Article

Development of Semi-Empirical and Machine Learning Models for Photoelectrochemical Cells

Niranjan Sunderraj [1], Shankar Raman Dhanushkodi [1,*], Ramesh Kumar Chidambaram [2],
Bohdan Węglowski [3,*], Dorota Skrzyniowska [3], Mathias Schmid [4] and Michael William Fowler [5,*]

1 Dhanushkodi Research Group, Department of Chemical Engineering, Vellore Institute of Technology, Vellore 632014, India
2 Automotive Research Center, School of Mechanical Engineering, Vellore Institute of Technology, Vellore 632014, India
3 Institute of Thermal Power Engineering, Cracow University of Technology, 31-864 Cracow, Poland
4 ZHAW School of Engineering, ICP—Institute of Computational Physics, Technikumstrasse 71, CH-8401 Winterthur, Switzerland
5 Department of Chemical Engineering, University of Waterloo, Waterloo, ON N2L3G1, Canada
* Correspondence: shankarraman.d@vit.ac.in (S.R.D.); bohdan.weglowski@pk.edu.pl (B.W.); mfowler@uwaterloo.ca (M.W.F.)

Abstract: We introduce a theoretical model for the photocurrent-voltage (I-V) characteristics designed to elucidate the interfacial phenomena in photoelectrochemical cells (PECs). This model investigates the sources of voltage losses and the distribution of photocurrent across the semiconductor–electrolyte interface (SEI). It calculates the whole exchange current parameter to derive cell polarization data at the SEI and visualizes the potential drop across n-type cells. The I-V model's simulation outcomes are utilized to distinguish between the impacts of bulk recombination and space charge region (SCR) recombination within semiconductor cells. Furthermore, we develop an advanced deep neural network model to analyze the electron–hole transfer dynamics using the I-V characteristic curve. The model's robustness is evaluated and validated with real-time experimental data, demonstrating a high degree of concordance with observed results.

Keywords: photochemical cell; space charge width; recombination; I-V model and electron hole transfer

1. Introduction

Hydrogen is a clean and versatile energy carrier, pivotal for generating heat and power in industrial applications. The adoption of hydrogen-powered fuel cell vehicles (FCVs) is rapidly accelerating worldwide, largely due to their remarkable efficiency and durability. Despite this, the widespread commercialization of FCVs is constrained by the high costs associated with hydrogen production methods, such as steam electrolysis and steam or natural gas reforming. Water electrolysis offers a promising alternative, demonstrating 70% efficiency with zero greenhouse gas emissions. Since 2010, the volume of hydrogen produced via electrolysis has doubled, leading to an increased integration of renewable energy sources (RES) with water electrolyzer cells for micro grid energy storage. RES electricity is converted into hydrogen gas by these cells and stored in metal hydride reservoirs. This method offers a viable strategy for managing smart grids, providing stability and security. However, hydrogen storage systems, while enhancing smart grid reliability, are not yet scaled to meet local electricity-gas grid requirements. Additionally, the electrolysis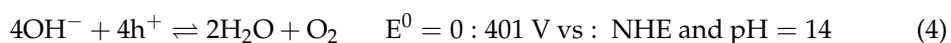 process remains costly and energy-intensive, with significant component degradation occurring during peak electricity demand periods. To address these challenges, adopting more cost-effective technologies, such as photoelectrochemical cells (PECs), could provide a viable alternative for hydrogen production. PEC systems can be integrated into "Power to Gas" grids, where hydrogen is injected into gas pipelines and transported to energy storage facilities. The components of PEC absorb the photons in the sunlight to

produce energy. The cell comprises light-sensitive nano-materials such as n- or p-type semiconductors to absorb photons. Figure 1 depicts the schematic of the cell. The incident photons irradiate the surface of the cell and form electron–hole pairs at space charge regions. The holes are separated from the electron by the electric field at the surface of the semiconductor. When the electrons depart toward the external circuit, the holes are pushed into the SEI of the cell [1]. In an acidic environment, the electrons flow via the external circuit and reach the interface of the counter electrode (CE). These electrons react with H+ ions and form hydrogen gas at the CE. The holes at the SEI of the photo-electrode react with water to form oxygen species [1]. The reaction steps of the cell are given in Equations (1) and (2).

$$4H^+ + 4e^- \rightleftharpoons 2H_2 \qquad E0 = 0 \text{ V vs}: \text{ NHE and pH} = 0 \tag{1}$$

$$2H_2O + 4h^+ \rightleftharpoons 4H^+ + O_2 \qquad E0 = +1.23 \text{ V vs}: \text{ NHE and pH} = 0 \tag{2}$$
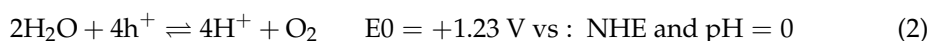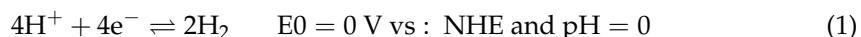
In a basic environment, the holes generated at the semiconductor react with the hydroxyl ions to form oxygen while the electrons in the CE decompose water into hydrogen and hydroxyl ions [1]. Equations (3) and (4) provide the reactions associated with the water splitting.

$$4H_2O + 4e^- \rightleftharpoons 2H_2 + 4OH^- \qquad E^0 = -0:828 \text{ V vs}: \text{ NHE and pH} = 14 \tag{3}$$

$$4OH^- + 4h^+ \rightleftharpoons 2H_2O + O_2 \qquad E^0 = 0:401 \text{ V vs}: \text{ NHE and pH} = 14 \tag{4}$$
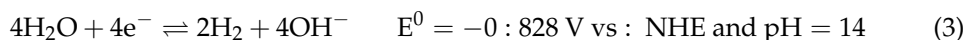
FeO, also known as wüstite, exhibits significant potential for enhancing the performance of photoelectrochemical (PEC) cells, particularly in terms of water splitting efficiency and solar energy conversion. Its band gap, which falls within the visible spectrum, facilitates effective sunlight absorption, and the abundant availability of iron makes FeO a cost-efficient material. The material demonstrates substantial photocatalytic activity, which is crucial for driving the redox reactions in PEC processes. However, FeO encounters several challenges, including instability under specific conditions, limited efficiency, and high synthesis costs. Researchers are addressing these issues through various approaches, such as doping with elements like cobalt or nickel to enhance electronic properties, creating composites with materials like TiO_2 to improve stability and catalytic activity, and developing nanostructures to increase surface area and efficiency. Despite these challenges, ongoing research continues to explore and refine FeO's potential, positioning it as a promising candidate for advanced PEC technologies. Detailed properties and performance metrics of the materials utilized in PECs are summarized in Table 1.

The hydrogen production rate in PEC cells is critically dependent on the stability of the semiconductor–electrolyte interface (SEI), which is comprised of three distinct regions: the bulk semiconductor, the space charge region (SCR), and the Helmholtz region, as illustrated in Figure 2. Understanding the interfacial interactions, mapping the photocurrent, and assessing the SEI thickness in relation to cell performance remain complex and underexplored. Accurate modeling of these layers requires a precise mathematical formulation of the SEI. Mapping the voltage drop across the SEI provides insights into the transport phenomena of hole–electron pairs and the hydrogen production rate. Key parameters such as surface charge width and band gap size are crucial for characterizing hole–pair dynamics and quantifying the rate of electