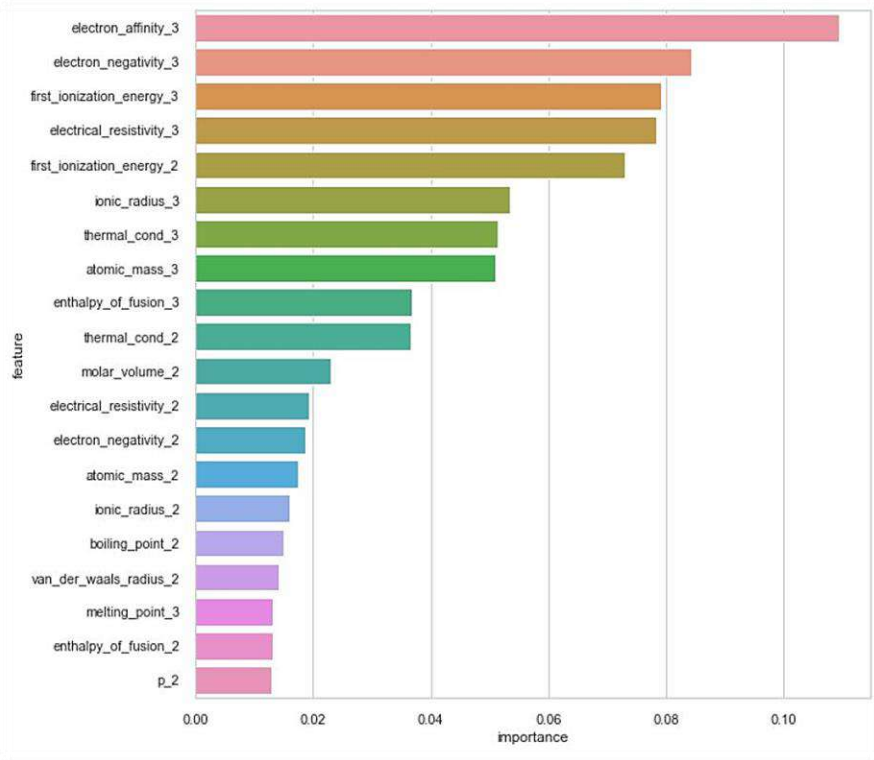


Figure 4.7: Feature importance made with generalized feature importance

The SHAPLey Additive Explanation (SHAP plot) eliminates the necessity of feature importance and partial dependence plots by intuitively representing the impact of each feature value along with their positive or negative effect on the final output values.

SHAP plot in Figure 4.8 clearly visualize the nature of impact of each feature on the target property (bandgap).

The vertical locations show the first 20 features affecting the bandgap with a feature value. Red color in the vertical axis indicates high feature impact and blue color indicates a low impact. Horizontal location of dots represents whether the effect of that feature value caused a higher or lower value of bandgap prediction. A positive shapley (SHAP) value increases the prediction probability by a particular feature and a negative value decreases the probability.

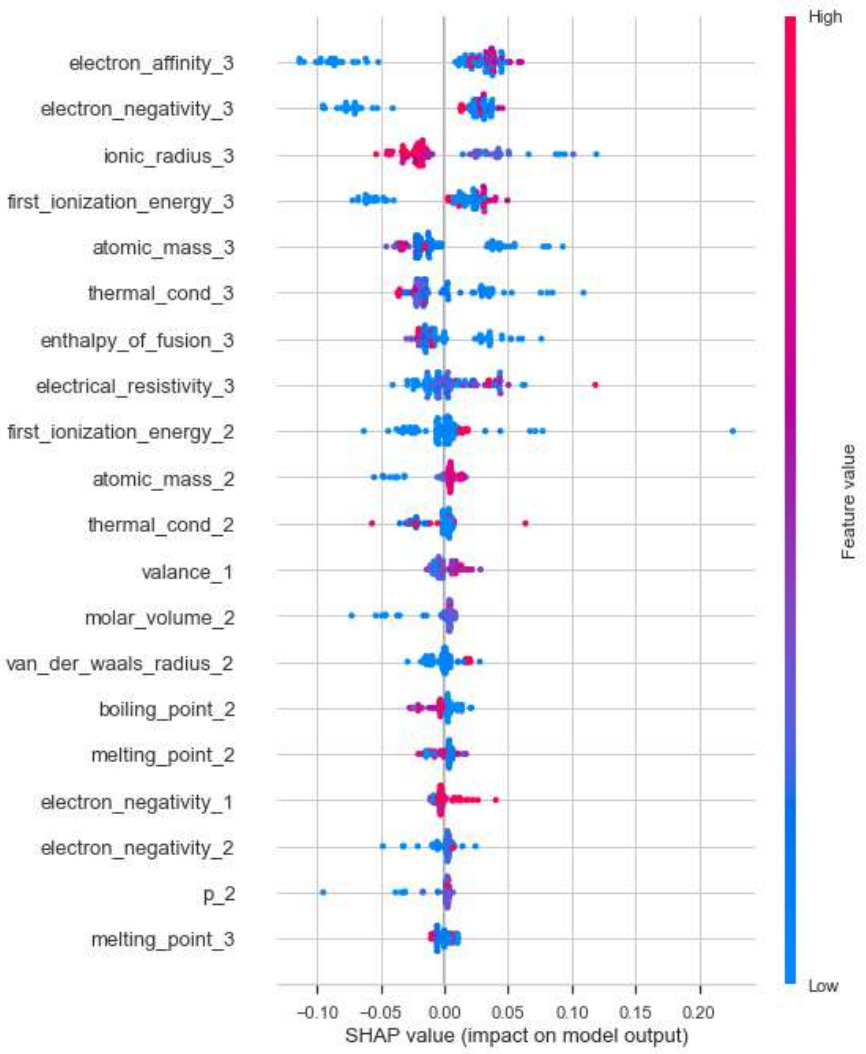


Figure 4.8: SHAP analysis for visualizing the impact of the feature on bandgap value

4.4 SCAPS Simulation

The schematic diagrams of the pristine $MAPbI_3$ perovskite ($CH_3NH_3PbI_3$) and Sn incorporated ($CH_3NH_3Pb_{0.75}Sn_{0.25}I_3$) solar cell absorber simulated in this work are shown in Figure 4.9 and 4.10 respectively. The planar device has an n-i-p configuration where the ETL, perovskite ($MAPbI_3$ or $MAPb_xSn_{3-x}I_3$) and HTL layer serve as the n-type (n), intrinsic (i), and p-type (p) regions, respectively [137].

4.4.1 Simulation of $MAPbI_3$ Absorber Materials

To account for the interface recombination, the perovskite/TiO_2 interface layer (IL_1) and the perovskite/CuSCN interface layer (IL_2) have been considered during device simulations. The input parameters of the structure and the different materials are summarized in Table 4.6.

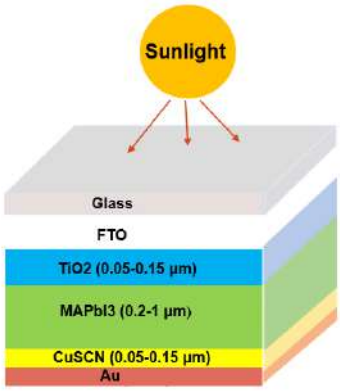


Figure 4.9: Schematic of simulated pristine $MAPbI_3$ absorber in SCAPS

We have tried various ETL and HTL combination but found TiO_2 and CuSCN perfect for $MAPbI_3$ absorber material. Characteristics properties of TiO_2 and CuSCN are taken from previous work [117], [193], [194]. Here fluorine doped tin oxide (FTO) coated on glass as anti-reflective coating is used on top of the architecture.

Table 4.6: The simulation input parameters for the n-i-p pristine $MAPbI_3$ absorber

Parameter	**Term**	**ETL (TiO_2)**	**Perovskite ($MAPbI_3$)**	**HTL (CuSCN)**
d (µm)	*Thickness*	0.05-0.15	0.2-1.0	0.05-0.15
E_g *(eV)*	*Band gap*	3.2	**1.58 (voting)**	3.4
χ *(eV)*	*Electron Affinity*	3.9	3.9	1.9
ε_r	*Relative Permittivity*	9	6.5	10
N_c (cm^{-3})	*Effective density of states at CB*	1×10^{21}	2.2×10^{18}	1.7×10^{19}
N_v (cm^{-3})	*Effective density of states at VB*	2×10^{20}	1.8×10^{19}	2.5×10^{21}
μ_n $(cm^2/V\,s)$	*Mobility of electrons*	20	2	1×10^{-4}
μ_p $(cm^2/V\,s)$	*Mobility of holes*	10	2	1×10^{-4}
N_d (cm^{-3})	*Density of n-type doping*	1×10^{19}	5.21×10^{9}	0.0
N_a (cm^{-3})	*Density of p-type doping*	1.0	5.21×10^{9}	1×10^{18}
N_t (cm^{-3})	*Density of defects*	1×10^{15} neutral	2.5×10^{13}	1×10^{14}
References		[117], [193]	[117], [193]	[117], [194]

However, we have utilized the bandgap value of $MAPbI_3$ obtained from voting regression ML model. Other characteristics of $MAPbI_3$ except layer thickness are collected from previous work [117], [193].

4.4.2 Simulation of $MAPb_{0.75}Sn_{0.25}I_3$ Absorber Materials

Several combinations of ETL and HTL are tried also for this simulation study but found better result with TiO_2 and NiO. Schematic of this device simulation is presented in Figure 4.10. Characteristics property of ETL (TiO_2) and HTL (NiO) are consolidated in Table 4.7.

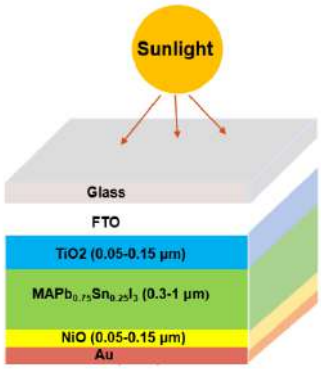


Figure 4.10: Schematic of simulated pristine $MAPb_{0.75}Sn_{0.25}I_3$ absorber in SCAPS

Table 4.7: The simulation input parameters for the n-i-p $MAPb_{0.75}Sn_{0.25}I_3$ absorber

Parameter	**Term**	**ETL (TiO_2)** [117], [193]	**Perovskite ($MAPb_{0.75}Sn_{0.25}I_3$)**	**HTL (NiO)**[195]
d (µm)	*Thickness*	0.05-0.15	0.3-1	0.05-0.15
E_g *(eV)*	*Band gap*	3.2	**1.53 (voting ML)**	3.6
χ *(eV)*	*Electron Affinity*	3.9	3.7-3.9 [192]	1.46
ε_r	*Relative Permittivity*	9	6.5	11
N_c *(cm−3)*	*Effective density of states at CB*	1×10^{21}	2.2×10^{18}	1.6×10^{19}
N_v *(cm−3)*	*Effective density of states at VB*	2×10^{20}	1.8×10^{19}	1.1×10^{19}
μ_n *($cm^2/V\,s$)*	*Mobility of electrons*	20	2	50
μ_p *($cm^2/V\,s$)*	*Mobility of holes*	10	2	50
N_d *(cm^{-3})*	*Density of n-type doping*	1×10^{19}	5.21×10^{9}	0.0
N_a *(cm^{-3})*	*Density of p-type doping*	1.0	5.21×10^{9}	1.8×10^{18}
N_t *(cm^{-3})*	*Density of defects*	1×10^{15} neutral	2.5×10^{13}	1×10^{14}

To account for the interface recombination, the perovskite/TiO_2 interface layer (IL_1) and the perovskite/NiO interface layer (IL_2) have been considered during device simulations.

The absorption co-efficient vs. wavelength plot of $MAPbI_3$, $MAPb_{0.75}Sn_{0.25}I_3$, TiO_2, CuSCN, and NiO were collected from various published papers [193], [196]–[199]. FTO (work function 4.4) and Au (work function 5.1) are used as front metal contact and back metal contact in both of the simulation [135], [136].

4.5 SCAPS Output

$MAPbI_3$ and $MAPb_{0.75}Sn_{0.25}I_3$ have been simulated with the variation of their layer thickness. Though other parameters such as electron affinity, hole mobility, defect concentrations of absorbers are also important in affecting the overall performance of solar cell, those parameters are not studied extensively. It is important to reiterate the main focus of this thesis work here. The main objective of this thesis work was to develop a machine learning model for predicting the most critical parameter, bandgap of halide absorber perovskite material and to analyse the solar cell performance based on the bandgap.

ETL and HTL parameters are also kept constant because we are not concerned about their impact in this work. But it is very important to select a suitable ETL and HTL analysing the band alignment with the absorber. For efficient electron extraction by the ETL, the electron affinity of the ETL must be greater than the absorber material. Similarly, for efficient hole extraction by the HTL, the electron affinity of perovskite should be higher than the HTL layer. The ETL should have a high bandgap energy to facilitate considerable portion of the electromagnetic spectrum to flow through and reach the perovskite (absorber) layer. However, these are the basic requirements for being compatible in a solar cell device but are not sufficient to achieve highest efficiency.

After analysing different combination of ETL and HTL with $MAPbI_3$, we found TiO_2 and CuSCN as the most suitable electron transport layer and hole transport layer, respectively. For the partially Pb replaced $MAPb_{0.75}Sn_{0.25}I_3$ absorber, TiO_2 and NiO performed better.

4.5.1 Effect of Absorber Layer Thickness on J-V Curve:

It is crucial to estimate the device performance with different thicknesses of the absorber layer. In this subsection, the analysis of the perovskite devices has been made, where the influence of $MAPbI_3$ and $MAPb_{0.75}Sn_{0.25}I_3$ thicknesses on the solar parameters is evaluated. In most cases, high bulk layer thickness of absorbers leads to the high solar energy absorption. Hence, increasing thickness might be a lucrative option but most perovskite absorbers become susceptible to intrinsic and extrinsic defects after reaching a certain thickness. The variation of thickness for $MAPbI_3$ has been performed from 0.02 µm to 1 µm in five equal steps in a linear scale, keeping the remaining device parameters intact. Corresponding impact for $MAPbI_3$ layer thickness in open circuit voltage, current density, fill factor and PCE are summarized in Figure 4.11 (a), (b), (c), and (d), respectively.

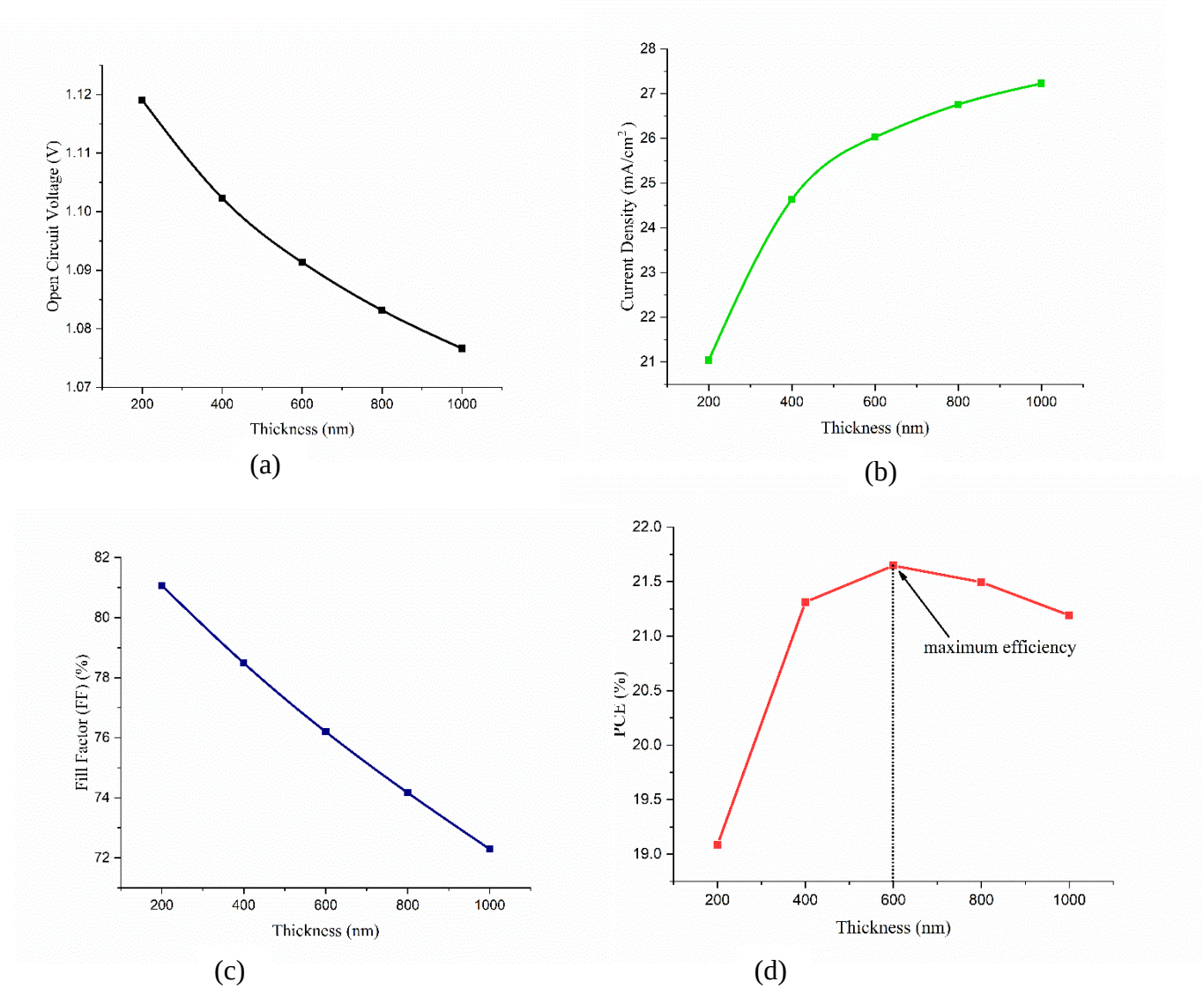


Figure 4.11: Effect of $MAPbI_3$ layer thickness on (a) Open Circuit Voltage (b) Current Density (c) Fill Factor (d) PCE

As it can be seen from Figure 4.11 (d), the peak PCE of 21.645% is found at 600 nm perovskite layer thickness, provided that there is no significant defect in the bulk absorber layer. Hence, the optimum layer thickness for $MAPbI_3$ absorber is 600 nm.

Figure 4.11 (a) and (c) indicate that open circuit voltage and fill factor is reducing with increasing absorber layer thickness due to the increasing proportion defect sites within the film.

In case of short circuit current density, more absorber thickness generates more electron−hole pairs and increases J_{sc} resulting in a higher photocurrent [30]. The evidence of this enhancement can also be seen from the J-V curves. Figure 4.12 shows the J-V curves in a single figure as a function of $MAPbI_3$ absorber layer thickness.

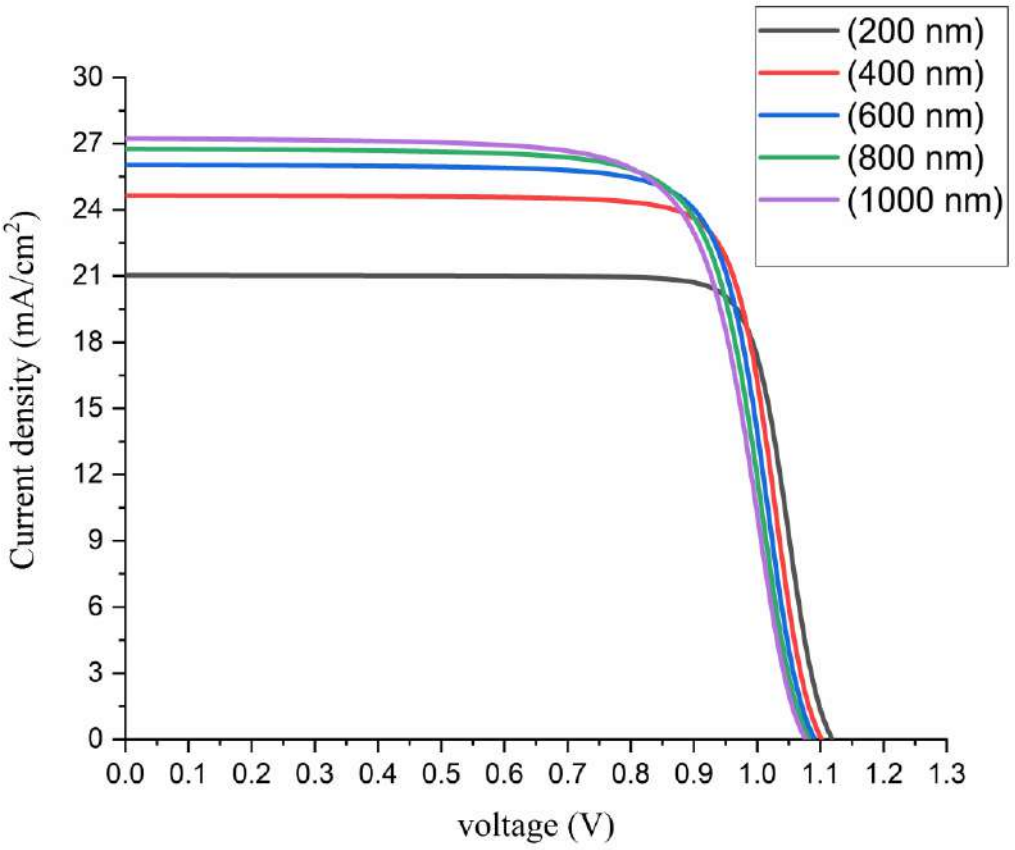


Figure 4.12: J-V curve as a function of absorber layer $MAPbI_3$ thickness

It is clear from the above figure that short circuit current density is increasing and open circuit voltage is decreasing with increase in layer thickness of $MAPbI_3$. J-V curve with important parameters for 600 nm layer thickness is shown in Figure 4.13.

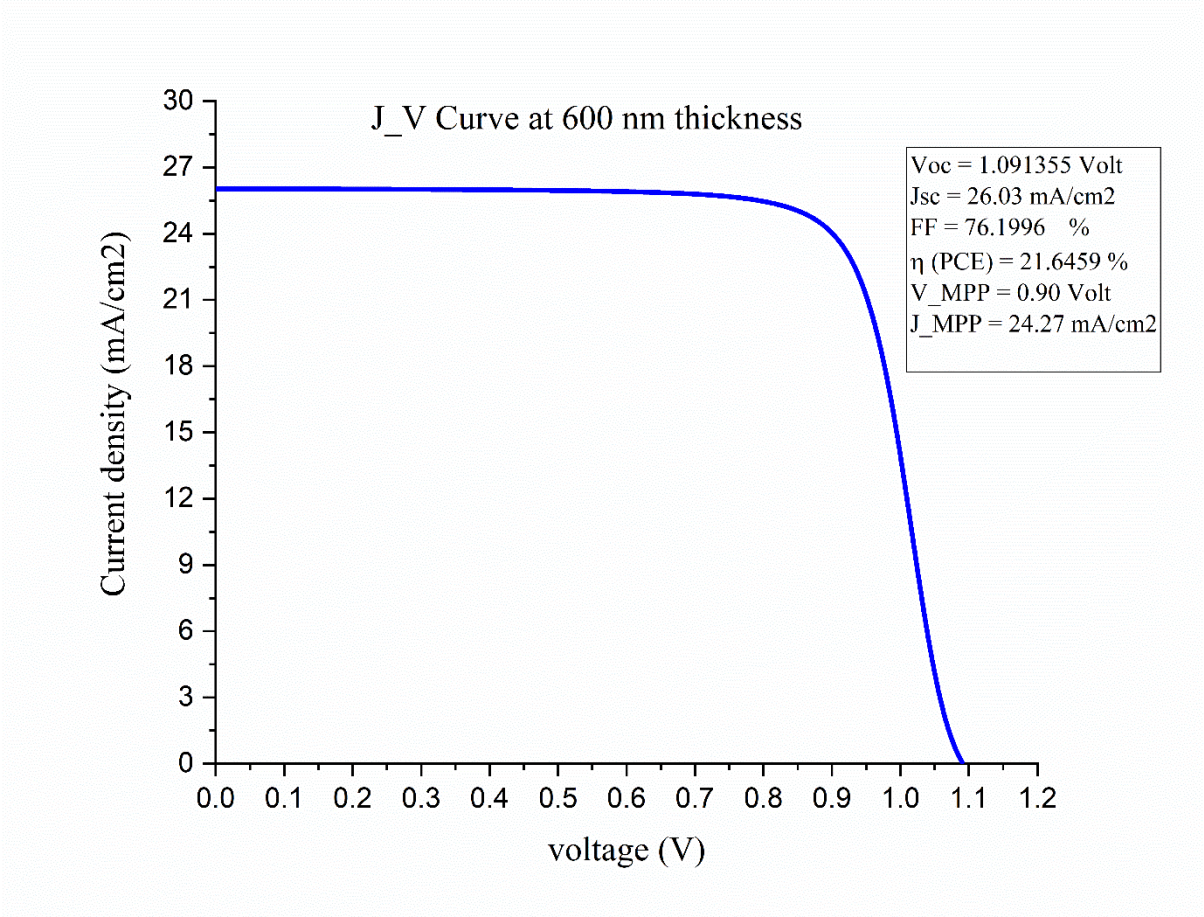


Figure 4.13: Current Density vs. Open Circuit Voltage diagram for $MAPbI_3$ at 600 nm thickness

However, the effect of absorber layer thickness for $MAPb_{0.75}Sn_{0.25}I_3$ shows a unique and exciting impact on solar cell parameters. The absorber layer thickness varies from 0.3 µm (300 nm) to 1 µm (1000 nm) in eight equal steps. ETL (TiO_2) and HTL (NiO) layer thickness were kept at 150 nm. The layer thickness of ETL and HTL was modulated, but we could not find better results with other thickness values.

Figure 4.14 to Figure 4.18 show the effect of perovskite layer thickness on various properties of solar cell. For the single junction solar cell with $MAPb_{0.75}Sn_{0.25}I_3$ absorber, open circuit

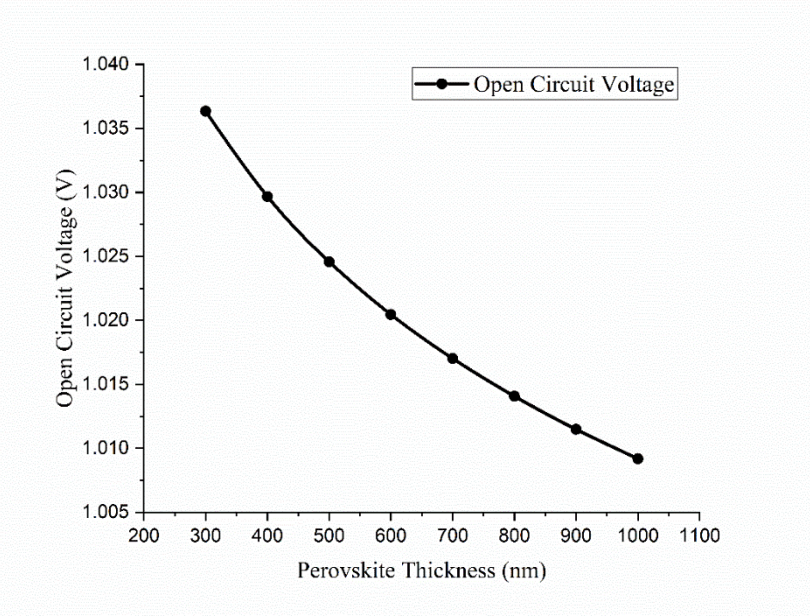


Figure 4.14: Effect of the variation of $MAPb_{0.75}Sn_{0.25}I_3$ layer thickness on Open Circuit Voltage

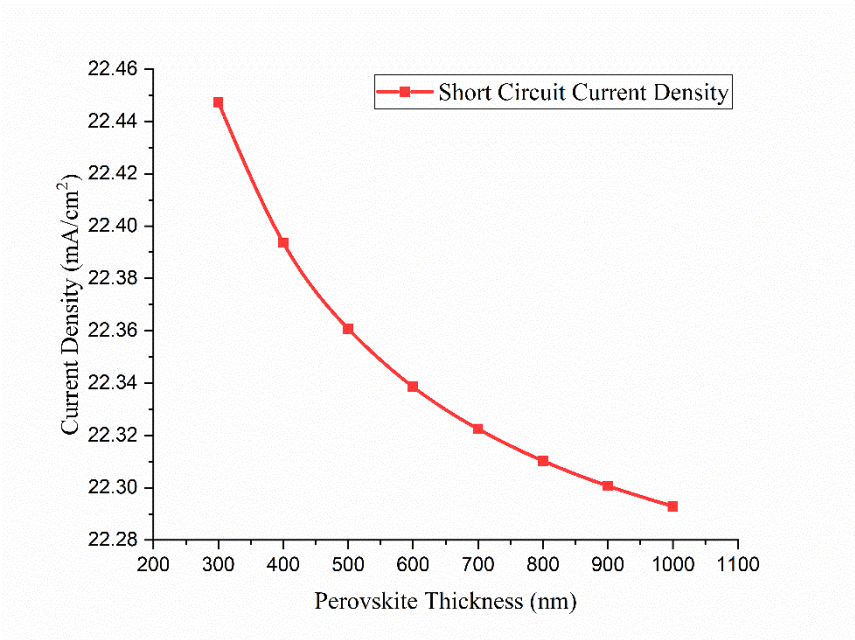


Figure 4.15: Effect of the variation of $MAPb_{0.75}Sn_{0.25}I_3$ layer thickness on Short Circuit Current Density

voltage is also reduced with increased absorber layer thickness, and the maximum open circuit voltage attainable for this solar cell is also lower than the solar cell with $MAPbI_3$ absorber layer. Figure 4.15 shows that short circuit current density of the solar cell is also

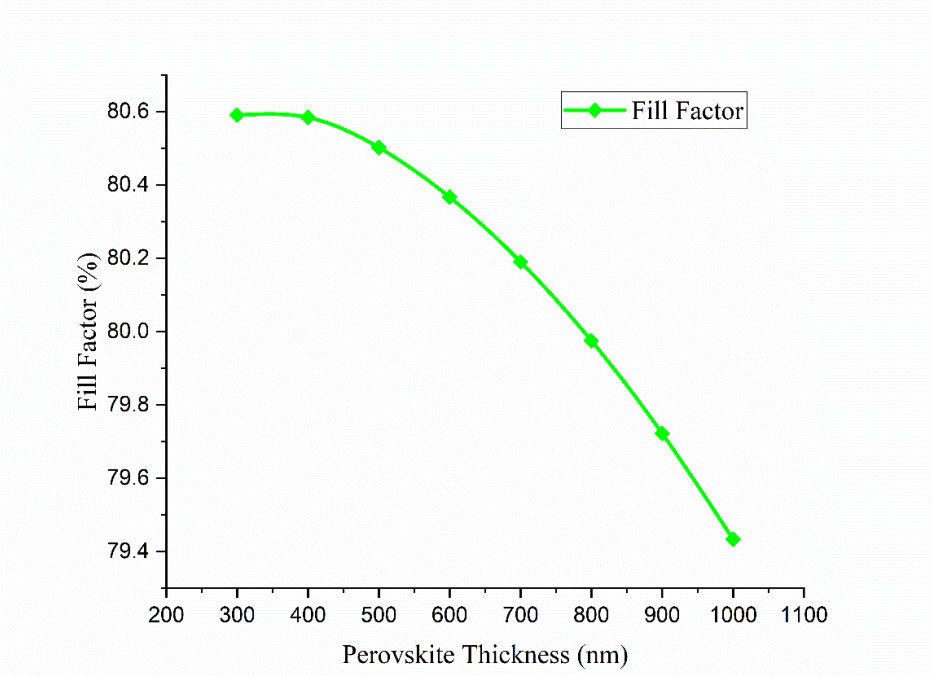


Figure 4.16: Effect of $MAPb_{0.75}Sn_{0.25}I_3$ layer thickness on Fill Factor

decreasing with increase in $MAPb_{0.75}Sn_{0.25}I_3$ absorber layer thickness which is contrary to the solar cell behavior with pristine $MAPbI_3$ absorber. It is clear from Figure 4.16 that fill factor is also decreasing with increase in absorber thickness. Since, open circuit voltage, short circuit current density, and fill factor decrease with increasing absorber layer thickness,

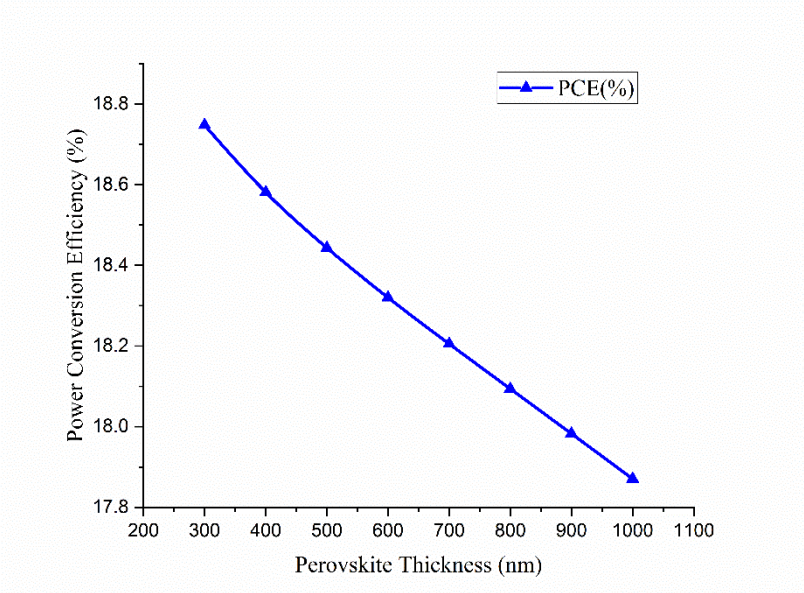


Figure 4.17: Effect of $MAPb_{0.75}Sn_{0.25}I_3$ layer thickness on PCE

overall power conversion efficiency (PCE) should also decrease with increasing absorber layer thickness which is evident from Figure 4.17.

The highest efficiency of 18.58% is found for 300 nm absorber layer thickness, which decreased for increasing absorber layer thickness. Hence, the optimum $MAPb_{0.75}Sn_{0.25}I_3$ absorber layer thickness is 300 nm when the ETL and HTL used in the solar cell is TiO_2 and NiO, respectively.

Tin based perovskites has the tendency to be oxidized to Sn^{4+} with or without the presence of oxygen which also causes a lower efficiency [200]. However, an excessive amount of Sn^{4+} within the absorber layer could increase the defects of tin-based perovskite, which could degrade the overall performance of perovskite solar cells [83], [201]. All of the previous four figures (Figure 4.14 to Figure 4.17) are superimposed on a single figure (Figure 4.18) for visualizing the net effects of the variation of absorber layer thickness.

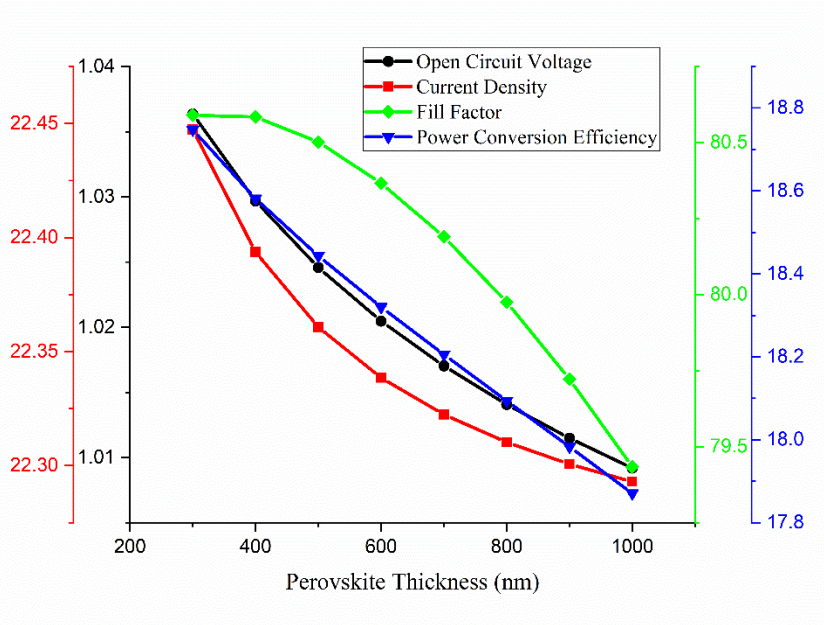


Figure 4.18: Effect of J-V parameters with the variation of $MAPb_{0.75}Sn_{0.25}I_3$ layer thickness

Higher defects within the absorber layer might attribute to the trend of lowering all the values (V_{oc}, J_{sc}, FF and PCE) when the thickness of the absorber layer increases above 300 nm.

The biggest part of the computational work is to compare the results with experimental work.

Liu *et al*. demonstrated that $MAPb_{0.75}Sn_{0.25}I_3$ perovskite solar cells with an inverted structure can obtain maximum power conversion efficiency (PCE) of 14.12% which is lower than the obtained result.

SCAPS analysis will give more reliable result if the bandgap prediction accuracy is increased. However, it can be asserted here that the partially Pb replaced composition of $MAPb_{0.75}Sn_{0.25}I_3$ should gain more attention from the scientific community as the performance found in this study is quite impressive. If the instability issue can be solved either by using encapsulating process or any other methods then it might show excellent solar cell property.

CHAPTER 5: CONCLUSION

Using the elemental properties of various types of perovskites and their experimental bandgaps, four different ML models have been developed to predict the bandgap of unknown halide perovskites. After developing the models their performances have been analyzed using various error metrics and cross-validation score and a suitable model is selected for further analysis. Finally, utilizing the bandgap obtained from the selected model, the solar cell performance and potentiality of two different composition of perovskites are analyzed in SCAPS.

The key findings of this thesis work have been summarized below:

1. Among the four ML models developed, voting regression model has been selected for predicting the bandgap with highest R^2 value of 0.95 and lowest RMSE of 0.076.

2. The prediction accuracy of voting regression is excellent for known dataset. For unknown data, it also predicted well, however with minor anomalies for a few compositions in the prediction set.

3. To reduce the toxicity of lead in $MAPbI_3$, three potential replacements of Ge, Sn, and Si were considered for bandgap prediction and Sn based $MAPb_{1-x}Sn_xI_3$ was found the best.

4. From voting regression ML model, predicted bandgap of pristine $MAPbI_3$ (1.58 eV) and $MAPb_{0.75}Sn_{0.25}I_3$ (1.53 eV) are selected for SCAPS Analysis.

5. From SCAPS analysis, it is found that power conversion efficiency (PCE) of pristine $MAPbI_3$ is 21.65% but the partially Pb replaced composition $MAPb_{0.75}Sn_{0.25}I_3$ has somewhat lower PCE of 18.58% which is higher than previously published result.

6. The highest PCE of pristine $MAPbI_3$ was found for 600 nm absorber layer thickness with whereas the maximum PCE achieved for 300 nm of $MAPb_{0.75}Sn_{0.25}I_3$ absorber layer thickness along with 150 nm ETL and 150 nm HTL thickness.

7. Although there are some anomalies in the predicted bandgap value, this analysis would serve as a commendable gateway toward further investigations in solar cell research.

CHAPTER 6: RECOMMENDATIONS FOR FUTURE WORK

- The contributions of this research work can be utilized to find a better ML model for predicting the bandgap with higher accuracy. The primary drawback of this research was that the dataset was relatively small and limited as the number of the experimentally determined bandgap of perovskites is limited. Dataset can be increased further to get a better prediction accuracy.

- The stability of perovskite solar cells in service condition i.e., the longest lifetime reported for PSCs is about one year [202], which is much shorter than 25 years as expected from commercialized PV technologies. It is thus clear that the short lifetime is another obstacle hindering the commercialization of PSC [203]. There are various methods to determine the formability of perovskites but these are not sufficient. The knowledge, gained from building the machine learning models, can be utilized to predict the formability and stability of PSC in service condition.

BIBLIOGRAPHY

[1] "IEA (2022), Solar PV, IEA, Paris https://www.iea.org/reports/solar-pv, License: CC BY 4.0." Accessed: May 13, 2023 [Online].

[2] F. F. Targhi, Y. S. Jalili, and F. Kanjouri, "$MAPbI_3$ and $FAPbI_3$ perovskites as solar cells: Case study on structural, electrical and optical properties," *Results Phys*, vol. 10, pp. 616–627, Sep. 2018, doi: 10.1016/j.rinp.2018.07.007.

[3] N.-G. Park, "Perovskite solar cells: an emerging photovoltaic technology," *Materials Today*, vol. 18, no. 2, pp. 65–72, Mar. 2015, doi: 10.1016/j.mattod.2014.07.007.

[4] H. J. Snaith, "Perovskites: The Emergence of a New Era for Low-Cost, High-Efficiency Solar Cells," *J Phys Chem Lett*, vol. 4, no. 21, pp. 3623–3630, Nov. 2013, doi: 10.1021/jz4020162.

[5] W. S. Yang *et al.*, "Iodide management in formamidinium-lead-halide–based perovskite layers for efficient solar cells," *Science (1979)*, vol. 356, no. 6345, pp. 1376–1379, Jun. 2017, doi: 10.1126/science.aan2301.

[6] V. Sharma *et al.*, "Rational design of all organic polymer dielectrics," *Nat Commun*, vol. 5, no. 1, p. 4845, Sep. 2014, doi: 10.1038/ncomms5845.

[7] L. Meng, J. You, and Y. Yang, "Addressing the stability issue of perovskite solar cells for commercial applications," *Nat Commun*, vol. 9, no. 1, p. 5265, Dec. 2018, doi: 10.1038/s41467-018-07255-1.

[8] J. Li, B. Pradhan, S. Gaur, and J. Thomas, "Predictions and Strategies Learned from Machine Learning to Develop High-Performing Perovskite Solar Cells," *Adv Energy Mater*, vol. 9, no. 46, p. 1901891, Dec. 2019, doi: 10.1002/aenm.201901891.

[9] N. J. Jeon, J. H. Noh, Y. C. Kim, W. S. Yang, S. Ryu, and S. Il Seok, "Solvent engineering for high-performance inorganic–organic hybrid perovskite solar cells," *Nat Mater*, vol. 13, no. 9, pp. 897–903, Sep. 2014, doi: 10.1038/nmat4014.

[10] M. M. Lee, J. Teuscher, T. Miyasaka, T. N. Murakami, and H. J. Snaith, "Efficient Hybrid Solar Cells Based on Meso-Superstructured Organometal Halide Perovskites,"

Science (1979), vol. 338, no. 6107, pp. 643–647, Nov. 2012, doi: 10.1126/science.1228604.

[11] M. A. Green, A. Ho-Baillie, and H. J. Snaith, "The emergence of perovskite solar cells," *Nature Photonics*, vol. 8, no. 7. Nature Publishing Group, pp. 506–514, 2014. doi: 10.1038/nphoton.2014.134.

[12] A. Kojima, K. Teshima, Y. Shirai, and T. Miyasaka, "Organometal halide perovskites as visible-light sensitizers for photovoltaic cells," *J Am Chem Soc*, vol. 131, no. 17, pp. 6050–6051, May 2009, doi: 10.1021/ja809598r.

[13] J. Jeong *et al.*, "Pseudo-halide anion engineering for α-$FAPbI_3$ perovskite solar cells," *Nature*, vol. 592, no. 7854, 2021, doi: 10.1038/s41586-021-03406-5.

[14] W. Ke, C. C. Stoumpos, and M. G. Kanatzidis, "'Unleaded' Perovskites: Status Quo and Future Prospects of Tin-Based Perovskite Solar Cells," *Advanced Materials*, vol. 31, no. 47, p. 1803230, Nov. 2019, doi: 10.1002/adma.201803230.

[15] A. Babayigit, A. Ethirajan, M. Muller, and B. Conings, "Toxicity of organometal halide perovskite solar cells," *Nature Materials*, vol. 15, no. 3. 2016. doi: 10.1038/nmat4572.

[16] A. Bibi *et al.*, "Lead-free halide double perovskites: Toward stable and sustainable optoelectronic devices," *Materials Today*, vol. 49. 2021. doi: 10.1016/j.mattod.2020.11.026.

[17] R. Wang, J. Wang, S. Tan, Y. Duan, Z. K. Wang, and Y. Yang, "Opportunities and Challenges of Lead-Free Perovskite Optoelectronic Devices," *Trends in Chemistry*, vol. 1, no. 4. 2019. doi: 10.1016/j.trechm.2019.04.004.

[18] F. Hao, C. C. Stoumpos, D. H. Cao, R. P. H. Chang, and M. G. Kanatzidis, "Lead-free solid-state organic–inorganic halide perovskite solar cells," *Nat Photonics*, vol. 8, no. 6, pp. 489–494, Jun. 2014, doi: 10.1038/nphoton.2014.82.

[19] W. Ke and M. G. Kanatzidis, "Prospects for low-toxicity lead-free perovskite solar cells," *Nat Commun*, vol. 10, no. 1, p. 965, Feb. 2019, doi: 10.1038/s41467-019-08918-3.

[20] J. Li, J. Duan, X. Yang, Y. Duan, P. Yang, and Q. Tang, “Review on recent progress of lead-free halide perovskites in optoelectronic applications,” *Nano Energy*, vol. 80, p. 105526, Feb. 2021, doi: 10.1016/j.nanoen.2020.105526.

[21] S. S. Mali, C. S. Shim, and C. K. Hong, “Highly stable and efficient solid-state solar cells based on methylammonium lead bromide ($CH_3NH_3PbBr_3$) perovskite quantum dots,” *NPG Asia Mater*, vol. 7, no. 8, pp. e208–e208, Aug. 2015, doi: 10.1038/am.2015.86.

[22] A. Sharenko and M. F. Toney, “Relationships between Lead Halide Perovskite Thin-Film Fabrication, Morphology, and Performance in Solar Cells,” *J Am Chem Soc*, vol. 138, no. 2, pp. 463–470, Jan. 2016, doi: 10.1021/jacs.5b10723.

[23] M. L. Petrus *et al.*, “New Generation Hole Transporting Materials for Perovskite Solar Cells: Amide-Based Small-Molecules with Nonconjugated Backbones,” *Adv Energy Mater*, vol. 8, no. 32, p. 1801605, Nov. 2018, doi: 10.1002/aenm.201801605.

[24] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” *Science (1979)*, vol. 349, no. 6245, pp. 255–260, Jul. 2015, doi: 10.1126/science.aaa8415.

[25] Y. Liu, W. Yan, H. Zhu, Y. Tu, L. Guan, and X. Tan, “Study on bandgap predications of ABX3-type perovskites by machine learning,” *Org Electron*, vol. 101, p. 106426, Feb. 2022, doi: 10.1016/j.orgel.2021.106426.

[26] X. Gao *et al.*, “Stable and High-Efficiency Methylammonium-Free Perovskite Solar Cells,” *Advanced Materials*, vol. 32, no. 9, p. 1905502, Mar. 2020, doi: 10.1002/adma.201905502.

[27] Z. H. Bakr, Q. Wali, A. Fakharuddin, L. Schmidt-Mende, T. M. Brown, and R. Jose, “Advances in hole transport materials engineering for stable and efficient perovskite solar cells,” *Nano Energy*, vol. 34, pp. 271–305, Apr. 2017, doi: 10.1016/j.nanoen.2017.02.025.

[28] N. J. Jeon *et al.*, “A fluorene-terminated hole-transporting material for highly efficient and stable perovskite solar cells,” *Nat Energy*, vol. 3, no. 8, pp. 682–689, Jul. 2018, doi: 10.1038/s41560-018-0200-6.

[29] K.-G. Lim, S. Ahn, Y.-H. Kim, Y. Qi, and T.-W. Lee, "Universal energy level tailoring of self-organized hole extraction layers in organic solar cells and organic–inorganic hybrid perovskite solar cells," *Energy Environ Sci*, vol. 9, no. 3, pp. 932–939, 2016, doi: 10.1039/C5EE03560K.

[30] Md. S. Islam *et al.*, "Machine Learning Approach to Delineate the Impact of Material Properties on Solar Cell Device Physics," *ACS Omega*, vol. 7, no. 26, pp. 22263–22278, Jul. 2022, doi: 10.1021/acsomega.2c01076.

[31] J. Alazraque-Cherni, "Renewable Energy for Rural Sustainability in Developing Countries," *Bull Sci Technol Soc*, vol. 28, no. 2, pp. 105–114, Apr. 2008, doi: 10.1177/0270467607313956.

[32] IEA, "Share of cumulative power capacity by technology, 2010-2027," *2022*. https://www.iea.org/data-and-statistics/charts/share-of-cumulative-power-capacity-by-technology-2010-2027 (accessed May 13, 2023).

[33] S. Asumadu-Sarkodie and P. A. Owusu, "Carbon dioxide emissions, GDP, energy use, and population growth: a multivariate and causality analysis for Ghana, 1971–2013," *Environmental Science and Pollution Research*, vol. 23, no. 13, pp. 13508–13520, Jul. 2016, doi: 10.1007/s11356-016-6511-x.

[34] A. Ajanovic, "Biofuels versus food production: Does biofuels production increase food prices?," *Energy*, vol. 36, no. 4, pp. 2070–2076, Apr. 2011, doi: 10.1016/j.energy.2010.05.019.

[35] R. Baños, F. Manzano-Agugliaro, F. G. Montoya, C. Gil, A. Alcayde, and J. Gómez, "Optimization methods applied to renewable and sustainable energy: A review," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 4, pp. 1753–1766, May 2011, doi: 10.1016/j.rser.2010.12.008.

[36] J. Timperley, "Why fossil fuel subsidies are so hard to kill," *Nature*, vol. 598, no. 7881, pp. 403–405, Oct. 2021, doi: 10.1038/d41586-021-02847-2.

[37] M. Z. Jacobson *et al.*, "100% clean and renewable wind, water, and sunlight (WWS) all-sector energy roadmaps for the 50 United States," *Energy Environ Sci*, vol. 8, no. 7, pp. 2093–2117, 2015, doi: 10.1039/C5EE01283J.

[38] N. Scovronick *et al.*, “The impact of human health co-benefits on evaluations of global climate policy,” *Nat Commun*, vol. 10, no. 1, p. 2095, May 2019, doi: 10.1038/s41467-019-09499-x.

[39] G. Luderer *et al.*, “Environmental co-benefits and adverse side-effects of alternative power sector decarbonization strategies,” *Nat Commun*, vol. 10, no. 1, p. 5229, Nov. 2019, doi: 10.1038/s41467-019-13067-8.

[40] T. Laing, “Solar power challenges,” *Nat Sustain*, vol. 5, no. 4, pp. 285–286, Jan. 2022, doi: 10.1038/s41893-021-00845-w.

[41] M. Iqbal, “THE SOLAR CONSTANT AND ITS SPECTRAL DISTRIBUTION,” in *An Introduction to Solar Radiation*, Elsevier, 1983, pp. 43–58. doi: 10.1016/B978-0-12-373750-2.50008-2.

[42] U. W. Peter Würfel, *Physics of Solar Cells: Basic Principles to Advanced Concepts*. 2016.

[43] “https://www.solarquotes.com.au/blog/uv-solar-panels/.” Accessed: May 13, 2023. [Online]. Available: https://www.solarquotes.com.au/blog/uv-solar-panels/

[44] Tetsuo Soga, “Fundamentals of Solar Cell,” in *Nanostructured Materials for Solar Energy Conversion*, First Ed.ELSEVIER, 2006.

[45] “energyeducation.ca.” https://energyeducation.ca/encyclopedia/Photovoltaic_cell (accessed May 13, 2023).

[46] “pveducation.org.” https://www.pveducation.org/pvcdrom/solar-cell-operation/iv-curve (accessed May 13, 2023).

[47] A. Augusto, S. Y. Herasimenka, R. R. King, S. G. Bowden, and C. Honsberg, “Analysis of the recombination mechanisms of a silicon solar cell with low bandgap-voltage offset,” *J Appl Phys*, vol. 121, no. 20, p. 205704, May 2017, doi: 10.1063/1.4984071.

[48] W. Arpavate, K. Roongraung, and S. Chuangchote, “Photochemical solid-state reactions,” in *Green Sustainable Process for Chemical and Environmental Engineering and Science*, Elsevier, 2021, pp. 189–203. doi: 10.1016/B978-0-12-819720-2.00012-6.

[49] F. Igbari, Z. Wang, and L. Liao, “Progress of Lead-Free Halide Double Perovskites,” *Adv Energy Mater*, vol. 9, no. 12, p. 1803150, Mar. 2019, doi: 10.1002/aenm.201803150.

[50] M. Srivastava, J. M. Howard, T. Gong, M. Rebello Sousa Dias, and M. S. Leite, “Machine Learning Roadmap for Perovskite Photovoltaics,” *J Phys Chem Lett*, vol. 12, no. 32, pp. 7866–7877, Aug. 2021, doi: 10.1021/acs.jpclett.1c01961.

[51] C. Li, X. Lu, W. Ding, L. Feng, Y. Gao, and Z. Guo, “Formability of ABX_3 (X = F, Cl, Br, I) halide perovskites,” *Acta Crystallogr B*, vol. 64, no. 6, pp. 702–707, Dec. 2008, doi: 10.1107/S0108768108032734.

[52] A. Chilvery, S. Das, P. Guggilla, C. Brantley, and A. Sunda-Meya, “A perspective on the recent progress in solution-processed methods for highly efficient perovskite solar cells,” *Sci Technol Adv Mater*, vol. 17, no. 1, pp. 650–658, Jan. 2016, doi: 10.1080/14686996.2016.1226120.

[53] W. Ke, C. C. Stoumpos, and M. G. Kanatzidis, “‘Unleaded’ Perovskites: Status Quo and Future Prospects of Tin-Based Perovskite Solar Cells,” *Advanced Materials*, vol. 31, no. 47, p. 1803230, Nov. 2019, doi: 10.1002/adma.201803230.

[54] H. Tsai *et al.*, “High-efficiency two-dimensional Ruddlesden–Popper perovskite solar cells,” *Nature*, vol. 536, no. 7616, pp. 312–316, Aug. 2016, doi: 10.1038/nature18306.

[55] M. Grätzel, “The light and shade of perovskite solar cells,” *Nat Mater*, vol. 13, no. 9, pp. 838–842, Sep. 2014, doi: 10.1038/nmat4065.

[56] M.-G. Ju *et al.*, “Toward Eco-friendly and Stable Perovskite Materials for Photovoltaics,” *Joule*, vol. 2, no. 7, pp. 1231–1241, Jul. 2018, doi: 10.1016/j.joule.2018.04.026.

[57] Y. Zong *et al.*, “Continuous Grain-Boundary Functionalization for High-Efficiency Perovskite Solar Cells with Exceptional Stability,” *Chem*, vol. 4, no. 6, pp. 1404–1415, Jun. 2018, doi: 10.1016/j.chempr.2018.03.005.

[58] M. Grätzel, “The light and shade of perovskite solar cells,” *Nat Mater*, vol. 13, no. 9, pp. 838–842, Sep. 2014, doi: 10.1038/nmat4065.

[59] H. S. Jung and N.-G. Park, “Perovskite Solar Cells: From Materials to Devices,” *Small*, vol. 11, no. 1, pp. 10–25, Jan. 2015, doi: 10.1002/smll.201402767.

[60] Q. Chen *et al.*, “Under the spotlight: The organic–inorganic hybrid halide perovskite for optoelectronic applications,” *Nano Today*, vol. 10, no. 3, pp. 355–396, Jun. 2015, doi: 10.1016/j.nantod.2015.04.009.

[61] M. Petrović, V. Chellappan, and S. Ramakrishna, “Perovskites: Solar cells & engineering applications – materials and device developments,” *Solar Energy*, vol. 122, pp. 678–699, Dec. 2015, doi: 10.1016/j.solener.2015.09.041.

[62] M. Shahbazi and H. Wang, “Progress in research on the stability of organometal perovskite solar cells,” *Solar Energy*, vol. 123, pp. 74–87, Jan. 2016, doi: 10.1016/j.solener.2015.11.008.

[63] F. Hao, C. C. Stoumpos, D. H. Cao, R. P. H. Chang, and M. G. Kanatzidis, “Lead-free solid-state organic–inorganic halide perovskite solar cells,” *Nat Photonics*, vol. 8, no. 6, pp. 489–494, Jun. 2014, doi: 10.1038/nphoton.2014.82.

[64] S. D. Stranks and H. J. Snaith, “Metal-halide perovskites for photovoltaic and light-emitting devices,” *Nat Nanotechnol*, vol. 10, no. 5, pp. 391–402, May 2015, doi: 10.1038/nnano.2015.90.

[65] K. Du, W. Meng, X. Wang, Y. Yan, and D. B. Mitzi, “Bandgap Engineering of Lead-Free Double Perovskite $Cs_2AgBiBr_6$ through Trivalent Metal Alloying,” *Angewandte Chemie International Edition*, vol. 56, no. 28, pp. 8158–8162, Jul. 2017, doi: 10.1002/anie.201703970.

[66] Z. Zhang *et al.*, “High-Quality $(CH_3NH_3)_3Bi_2I_9$ Film-Based Solar Cells: Pushing Efficiency up to 1.64%,” *J Phys Chem Lett*, vol. 8, no. 17, pp. 4300–4307, Sep. 2017, doi: 10.1021/acs.jpclett.7b01952.

[67] M. Wang *et al.*, “High-Quality Sequential-Vapor-Deposited $Cs_2AgBiBr_6$ Thin Films for Lead-Free Perovskite Solar Cells,” *Solar RRL*, vol. 2, no. 12, p. 1800217, Dec. 2018, doi: 10.1002/solr.201800217.

[68] C. Zuo and L. Ding, “Lead-free Perovskite Materials (NH $_4$) $_3$ Sb $_2$ I $_x$ Br $_{9-x}$,” *Angewandte Chemie International Edition*, vol. 56, no. 23, pp. 6528–6532, Jun. 2017, doi: 10.1002/anie.201702265.

[69] M. Chen *et al.*, “Cesium Titanium(IV) Bromide Thin Films Based Stable Lead-free Perovskite Solar Cells,” *Joule*, vol. 2, no. 3, pp. 558–570, Mar. 2018, doi: 10.1016/j.joule.2018.01.009.

[70] M.-G. Ju *et al.*, “Earth-Abundant Nontoxic Titanium(IV)-based Vacancy-Ordered Double Perovskite Halides with Tunable 1.0 to 1.8 eV Bandgaps for Photovoltaic Applications,” *ACS Energy Lett*, vol. 3, no. 2, pp. 297–304, Feb. 2018, doi: 10.1021/acsenergylett.7b01167.

[71] T. Krishnamoorthy *et al.*, “Lead-free germanium iodide perovskite materials for photovoltaic applications,” *J Mater Chem A Mater*, vol. 3, no. 47, pp. 23829–23832, 2015, doi: 10.1039/C5TA05741H.

[72] F. Hao, C. C. Stoumpos, D. H. Cao, R. P. H. Chang, and M. G. Kanatzidis, “Lead-free solid-state organic–inorganic halide perovskite solar cells,” *Nat Photonics*, vol. 8, no. 6, pp. 489–494, Jun. 2014, doi: 10.1038/nphoton.2014.82.

[73] M. I. Khan *et al.*, “Bi and Sn Doping Improved the Structural, Optical and Photovoltaic Properties of MAPbI3-Based Perovskite Solar Cells,” *Materials*, vol. 15, no. 15, p. 5216, Jul. 2022, doi: 10.3390/ma15155216.

[74] L. Mao, C. C. Stoumpos, and M. G. Kanatzidis, “Two-Dimensional Hybrid Halide Perovskites: Principles and Promises,” *J Am Chem Soc*, vol. 141, no. 3, pp. 1171–1190, Jan. 2019, doi: 10.1021/jacs.8b10851.

[75] W. Ke, C. C. Stoumpos, and M. G. Kanatzidis, “‘Unleaded’ Perovskites: Status Quo and Future Prospects of Tin-Based Perovskite Solar Cells,” *Advanced Materials*, vol. 31, no. 47, p. 1803230, Nov. 2019, doi: 10.1002/adma.201803230.

[76] C. C. Stoumpos and M. G. Kanatzidis, “Halide Perovskites: Poor Man’s High-Performance Semiconductors,” *Advanced Materials*, vol. 28, no. 28, pp. 5778–5793, Jul. 2016, doi: 10.1002/adma.201600265.

[77] M. Chen *et al.*, “Highly stable and efficient all-inorganic lead-free perovskite solar cells with native-oxide passivation,” *Nat Commun*, vol. 10, no. 1, p. 16, Jan. 2019, doi: 10.1038/s41467-018-07951-y.

[78] F. Giustino and H. J. Snaith, “Toward Lead-Free Perovskite Solar Cells,” *ACS Energy Lett*, vol. 1, no. 6, pp. 1233–1240, Dec. 2016, doi: 10.1021/acsenergylett.6b00499.

[79] S. F. Hoefler, G. Trimmel, and T. Rath, “Progress on lead-free metal halide perovskites for photovoltaic applications: a review,” *Monatshefte für Chemie - Chemical Monthly*, vol. 148, no. 5, pp. 795–826, May 2017, doi: 10.1007/s00706-017-1933-9.

[80] S. Tsarev *et al.*, “Hydrazinium-assisted stabilisation of methylammonium tin iodide for lead-free perovskite solar cells,” *J Mater Chem A Mater*, vol. 6, no. 43, pp. 21389–21395, 2018, doi: 10.1039/C8TA07699E.

[81] Z. Chen *et al.*, “Photoluminescence study of polycrystalline CsSnI3 thin films: Determination of exciton binding energy,” *J Lumin*, vol. 132, no. 2, pp. 345–349, Feb. 2012, doi: 10.1016/j.jlumin.2011.09.006.

[82] L. and L. W. R. L. Huang, “Electronic band structure, phonons, and exciton binding energies of halide perovskites CsSnCl3, CsSnBr3, and CsSnI3,” *Phys. Rev. B*, vol. 88, no. 16, p. 165203, Oct. 2013.

[83] T. M. Koh *et al.*, “Formamidinium tin-based perovskite with low E_g for photovoltaic applications,” *J Mater Chem A Mater*, vol. 3, no. 29, pp. 14996–15000, 2015, doi: 10.1039/C5TA00190K.

[84] N. K. Noel *et al.*, “Lead-free organic–inorganic tin halide perovskites for photovoltaic applications,” *Energy Environ. Sci.*, vol. 7, no. 9, pp. 3061–3068, 2014, doi: 10.1039/C4EE01076K.

[85] M. H. Kumar *et al.*, “Lead-Free Halide Perovskite Solar Cells with High Photocurrents Realized Through Vacancy Modulation,” *Advanced Materials*, vol. 26, no. 41, pp. 7122–7127, Nov. 2014, doi: 10.1002/adma.201401991.

[86] F. Hao, C. C. Stoumpos, D. H. Cao, R. P. H. Chang, and M. G. Kanatzidis, “Lead-free solid-state organic–inorganic halide perovskite solar cells,” *Nat Photonics*, vol. 8, no. 6, pp. 489–494, Jun. 2014, doi: 10.1038/nphoton.2014.82.

[87] Y. Ogomi *et al.*, “$CH_3NH_3Sn_xPb_{(1-x)}I_3$ Perovskite Solar Cells Covering up to 1060 nm,” *J Phys Chem Lett*, vol. 5, no. 6, pp. 1004–1011, Mar. 2014, doi: 10.1021/jz5002117.

[88] R. Kour *et al.*, “Potential Substitutes for Replacement of Lead in Perovskite Solar Cells: A Review,” *Global Challenges*, vol. 3, no. 11, p. 1900050, Nov. 2019, doi: 10.1002/gch2.201900050.

[89] Y. Dang *et al.*, “Formation of Hybrid Perovskite Tin Iodide Single Crystals by Top-Seeded Solution Growth,” *Angewandte Chemie International Edition*, vol. 55, no. 10, pp. 3447–3450, Mar. 2016, doi: 10.1002/anie.201511792.

[90] C. C. Stoumpos *et al.*, “Hybrid Germanium Iodide Perovskite Semiconductors: Active Lone Pairs, Structural Distortions, Direct and Indirect Energy Gaps, and Strong Nonlinear Optical Properties,” *J Am Chem Soc*, vol. 137, no. 21, pp. 6804–6819, Jun. 2015, doi: 10.1021/jacs.5b01025.

[91] M. Chen *et al.*, “Highly stable and efficient all-inorganic lead-free perovskite solar cells with native-oxide passivation,” *Nat Commun*, vol. 10, no. 1, p. 16, Jan. 2019, doi: 10.1038/s41467-018-07951-y.

[92] I. Mora-Seró and J. Bisquert, “Breakthroughs in the Development of Semiconductor-Sensitized Solar Cells,” *J Phys Chem Lett*, vol. 1, no. 20, pp. 3046–3052, Oct. 2010, doi: 10.1021/jz100863b.

[93] K. T. Butler, S. McKechnie, P. Azarhoosh, M. van Schilfgaarde, D. O. Scanlon, and A. Walsh, “Quasi-particle electronic band structure and alignment of the V-VI-VII semiconductors SbSI, SbSBr, and SbSeI for solar cells,” *Appl Phys Lett*, vol. 108, no. 11, p. 112103, Mar. 2016, doi: 10.1063/1.4943973.

[94] S.-J. Moon, Y. Itzhaik, J.-H. Yum, S. M. Zakeeruddin, G. Hodes, and M. Grätzel, “Sb_2S_3-Based Mesoscopic Solar Cell using an Organic Hole Conductor,” *J Phys Chem Lett*, vol. 1, no. 10, pp. 1524–1527, May 2010, doi: 10.1021/jz100308q.

[95] L. Zhang *et al.*, “Sequential deposition route to efficient Sb_2S_3 solar cells,” *J Mater Chem A Mater*, vol. 6, no. 43, pp. 21320–21326, 2018, doi: 10.1039/C8TA08296K.

[96] R. Nie, A. Mehta, B. Park, H.-W. Kwon, J. Im, and S. Il Seok, “Mixed Sulfur and Iodide-Based Lead-Free Perovskite Solar Cells,” *J Am Chem Soc*, vol. 140, no. 3, pp. 872–875, Jan. 2018, doi: 10.1021/jacs.7b11332.

[97] L. Whittaker-Brooks, J. Gao, A. K. Hailey, C. R. Thomas, N. Yao, and Y.-L. Loo, “Bi_2S_3 nanowire networks as electron acceptor layers in solution-processed hybrid solar cells,” *J. Mater. Chem. C*, vol. 3, no. 11, pp. 2686–2692, 2015, doi: 10.1039/C4TC02534B.

[98] R. E. Brandt *et al.*, “Investigation of Bismuth Triiodide (BiI_3) for Photovoltaic Applications,” *J Phys Chem Lett*, vol. 6, no. 21, pp. 4297–4302, Nov. 2015, doi: 10.1021/acs.jpclett.5b02022.

[99] B. Wagner, A. Huttner, D. Bischof, A. Engel, G. Witte, and J. Heine, “Chemical Surface Reactivity and Morphological Changes of Bismuth Triiodide (BiI_3) under Different Environmental Conditions,” *Langmuir*, vol. 36, no. 23, pp. 6458–6464, Jun. 2020, doi: 10.1021/acs.langmuir.0c00740.

[100] N. T. Hahn, A. J. E. Rettie, S. K. Beal, R. R. Fullon, and C. B. Mullins, “n-BiSI Thin Films: Selenium Doping and Solar Cell Behavior,” *The Journal of Physical Chemistry C*, vol. 116, no. 47, pp. 24878–24886, Nov. 2012, doi: 10.1021/jp3088397.

[101] A. M. Ganose, K. T. Butler, A. Walsh, and D. O. Scanlon, “Relativistic electronic structure and band alignment of BiSI and BiSeI: candidate photovoltaic materials,” *J Mater Chem A Mater*, vol. 4, no. 6, pp. 2060–2068, 2016, doi: 10.1039/C5TA09612J.

[102] B.-W. Park, B. Philippe, X. Zhang, H. Rensmo, G. Boschloo, and E. M. J. Johansson, “Bismuth Based Hybrid Perovskites $A_3Bi_2I_9$ (A: Methylammonium or Cesium) for Solar Cell Application,” *Advanced Materials*, vol. 27, no. 43, pp. 6806–6813, Nov. 2015, doi: 10.1002/adma.201501978.

[103] Y. Kim *et al.*, “Pure Cubic-Phase Hybrid Iodobismuthates $AgBi_2I_7$ for Thin-Film Photovoltaics,” *Angewandte Chemie International Edition*, vol. 55, no. 33, pp. 9586–9590, Aug. 2016, doi: 10.1002/anie.201603608.

[104] W. Gao *et al.*, "High-Quality $Cs_2AgBiBr_6$ Double Perovskite Film for Lead-Free Inverted Planar Heterojunction Solar Cells with 2.2 % Efficiency," *ChemPhysChem*, vol. 19, no. 14, pp. 1696–1700, Jul. 2018, doi: 10.1002/cphc.201800346.

[105] E. Greul, M. L. Petrus, A. Binek, P. Docampo, and T. Bein, "Highly stable, phase pure $Cs_2AgBiBr_6$ double perovskite thin films for optoelectronic applications," *J Mater Chem A Mater*, vol. 5, no. 37, pp. 19972–19981, 2017, doi: 10.1039/C7TA06816F.

[106] Z. Chen *et al.*, "Processing and Preparation Method for High-Quality Opto-Electronic Perovskite Film," *Front Mater*, vol. 8, Oct. 2021, doi: 10.3389/fmats.2021.723169.

[107] M.-C. Wu and Y.-H. Chang, "Perovskite-Structured Photovoltaic Materials," in *Solar Panels and Photovoltaic Materials*, InTech, 2018. doi: 10.5772/intechopen.74997.

[108] K.-M. Lee, W.-J. Lin, S.-H. Chen, and M.-C. Wu, "Control of TiO2 electron transport layer properties to enhance perovskite photovoltaics performance and stability," *Org Electron*, vol. 77, p. 105406, Feb. 2020, doi: 10.1016/j.orgel.2019.105406.

[109] M.-C. Wu *et al.*, "Enhancing the efficiency of perovskite solar cells using mesoscopic zinc-doped TiO_2 as the electron extraction layer through band alignment," *J Mater Chem A Mater*, vol. 6, no. 35, pp. 16920–16931, 2018, doi: 10.1039/C8TA05291C.

[110] M. F. Mohamad Noh *et al.*, "The architecture of the electron transport layer for a perovskite solar cell," *J Mater Chem C Mater*, vol. 6, no. 4, pp. 682–712, 2018, doi: 10.1039/C7TC04649A.

[111] A. Rajagopal, K. Yao, and A. K. -Y. Jen, "Toward Perovskite Solar Cell Commercialization: A Perspective and Research Roadmap Based on Interfacial Engineering," *Advanced Materials*, vol. 30, no. 32, p. 1800455, Aug. 2018, doi: 10.1002/adma.201800455.

[112] H. Lu, W. Tian, B. Gu, Y. Zhu, and L. Li, "TiO_2 Electron Transport Bilayer for Highly Efficient Planar Perovskite Solar Cell," *Small*, vol. 13, no. 38, p. 1701535, Oct. 2017, doi: 10.1002/smll.201701535.

[113] W. Ye, J. Xiang, F. Huang, and D. Zhong, "Towards large-area perovskite solar cells: the influence of compact and mesoporous TiO_2 electron transport layers," *Mater Res Express*, vol. 5, no. 8, p. 085506, Jul. 2018, doi: 10.1088/2053-1591/aad311.

[114] A. J. Frank, N. Kopidakis, and J. van de Lagemaat, “Electrons in nanostructured TiO2 solar cells: transport, recombination and photovoltaic properties,” *Coord Chem Rev*, vol. 248, no. 13–14, pp. 1165–1179, Jul. 2004, doi: 10.1016/j.ccr.2004.03.015.

[115] L. Huang *et al.*, “Efficient planar perovskite solar cells without a high temperature processed titanium dioxide electron transport layer,” *Solar Energy Materials and Solar Cells*, vol. 149, pp. 1–8, May 2016, doi: 10.1016/j.solmat.2015.12.033.

[116] J. Choi, S. Song, M. T. Hörantner, H. J. Snaith, and T. Park, “Well-Defined Nanostructured, Single-Crystalline TiO_2 Electron Transport Layer for Efficient Planar Perovskite Solar Cells,” *ACS Nano*, vol. 10, no. 6, pp. 6029–6036, Jun. 2016, doi: 10.1021/acsnano.6b01575.

[117] F. Azri, A. Meftah, N. Sengouga, and A. Meftah, “Electron and hole transport layers optimization by numerical simulation of a perovskite solar cell,” *Solar Energy*, vol. 181, pp. 372–378, Mar. 2019, doi: 10.1016/j.solener.2019.02.017.

[118] J. H. Heo, H. J. Han, D. Kim, T. K. Ahn, and S. H. Im, “Hysteresis-less inverted $CH_3NH_3PbI_3$ planar perovskite hybrid solar cells with 18.1% power conversion efficiency,” *Energy Environ Sci*, vol. 8, no. 5, pp. 1602–1608, 2015, doi: 10.1039/C5EE00120J.

[119] J. H. Heo *et al.*, “Efficient inorganic–organic hybrid heterojunction solar cells containing perovskite compound and polymeric hole conductors,” *Nat Photonics*, vol. 7, no. 6, pp. 486–491, Jun. 2013, doi: 10.1038/nphoton.2013.80.

[120] S. W. Lee *et al.*, “Improved Cu_2O-Based Solar Cells Using Atomic Layer Deposition to Control the Cu Oxidation State at the p-n Junction,” *Adv Energy Mater*, vol. 4, no. 11, p. 1301916, Aug. 2014, doi: 10.1002/aenm.201301916.

[121] P. Pattanasattayavong *et al.*, “Hole-Transporting Transistors and Circuits Based on the Transparent Inorganic Semiconductor Copper(I) Thiocyanate (CuSCN) Processed from Solution at Room Temperature,” *Advanced Materials*, vol. 25, no. 10, pp. 1504–1509, Mar. 2013, doi: 10.1002/adma.201202758.

[122] P. Pattanasattayavong *et al.*, “Electric field-induced hole transport in copper(i) thiocyanate (CuSCN) thin-films processed from solution at room temperature,” *Chem. Commun.*, vol. 49, no. 39, pp. 4154–4156, 2013, doi: 10.1039/C2CC37065D.

[123] N. D. Treat *et al.*, “Copper thiocyanate: An attractive hole transport/extraction layer for use in organic photovoltaic cells,” *Appl Phys Lett*, vol. 107, no. 1, p. 013301, Jul. 2015, doi: 10.1063/1.4926408.

[124] N. Yaacobi-Gross *et al.*, “High-Efficiency Organic Photovoltaic Cells Based on the Solution-Processable Hole Transporting Interlayer Copper Thiocyanate (CuSCN) as a Replacement for PEDOT:PSS,” *Adv Energy Mater*, vol. 5, no. 3, p. 1401529, Feb. 2015, doi: 10.1002/aenm.201401529.

[125] S. Ye *et al.*, “CuSCN-Based Inverted Planar Perovskite Solar Cell with an Average PCE of 15.6%,” *Nano Lett*, vol. 15, no. 6, pp. 3723–3728, Jun. 2015, doi: 10.1021/acs.nanolett.5b00116.

[126] K. Zhao, R. Munir, B. Yan, Y. Yang, T. Kim, and A. Amassian, “Solution-processed inorganic copper(<scp>i</scp>) thiocyanate (CuSCN) hole transporting layers for efficient p–i–n perovskite solar cells,” *J Mater Chem A Mater*, vol. 3, no. 41, pp. 20554–20559, 2015, doi: 10.1039/C5TA04028K.

[127] P. Qin *et al.*, “Inorganic hole conductor-based lead halide perovskite solar cells with 12.4% conversion efficiency,” *Nat Commun*, vol. 5, no. 1, p. 3834, May 2014, doi: 10.1038/ncomms4834.

[128] S. Chavhan *et al.*, “Organo-metal halide perovskite-based solar cells with CuSCN as the inorganic hole selective contact,” *J. Mater. Chem. A*, vol. 2, no. 32, pp. 12754–12760, 2014, doi: 10.1039/C4TA01310G.

[129] M. D. Irwin, D. B. Buchholz, A. W. Hains, R. P. H. Chang, and T. J. Marks, “*p* -Type semiconducting nickel oxide as an efficiency-enhancing anode interfacial layer in polymer bulk-heterojunction solar cells,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 8, pp. 2783–2787, Feb. 2008, doi: 10.1073/pnas.0711990105.

[130] J. A. Caraveo-Frescas, P. K. Nayak, H. A. Al-Jawhari, D. B. Granato, U. Schwingenschlögl, and H. N. Alshareef, “Record Mobility in Transparent p-Type Tin Monoxide Films and Devices by Phase Engineering,” *ACS Nano*, vol. 7, no. 6, pp. 5160–5167, Jun. 2013, doi: 10.1021/nn400852r.

[131] E. Fortunato *et al.*, “Thin-film transistors based on p-type Cu2O thin films produced at room temperature,” *Appl Phys Lett*, vol. 96, no. 19, p. 192102, May 2010, doi: 10.1063/1.3428434.

[132] M. I. Hossain and F. H. Alharbi, “Recent advances in alternative material photovoltaics,” *Materials Technology*, vol. 28, no. 1–2, pp. 88–97, Mar. 2013, doi: 10.1179/1753555712Y.0000000039.

[133] T. Minami, T. Miyata, and Y. Nishi, “Cu2O-based heterojunction solar cells with an Al-doped ZnO/oxide semiconductor/thermally oxidized Cu2O sheet structure,” *Solar Energy*, vol. 105, pp. 206–217, Jul. 2014, doi: 10.1016/j.solener.2014.03.036.

[134] F. Anwar, R. Mahbub, S. S. Satter, and S. M. Ullah, “Effect of Different HTM Layers and Electrical Parameters on ZnO Nanorod-Based Lead-Free Perovskite Solar Cell for High-Efficiency Performance,” *International Journal of Photoenergy*, vol. 2017, pp. 1–9, 2017, doi: 10.1155/2017/9846310.

[135] F. Behrouznejad, S. Shahbazi, N. Taghavinia, H.-P. Wu, and E. Wei-Guang Diau, “A study on utilizing different metals as the back contact of $CH_3NH_3PbI_3$ perovskite solar cells,” *J Mater Chem A Mater*, vol. 4, no. 35, pp. 13488–13498, 2016, doi: 10.1039/C6TA05938D.

[136] H. B. Michaelson, “The work function of the elements and its periodicity,” *J Appl Phys*, vol. 48, no. 11, pp. 4729–4733, Nov. 1977, doi: 10.1063/1.323539.

[137] D. Zhou, T. Zhou, Y. Tian, X. Zhu, and Y. Tu, “Perovskite-Based Solar Cells: Materials, Methods, and Future Perspectives,” *J Nanomater*, vol. 2018, pp. 1–15, 2018, doi: 10.1155/2018/8148072.

[138] M. R. Jani *et al.*, “Exploring solar cell performance of inorganic Cs2TiBr6 halide double perovskite: A numerical study,” *Superlattices Microstruct*, vol. 146, p. 106652, Oct. 2020, doi: 10.1016/j.spmi.2020.106652.

[139] T. Leijtens, K. A. Bush, R. Prasanna, and M. D. McGehee, “Opportunities and challenges for tandem solar cells using metal halide perovskite semiconductors,” *Nat Energy*, vol. 3, no. 10, pp. 828–838, Jul. 2018, doi: 10.1038/s41560-018-0190-4.

[140] M. Saliba *et al.*, “Incorporation of rubidium cations into perovskite solar cells improves photovoltaic performance,” *Science (1979)*, vol. 354, no. 6309, pp. 206–209, Oct. 2016, doi: 10.1126/science.aah5557.

[141] K. A. Bush *et al.*, “23.6%-efficient monolithic perovskite/silicon tandem solar cells with improved stability,” *Nat Energy*, vol. 2, no. 4, p. 17009, Feb. 2017, doi: 10.1038/nenergy.2017.9.

[142] J. Werner *et al.*, “Efficient Near-Infrared-Transparent Perovskite Solar Cells Enabling Direct Comparison of 4-Terminal and Monolithic Perovskite/Silicon Tandem Cells,” *ACS Energy Lett*, vol. 1, no. 2, pp. 474–480, Aug. 2016, doi: 10.1021/acsenergylett.6b00254.

[143] J. Werner *et al.*, “Efficient Monolithic Perovskite/Silicon Tandem Solar Cell with Cell Area >1 cm ^{2},” *J Phys Chem Lett*, vol. 7, no. 1, pp. 161–166, Jan. 2016, doi: 10.1021/acs.jpclett.5b02686.

[144] J. P. Mailoa *et al.*, “A 2-terminal perovskite/silicon multijunction solar cell enabled by a silicon tunnel junction,” *Appl Phys Lett*, vol. 106, no. 12, p. 121105, Mar. 2015, doi: 10.1063/1.4914179.

[145] Y. Liu, W. Yan, H. Zhu, Y. Tu, L. Guan, and X. Tan, “Study on bandgap predications of ABX3-type perovskites by machine learning,” *Org Electron*, vol. 101, p. 106426, Feb. 2022, doi: 10.1016/j.orgel.2021.106426.

[146] F. Hao, C. C. Stoumpos, R. P. H. Chang, and M. G. Kanatzidis, “Anomalous Band Gap Behavior in Mixed Sn and Pb Perovskites Enables Broadening of Absorption Spectrum in Solar Cells,” *J Am Chem Soc*, vol. 136, no. 22, pp. 8094–8099, Jun. 2014, doi: 10.1021/ja5033259.

[147] E. Mosconi, P. Umari, and F. De Angelis, “Electronic and optical properties of mixed Sn–Pb organohalide perovskites: a first principles investigation,” *J Mater Chem A Mater*, vol. 3, no. 17, pp. 9208–9215, 2015, doi: 10.1039/C4TA06230B.

[148] F. Zuo, S. T. Williams, P.-W. Liang, C.-C. Chueh, C.-Y. Liao, and A. K.-Y. Jen, “Binary-Metal Perovskites Toward High-Performance Planar-Heterojunction Hybrid Solar Cells,” *Advanced Materials*, vol. 26, no. 37, pp. 6454–6460, Oct. 2014, doi: 10.1002/adma.201401641.

[149] H.-J. Feng, T. R. Paudel, E. Y. Tsymbal, and X. C. Zeng, “Tunable Optical Properties and Charge Separation in $CH_3NH_3Sn_xPb_{1-x}I_3/TiO_2$ -Based Planar Perovskites Cells,” *J Am Chem Soc*, vol. 137, no. 25, pp. 8227–8236, Jul. 2015, doi: 10.1021/jacs.5b04015.

[150] J. Im, C. C. Stoumpos, H. Jin, A. J. Freeman, and M. G. Kanatzidis, “Antagonism between Spin–Orbit Coupling and Steric Effects Causes Anomalous Band Gap Evolution in the Perovskite Photovoltaic Materials $CH_3NH_3Sn_{1-x}Pb_xI_3$,” *J Phys Chem Lett*, vol. 6, no. 17, pp. 3503–3509, Sep. 2015, doi: 10.1021/acs.jpclett.5b01738.

[151] P. Hohenberg and W. Kohn, “Inhomogeneous Electron Gas,” *Physical Review*, vol. 136, no. 3B, pp. B864–B871, Nov. 1964, doi: 10.1103/PhysRev.136.B864.

[152] A. Hussain *et al.*, “Monte Carlo simulation study of electron yields from compound semiconductor materials,” *J Appl Phys*, vol. 128, no. 1, p. 015305, Jul. 2020, doi: 10.1063/5.0012154.

[153] B. J. Alder and T. E. Wainwright, “Studies in Molecular Dynamics. I. General Method,” *J Chem Phys*, vol. 31, no. 2, pp. 459–466, Aug. 1959, doi: 10.1063/1.1730376.

[154] A. Agrawal and A. Choudhary, “Perspective: Materials informatics and big data: Realization of the ‘fourth paradigm’ of science in materials science,” *APL Mater*, vol. 4, no. 5, p. 053208, May 2016, doi: 10.1063/1.4946894.

[155] P. Raccuglia *et al.*, “Machine-learning-assisted materials discovery using failed experiments,” *Nature*, vol. 533, no. 7601, pp. 73–76, May 2016, doi: 10.1038/nature17439.

[156] P. V. Balachandran, B. Kowalski, A. Sehirlioglu, and T. Lookman, “Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning,” *Nat Commun*, vol. 9, no. 1, p. 1668, Apr. 2018, doi: 10.1038/s41467-018-03821-9.

[157] D. Dai *et al.*, “Using machine learning and feature engineering to characterize limited material datasets of high-entropy alloys,” *Comput Mater Sci*, vol. 175, p. 109618, Apr. 2020, doi: 10.1016/j.commatsci.2020.109618.

[158] W. Sun *et al.*, “Machine learning–assisted molecular design and efficiency prediction for high-performance organic photovoltaic materials,” *Sci Adv*, vol. 5, no. 11, Nov. 2019, doi: 10.1126/sciadv.aay4275.

[159] V. Stanev *et al.*, “Machine learning modeling of superconducting critical temperature,” *NPJ Comput Mater*, vol. 4, no. 1, p. 29, Jun. 2018, doi: 10.1038/s41524-018-0085-8.

[160] M. Rupp, “Machine learning for quantum mechanics in a nutshell,” *Int J Quantum Chem*, vol. 115, no. 16, pp. 1058–1073, Aug. 2015, doi: 10.1002/qua.24954.

[161] B. R. Goldsmith, J. Esterhuizen, J. Liu, C. J. Bartel, and C. Sutton, “Machine learning for heterogeneous catalyst design and discovery,” *AIChE Journal*, vol. 64, no. 7, pp. 2311–2323, Jul. 2018, doi: 10.1002/aic.16198.

[162] S. Körbel, M. A. L. Marques, and S. Botti, “Stability and electronic properties of new inorganic perovskites from high-throughput ab initio calculations,” *J Mater Chem C Mater*, vol. 4, no. 15, pp. 3157–3167, 2016, doi: 10.1039/C5TC04172D.

[163] J. Im, S. Lee, T.-W. Ko, H. W. Kim, Y. Hyon, and H. Chang, “Identifying Pb-free perovskites for solar cells by machine learning,” *NPJ Comput Mater*, vol. 5, no. 1, p. 37, Mar. 2019, doi: 10.1038/s41524-019-0177-0.

[164] Z. Guo and B. Lin, “Machine learning stability and band gap of lead-free halide double perovskite materials for perovskite solar cells,” *Solar Energy*, vol. 228, pp. 689–699, Nov. 2021, doi: 10.1016/j.solener.2021.09.030.

[165] Z. Gao *et al.*, “Screening for lead-free inorganic double perovskites with suitable band gaps and high stability using combined machine learning and DFT calculation,” *Appl Surf Sci*, vol. 568, 2021, doi: 10.1016/j.apsusc.2021.150916.

[166] G. Pilania, A. Mannodi-Kanakkithodi, B. P. Uberuaga, R. Ramprasad, J. E. Gubernatis, and T. Lookman, “Machine learning bandgaps of double perovskites,” *Sci Rep*, vol. 6, no. 1, p. 19375, Jan. 2016, doi: 10.1038/srep19375.

[167] L. Zhang, M. He, and S. Shao, “Machine learning for halide perovskite materials,” *Nano Energy*, vol. 78, p. 105380, Dec. 2020, doi: 10.1016/j.nanoen.2020.105380.

[168] X. Yang, L. Li, Q. Tao, W. Lu, and M. Li, “Rapid discovery of narrow bandgap oxide double perovskites using machine learning,” *Comput Mater Sci*, vol. 196, p. 110528, Aug. 2021, doi: 10.1016/j.commatsci.2021.110528.

[169] R. Vasudevan, G. Pilania, and P. V. Balachandran, “Machine learning for materials design and discovery,” *J Appl Phys*, vol. 129, no. 7, p. 070401, Feb. 2021, doi: 10.1063/5.0043300.

[170] S. Lu, Q. Zhou, Y. Ouyang, Y. Guo, Q. Li, and J. Wang, “Accelerated discovery of stable lead-free hybrid organic-inorganic perovskites via machine learning,” *Nat Commun*, vol. 9, no. 1, Dec. 2018, doi: 10.1038/s41467-018-05761-w.

[171] Ç. Odabaşı and R. Yıldırım, “Machine learning analysis on stability of perovskite solar cells,” *Solar Energy Materials and Solar Cells*, vol. 205, p. 110284, Feb. 2020, doi: 10.1016/j.solmat.2019.110284.

[172] Z. Guo and B. Lin, “Machine learning stability and band gap of lead-free halide double perovskite materials for perovskite solar cells,” *Solar Energy*, vol. 228, pp. 689–699, Nov. 2021, doi: 10.1016/j.solener.2021.09.030.

[173] B. Weng *et al.*, “Simple descriptor derived from symbolic regression accelerating the discovery of new perovskite catalysts,” *Nat Commun*, vol. 11, no. 1, p. 3513, Jul. 2020, doi: 10.1038/s41467-020-17263-9.

[174] Md. T. Islam *et al.*, “Investigation of non-Pb all-perovskite 4-T mechanically stacked and 2-T monolithic tandem solar devices utilizing SCAPS simulation,” *SN Appl Sci*, vol. 3, no. 4, p. 504, Apr. 2021, doi: 10.1007/s42452-021-04487-7.

[175] M. Burgelman, P. Nollet, and S. Degrave, “Modelling polycrystalline semiconductor solar cells,” *Thin Solid Films*, vol. 361–362, pp. 527–532, Feb. 2000, doi: 10.1016/S0040-6090(99)00825-1.

[176] H. Al Jame *et al.*, “Supervised Machine Learning-Aided SCAPS-Based Quantitative Analysis for the Discovery of Optimum Bromine Doping in Methylammonium Tin-Based Perovskite (MASnI3- xBr x),” *ACS Appl Mater Interfaces*, vol. 14, no. 1, 2022, doi: 10.1021/acsami.1c15030.

[177] “Simulation programme SCAPS-1D for thin film solar cells developed at ELIS, University of Gent.” https://scaps.elis.ugent.be/ (accessed Jul. 07, 2023).

[178] Y. Zhuo, A. Mansouri Tehrani, and J. Brgoch, “Predicting the Band Gaps of Inorganic Solids by Machine Learning,” *Journal of Physical Chemistry Letters*, vol. 9, no. 7, pp. 1668–1673, Apr. 2018, doi: 10.1021/acs.jpclett.8b00124.

[179] J. Im, S. Lee, T.-W. Ko, H. W. Kim, Y. Hyon, and H. Chang, “Identifying Pb-free perovskites for solar cells by machine learning,” *NPJ Comput Mater*, vol. 5, no. 1, p. 37, Mar. 2019, doi: 10.1038/s41524-019-0177-0.

[180] “Periodic Table,” *Royal Society of Chemistry*. https://www.rsc.org/periodic-table/ (accessed Aug. 02, 2023).

[181] “Web Elements.” https://webelements.com/ (accessed Aug. 02, 2023).

[182] G. Kieslich, S. Sun, and A. K. Cheetham, “Solid-state principles applied to organic–inorganic perovskites: new tricks for an old dog,” *Chem. Sci.*, vol. 5, no. 12, pp. 4712–4715, 2014, doi: 10.1039/C4SC02211D.

[183] R. D. Shannon, “Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides,” *Acta Crystallographica Section A*, vol. 32, no. 5, pp. 751–767, Sep. 1976, doi: 10.1107/S0567739476001551.

[184] B. Saparov and D. B. Mitzi, “Organic–Inorganic Perovskites: Structural Versatility for Functional Materials Design,” *Chem Rev*, vol. 116, no. 7, pp. 4558–4596, Apr. 2016, doi: 10.1021/acs.chemrev.5b00715.

[185] P. Bhagwan, *A Handbook of Inorganic Chemistry* . International Scientific Publishing Academy, 2005.

[186] D. Zhou, T. Zhou, Y. Tian, X. Zhu, and Y. Tu, “Perovskite-Based Solar Cells: Materials, Methods, and Future Perspectives,” *J Nanomater*, vol. 2018, pp. 1–15, 2018, doi: 10.1155/2018/8148072.

[187] Y. Liu, T. Zhao, W. Ju, and S. Shi, “Materials discovery and design using machine learning,” *Journal of Materiomics*, vol. 3, no. 3, pp. 159–177, Sep. 2017, doi: 10.1016/j.jmat.2017.08.002.

[188] T. Nakajima and K. Sawada, “Discovery of Pb-Free Perovskite Solar Cells via High-Throughput Simulation on the K Computer,” *J Phys Chem Lett*, vol. 8, no. 19, pp. 4826–4831, Oct. 2017, doi: 10.1021/acs.jpclett.7b02203.

[189] H. J. Snaith, “Estimating the Maximum Attainable Efficiency in Dye-Sensitized Solar Cells,” *Adv Funct Mater*, vol. 20, no. 1, pp. 13–19, Jan. 2010, doi: 10.1002/adfm.200901476.

[190] W. Shockley and H. J. Queisser, “Detailed Balance Limit of Efficiency of *p-n* Junction Solar Cells,” *J Appl Phys*, vol. 32, no. 3, pp. 510–519, Mar. 1961, doi: 10.1063/1.1736034.

[191] H. Zhou *et al.*, “Interface engineering of highly efficient perovskite solar cells,” *Science (1979)*, vol. 345, no. 6196, pp. 542–546, Aug. 2014, doi: 10.1126/science.1254050.

[192] Z. Yang *et al.*, “Stable Low-Bandgap Pb-Sn Binary Perovskites for Tandem Solar Cells,” *Advanced Materials*, vol. 28, no. 40, pp. 8990–8997, Oct. 2016, doi: 10.1002/adma.201602696.

[193] T. Minemoto and M. Murata, “Impact of work function of back contact of perovskite solar cells without hole transport material analyzed by device simulation,” *Current Applied Physics*, vol. 14, no. 11, pp. 1428–1433, Nov. 2014, doi: 10.1016/j.cap.2014.08.002.

[194] S. N. Raghuram Nanduri, M. K. Siddiki, G. M. Chaudhry, and Y. Z. Alharthi, “Numerical simulation and performance optimization of perovskite solar cell,” in *2017 IEEE 44th Photovoltaic Specialist Conference (PVSC)*, IEEE, Jun. 2017, pp. 1018–1021. doi: 10.1109/PVSC.2017.8366586.

[195] Sajid *et al.*, “Novel hole transport layer of nickel oxide composite with carbon for high-performance perovskite solar cells,” *Chinese Physics B*, vol. 27, no. 1, p. 017305, Jan. 2018, doi: 10.1088/1674-1056/27/1/017305.

[196] Y. Zou, Y. Liang, C. Mu, and J. Zhang, “Enhancement of Open-Circuit Voltage of Perovskite Solar Cells by Interfacial Modification with *p* -Aminobenzoic Acid,” *Adv Mater Interfaces*, vol. 7, no. 1, p. 1901584, Jan. 2020, doi: 10.1002/admi.201901584.

[197] C. Liu, J. Fan, H. Li, C. Zhang, and Y. Mai, “Highly Efficient Perovskite Solar Cells with Substantial Reduction of Lead Content,” *Sci Rep*, vol. 6, no. 1, p. 35705, Oct. 2016, doi: 10.1038/srep35705.

[198] F. Bouhjar, S. Ullah, M. L. Chourou, M. Mollar, B. Marí, and B. Bessaïs, “Electrochemical Fabrication and Characterization of p-CuSCN/n-Fe $_2$ O $_3$ Heterojunction Devices for Hydrogen Production,” *J Electrochem Soc*, vol. 164, no. 13, pp. H936–H945, Oct. 2017, doi: 10.1149/2.1431713jes.

[199] , Mustafa M. A. Hussein et al., Mustafa M. A. Hussein et al., “An Investigation of Optical Qualities of Nickel Oxide Thin Films Doped with Gold Nanoparticles by Thermal Evaporation Process,” *International Journal of Mechanical and Production Engineering Research and Development*, vol. 8, no. 5, pp. 353–358, 2018, doi: 10.24247/ijmperdoct201840.

[200] T. Leijtens, R. Prasanna, A. Gold-Parker, M. F. Toney, and M. D. McGehee, “Mechanism of Tin Oxidation and Stabilization by Lead Substitution in Tin Halide Perovskites,” *ACS Energy Lett*, vol. 2, no. 9, pp. 2159–2165, Sep. 2017, doi: 10.1021/acsenergylett.7b00636.

[201] N. K. Noel *et al.*, “Lead-free organic–inorganic tin halide perovskites for photovoltaic applications,” *Energy Environ. Sci.*, vol. 7, no. 9, pp. 3061–3068, 2014, doi: 10.1039/C4EE01076K.

[202] G. Grancini *et al.*, “One-Year stable perovskite solar cells by 2D/3D interface engineering,” *Nat Commun*, vol. 8, no. 1, p. 15684, Jun. 2017, doi: 10.1038/ncomms15684.

[203] Y. Rong *et al.*, “Challenges for commercializing perovskite solar cells,” *Science (1979)*, vol. 361, no. 6408, Sep. 2018, doi: 10.1126/science.aat8235.

[204] Z. Zhao *et al.*, “Mixed-Organic-Cation Tin Iodide for Lead-Free Perovskite Solar Cells with an Efficiency of 8.12%,” *Advanced Science*, vol. 4, no. 11, p. 1700204, Nov. 2017, doi: 10.1002/advs.201700204.

[205] F. Hao, C. C. Stoumpos, D. H. Cao, R. P. H. Chang, and M. G. Kanatzidis, “Lead-free solid-state organic–inorganic halide perovskite solar cells,” *Nat Photonics*, vol. 8, no. 6, pp. 489–494, Jun. 2014, doi: 10.1038/nphoton.2014.82.

[206] N. Wang *et al.*, “Heterojunction-Depleted Lead-Free Perovskite Solar Cells with Coarse-Grained B-γ-$CsSnI_3$ Thin Films,” *Adv Energy Mater*, vol. 6, no. 24, p. 1601130, Dec. 2016, doi: 10.1002/aenm.201601130.

[207] S. Gupta, T. Bendikov, G. Hodes, and D. Cahen, “$CsSnBr_3$, A Lead-Free Halide Perovskite for Long-Term Solar Cell Application: Insights on SnF_2 Addition,” *ACS Energy Lett*, vol. 1, no. 5, pp. 1028–1033, Nov. 2016, doi: 10.1021/acsenergylett.6b00402.

[208] B. Lee, A. Krenselewski, S. Il Baik, D. N. Seidman, and R. P. H. Chang, “Solution processing of air-stable molecular semiconducting iodosalts, $Cs_2SnI_{6-x}Br_x$, for potential solar cell applications,” *Sustain Energy Fuels*, vol. 1, no. 4, pp. 710–724, 2017, doi: 10.1039/C7SE00100B.

[209] T. Krishnamoorthy *et al.*, “Lead-free germanium iodide perovskite materials for photovoltaic applications,” *J Mater Chem A Mater*, vol. 3, no. 47, pp. 23829–23832, 2015, doi: 10.1039/C5TA05741H.

[210] D. M. Fabian and S. Ardo, “Hybrid organic–inorganic solar cells based on bismuth iodide and 1,6-hexanediammonium dication,” *J Mater Chem A Mater*, vol. 4, no. 18, pp. 6837–6841, 2016, doi: 10.1039/C6TA00517A.

[211] J. H. Heo *et al.*, “Highly Stable All-Inorganic Pb-Free Perovskite Solar Cells,” *Journal of Nanoelectronics and Optoelectronics*, vol. 13, no. 12, pp. 1764–1768, Dec. 2018, doi: 10.1166/jno.2018.2407.

[212] E. Greul, M. L. Petrus, A. Binek, P. Docampo, and T. Bein, “Highly stable, phase pure $Cs_2AgBiBr_6$ double perovskite thin films for optoelectronic applications,” *J Mater Chem A Mater*, vol. 5, no. 37, pp. 19972–19981, 2017, doi: 10.1039/C7TA06816F.

[213] B. Saparov *et al.*, “Thin-Film Preparation and Characterization of $Cs_3Sb_2I_9$: A Lead-Free Layered Perovskite Semiconductor,” *Chemistry of Materials*, vol. 27, no. 16, pp. 5622–5632, Aug. 2015, doi: 10.1021/acs.chemmater.5b01989.

[214] P. C. Harikesh *et al.*, “Rb as an Alternative Cation for Templating Inorganic Lead-Free Perovskites for Solution Processed Photovoltaics,” *Chemistry of Materials*, vol. 28, no. 20, pp. 7496–7504, Oct. 2016, doi: 10.1021/acs.chemmater.6b03310.

[215] J.-C. Hebig, I. Kühn, J. Flohre, and T. Kirchartz, "Optoelectronic Properties of (CH_3NH_3)$_3Sb_2I_9$ Thin Films for Photovoltaic Applications," *ACS Energy Lett*, vol. 1, no. 1, pp. 309–314, Jul. 2016, doi: 10.1021/acsenergylett.6b00170.

[216] A. M. Elseman, A. E. Shalan, S. Sajid, M. M. Rashad, A. M. Hassan, and M. Li, "Copper-Substituted Lead Perovskite Materials Constructed with Different Halides for Working (CH_3NH_3)$_2CuX_4$ -Based Perovskite Solar Cells from Experimental and Theoretical View," *ACS Appl Mater Interfaces*, vol. 10, no. 14, pp. 11699–11707, Apr. 2018, doi: 10.1021/acsami.8b00495.

[217] D. Cortecchia *et al.*, "Lead-Free $MA_2CuCl_xBr_{4-x}$ Hybrid Perovskites," *Inorg Chem*, vol. 55, no. 3, pp. 1044–1052, Feb. 2016, doi: 10.1021/acs.inorgchem.5b01896.

[218] W. Zhou *et al.*, "Lead-Free Small-Bandgap $Cs_2CuSbCl_6$ Double Perovskite Nanocrystals," *J Phys Chem Lett*, vol. 11, no. 15, pp. 6463–6467, Aug. 2020, doi: 10.1021/acs.jpclett.0c01968.

[219] W. Deng *et al.*, "Synthesis of Cs2AgSbCl6 and improved optoelectronic properties of Cs2AgSbCl6/TiO2 heterostructure driven by the interface effect for lead-free double perovskites solar cells," *Appl Phys Lett*, vol. 111, no. 15, Oct. 2017, doi: 10.1063/1.4999192.

[220] M. B. Gray, E. T. McClure, and P. M. Woodward, "$Cs_2AgBiBr_{6-x}Cl_x$ solid solutions – band gap engineering with halide double perovskites," *J Mater Chem C Mater*, vol. 7, no. 31, pp. 9686–9689, 2019, doi: 10.1039/C9TC02674F.

[221] F. Wei *et al.*, "Enhanced visible light absorption for lead-free double perovskite $Cs_2AgSbBr_6$," *Chemical Communications*, vol. 55, no. 26, pp. 3721–3724, 2019, doi: 10.1039/C9CC01134J.

[222] A. H. Slavney *et al.*, "Small-Band-Gap Halide Double Perovskites," *Angewandte Chemie International Edition*, vol. 57, no. 39, pp. 12765–12770, Sep. 2018, doi: 10.1002/anie.201807421.

[223] Y. Saeed *et al.*, "$Cs_2NaGaBr_6$: a new lead-free and direct band gap halide double perovskite," *RSC Adv*, vol. 10, no. 30, pp. 17444–17451, 2020, doi: 10.1039/D0RA01764G.

[224] Y. Saeed *et al.*, "$Cs_2NaGaBr_6$: a new lead-free and direct band gap halide double perovskite," *RSC Adv*, vol. 10, no. 30, pp. 17444–17451, 2020, doi: 10.1039/D0RA01764G.

[225] F. Locardi *et al.*, "Colloidal Synthesis of Double Perovskite $Cs_2AgInCl_6$ and Mn-Doped $Cs_2AgInCl_6$ Nanocrystals," *J Am Chem Soc*, vol. 140, no. 40, pp. 12989–12995, Oct. 2018, doi: 10.1021/jacs.8b07983.

[226] H. Wu *et al.*, "Mixed-Halide Double Perovskite Cs_2AgBiX_6 (X=Br, I) with Tunable Optical Properties via Anion Exchange," *ChemSusChem*, vol. 14, no. 20, pp. 4507–4515, Oct. 2021, doi: 10.1002/cssc.202101146.

[227] M. Chen *et al.*, "Cesium Titanium(IV) Bromide Thin Films Based Stable Lead-free Perovskite Solar Cells," *Joule*, vol. 2, no. 3, pp. 558–570, Mar. 2018, doi: 10.1016/j.joule.2018.01.009.

[228] T.-B. Song *et al.*, "Importance of Reducing Vapor Atmosphere in the Fabrication of Tin-Based Perovskite Solar Cells," *J Am Chem Soc*, vol. 139, no. 2, pp. 836–842, Jan. 2017, doi: 10.1021/jacs.6b10734.

[229] Y. Hu *et al.*, "Bismuth Incorporation Stabilized α-$CsPbI_3$ for Fully Inorganic Perovskite Solar Cells," *ACS Energy Lett*, vol. 2, no. 10, pp. 2219–2227, Oct. 2017, doi: 10.1021/acsenergylett.7b00508.

[230] K. Fatima *et al.*, "Surface modification of CsPbI2Br for improved performance of inorganic perovskite solar cells," *Physica E Low Dimens Syst Nanostruct*, vol. 142, p. 115265, Aug. 2022, doi: 10.1016/j.physe.2022.115265.

[231] Z. Fang *et al.*, "Bandgap alignment of α-CsPbI3 perovskites with synergistically enhanced stability and optical performance via B-site minor doping," *Nano Energy*, vol. 61, pp. 389–396, Jul. 2019, doi: 10.1016/j.nanoen.2019.04.084.

[232] J. K. Nam *et al.*, "Potassium Incorporation for Enhanced Performance and Stability of Fully Inorganic Cesium Lead Halide Perovskite Solar Cells," *Nano Lett*, vol. 17, no. 3, pp. 2028–2033, Mar. 2017, doi: 10.1021/acs.nanolett.7b00050.

[233] C. Liu, W. Li, C. Zhang, Y. Ma, J. Fan, and Y. Mai, “All-Inorganic CsPbI $_2$ Br Perovskite Solar Cells with High Efficiency Exceeding 13%,” *J Am Chem Soc*, vol. 140, no. 11, pp. 3825–3828, Mar. 2018, doi: 10.1021/jacs.7b13229.

[234] Y. Ding *et al.*, “A low-cost hole transport layer enables CsPbI2Br single-junction and tandem perovskite solar cells with record efficiencies of 17.8 % and 21.4 %,” *Nano Today*, vol. 46, p. 101586, Oct. 2022, doi: 10.1016/j.nantod.2022.101586.

[235] C. F. J. Lau *et al.*, “Strontium-Doped Low-Temperature-Processed CsPbI $_2$ Br Perovskite Solar Cells,” *ACS Energy Lett*, vol. 2, no. 10, pp. 2319–2325, Oct. 2017, doi: 10.1021/acsenergylett.7b00751.

[236] Y. Guo, X. Yin, J. Liu, and W. Que, “Highly efficient CsPbIBr $_2$ perovskite solar cells with efficiency over 9.8% fabricated using a preheating-assisted spin-coating method,” *J Mater Chem A Mater*, vol. 7, no. 32, pp. 19008–19016, 2019, doi: 10.1039/C9TA03336J.

[237] J. Liang *et al.*, “CsPb $_{0.9}$ Sn $_{0.1}$ IBr $_2$ Based All-Inorganic Perovskite Solar Cells with Exceptional Efficiency and Stability,” *J Am Chem Soc*, vol. 139, no. 40, pp. 14009–14012, Oct. 2017, doi: 10.1021/jacs.7b07949.

[238] R. Ali, G.-J. Hou, Z.-G. Zhu, Q.-B. Yan, Q.-R. Zheng, and G. Su, “Predicted Lead-Free Perovskites for Solar Cells,” *Chemistry of Materials*, vol. 30, no. 3, pp. 718–728, Feb. 2018, doi: 10.1021/acs.chemmater.7b04036.

[239] G. E. Eperon, S. D. Stranks, C. Menelaou, M. B. Johnston, L. M. Herz, and H. J. Snaith, “Formamidinium lead trihalide: a broadly tunable perovskite for efficient planar heterojunction solar cells,” *Energy Environ Sci*, vol. 7, no. 3, p. 982, 2014, doi: 10.1039/c3ee43822h.

[240] J. H. Heo, D. H. Song, and S. H. Im, “Planar CH $_3$ NH $_3$ PbBr $_3$ Hybrid Solar Cells with 10.4% Power Conversion Efficiency, Fabricated by Controlled Crystallization in the Spin-Coating Process,” *Advanced Materials*, vol. 26, no. 48, pp. 8179–8183, Dec. 2014, doi: 10.1002/adma.201403140.

[241] T. Jesper Jacobsson *et al.*, “Exploration of the compositional space for mixed lead halogen perovskites for high efficiency solar cells,” *Energy Environ Sci*, vol. 9, no. 5, pp. 1706–1724, 2016, doi: 10.1039/C6EE00030D.

[242] A. Bhorde *et al.*, “Structural, Electronic, and Optical Properties of Lead-Free Halide Double Perovskite Rb2AgBiI6: A Combined Experimental and DFT Study,” *ES Materials & Manufacturing*, 2021, doi: 10.30919/esmm5f1042.

[243] X. Liu *et al.*, “Improved efficiency and stability of Pb–Sn binary perovskite solar cells by Cs substitution,” *J Mater Chem A Mater*, vol. 4, no. 46, pp. 17939–17945, 2016, doi: 10.1039/C6TA07712A.

[244] A. D. Jodlowski *et al.*, “Large guanidinium cation mixed with methylammonium in lead iodide perovskites for 19% efficient solar cells,” *Nat Energy*, vol. 2, no. 12, pp. 972–979, Dec. 2017, doi: 10.1038/s41560-017-0054-3.

[245] C. Ma *et al.*, “2D/3D perovskite hybrids as moisture-tolerant and efficient light absorbers for solar cells,” *Nanoscale*, vol. 8, no. 43, pp. 18309–18314, 2016, doi: 10.1039/C6NR04741F.

[246] G. Longo, C. Momblona, M.-G. La-Placa, L. Gil-Escrig, M. Sessolo, and H. J. Bolink, “Fully Vacuum-Processed Wide Band Gap Mixed-Halide Perovskite Solar Cells,” *ACS Energy Lett*, vol. 3, no. 1, pp. 214–219, Jan. 2018, doi: 10.1021/acsenergylett.7b01217.

APPENDIX

Dataset Containing the bandgap of various perovskite halides:

Number of Samples	Composition (ABX_3)	Bandgap (E_g)	References
1	$MASnIBr_2$	1.75	[205]
2	$MASnBr_3$	2.15	[205]
3	$CsSnI_3$	1.30	[206]
4	$CsSnBr_3$	1.70	[207]
5	Cs_2SnI6	1.30	[208]
6	Cs_2SnI_5Br	1.38	[208]
7	$Cs_2SnI_4Br_2$	1.40	[208]
8	$Cs_2SnI_2Br_4$	1.63	[208]
9	Cs_2SnIBr_5	2.36	[208]
10	Cs_2SnBr_6	2.85	[208]
11	$CsGeI_3$	1.63	[209]
12	$MA_3Bi_2I_9$	2.11	[210]
13	$Cs_3Bi_2I_9$	2.1	[211]
14	$AgBi_2I_7$	1.87	[103]
15	$Cs_2AgBiBr_6$	2.21	[212]
16	$Cs_3Bi_2I_9$	2.03	[211]
17	$Cs_3Sb_2I_9$	2.05	[213]
18	$Rb_3Sb_2I_9$	2.10	[214]
19	$MA_3Sb_2I_9$	2.14	[215]
20	MA_2CuCl_4	2.36	[216]
21	MA_2CuCl_4	2.48	[217]
22	$MA_2CuCl_2Br_2$	2.12	[217]
23	$MA_2CuClBr_3$	1.90	[217]
24	$MA_2CuCl_{0.5}Br_{3.5}$	1.80	[217]
25	$Cs_2CuSbCl_6$	1.66	[218]
26	$Cs_2AgSbCl_6$	2.60	[219]
27	$Cs_2AgBiCl_6$	2.77	[220]
28	$Cs_2AgSbBr_6$	1.64	[221]
29	$Cs_2AgTiCl_6$	2.00	[222]
30	$Cs_2AgTiBr_6$	0.95	[222]
31	$Cs_2LiGaBr_6$	1.96	[223]
32	$Cs_2NaGaBr_6$	1.76	[224]
33	$Cs_2AgInCl_6$	3.33	[225]
34	Cs_2AgBiI_6	1.75	[226]
35	Cs_2TiBr_6	1.80	[227]

Number of Samples	Composition (ABX_3)	Bandgap (E_g)	References
36	$MA_2CuCl_2I_2$	1.99	[216]
37	$CsSnBr_3$	1.79	[228]
38	$CsPb_{0.96}Bi_{0.04}I_3$	1.56	[229]
39	$CsPbI_2Br$	1.91	[230]
40	$CsPbI_3$	1.73	[231]
41	$Cs_{0.925}K_{0.075}PbI_2Br$	1.90	[232]
42	$CsPbI_2Br$	1.92	[233]
43	$CsPbI_2Br$	1.90	[234]
44	$CsPb_{0.98}Sr_{0.02}I_2Br$	1.89	[235]
45	$CsPbIBr_2$	2.05	[236]
46	$CsPb_{0.9}Sn_{0.1}IBr_2$	1.79	[237]
47	$MAPbI_3$	1.55	[191][238][239]
48	$MAPbBr_3$	2.20	[240]
49	$FAPbI_3$	1.52	[241]
50	$FAPbBr_{0.5}I_{2.5}$	1.97	[241]
51	$FAPbBrI_2$	2.27	[241]
52	$FAPbBr_{1.5}I_{1.5}$	1.91	[241]
53	$FAPbBr_2I$	1.97	[241]
54	$FAPbBr_{2.5}I_{0.5}$	2.10	[241]
55	$FAPbBr_3$	2.27	[241]
56	Rb_2AgBiI_6	1.98	[242]
57	$FA_{0.25}MA_{0.75}SnI_3$	1.28	[204]
58	$FA_{0.50}MA_{0.50}SnI_3$	1.33	[204]
59	$FA_{0.75}MA_{0.25}SnI_3$	1.33	[204]
60	$FA_{0.8}Cs_{0.2}SnI_3$	1.51	[243]
61	$GA_{0.14}MA_{0.86}PbI_3$	1.57	[244]
62	$MAPbI_{2.7}Cl_{0.3}$	1.59	[245]
63	$MAPbI_{1.5}Br_{1.5}$	1.87	[246]
64	$MAPbI_{2.4}Br_{0.6}$	1.72	[246]
65	$FA_{1/6}MA_{5/6}PbI_3$	1.59	[241]
66	$FA_{1/6}MA_{5/6}PbBr_{0.5}I_{2.5}$	1.63	[241]
67	$FA_{1/6}MA_{5/6}PbBr_1I_2$	1.76	[241]
68	$FA_{1/6}MA_{5/6}PbB_{r1.5}I_{1.5}$	1.86	[241]
69	$FA_{1/6}MA_{5/6}PbBr_2I1$	1.99	[241]
70	$FA_{1/6}MA_{5/6}PbBr_{2.5}I_{0.5}$	2.15	[241]
71	$FA_{1/6}MA_{5/6}PbBr_3$	2.30	[241]
72	$FA_{2/6}MA_{4/6}PbI_3$	1.56	[241]
73	$FA_{2/6}MA_{4/6}PbBr_{0.5}I_{2.5}$	1.67	[241]
74	$FA_{2/6}MA_{4/6}PbBrI_2$	1.76	[241]

Number of Samples	Composition (ABX_3)	Bandgap (E_g)	References
75	$FA_{2/6}MA_{4/6}PbBr_{1.5}I_{1.5}$	1.85	[241]
76	$FA_{2/6}MA_{4/6}PbBr_2I$	1.97	[241]
77	$FA_{2/6}MA_{4/6}PbBr_{2.5}I_{0.5}$	2.14	[241]
78	$FA_{2/6}MA_{4/6}PbBr_3$	2.29	[241]
79	$FA_{3/6}MA_{3/6}PbI_3$	1.55	[241]
80	$FA_{3/6}MA_{3/6}PbBr0.5I_{2.5}$	1.61	[241]
81	$FA_{3/6}MA3/6PbBrI_2$	1.75	[241]
82	$FA_{3/6}MA_{3/6}PbBr_{1.5}I_{1.5}$	1.85	[241]
83	$FA_{3/6}MA_{3/6}PbBr_2I$	1.96	[241]
84	$FA_{3/6}MA_{3/6}PbBr_{2.5}I_{0.5}$	2.14	[241]
85	$FA_{3/6}MA_{3/6}PbBr_3$	2.28	[241]
86	$FA_{4/6}MA_{2/6}PbI_3$	1.53	[241]
87	$FA_{4/6}MA_{2/6}PbBr_{0.5}I_{2.5}$	1.64	[241]
88	$FA_{4/6}MA_{2/6}PbBrI_2$	1.68	[241]
89	$FA_{4/6MA2/6}PbBr_{1.5}I_{1.5}$	1.86	[241]
90	$FA_{4/6}MA_{2/6}PbBr_2I$	1.97	[241]
91	$FA_{4/6}MA_{2/6}PbBr_{2.5}I_{0.5}$	2.11	[241]
92	$FA_{4/6}MA_{2/6}PbBr_3$	2.28	[241]
93	$FA_{5/6}MA_{1/6}PbI_3$	1.53	[241]
94	$FA_{5/6}MA_{1/6}PbBr_{0.5}I_{2.5}$	1.63	[241]
95	$FA_{5/6}MA_{1/6}PbBrI_2$	1.68	[241]
96	$FA_{5/6}MA_{1/6}PbBr_{1.5}I_{1.5}$	1.87	[241]
97	$FA_{5/6}MA_{1/6}PbBr_2I$	1.98	[241]
98	$FA_{5/6}MA_{1/6}PbBr_{2.5}I_{0.5}$	2.11	[241]
99	$FA_{5/6}MA_{1/6}PbBr_3$	2.28	[241]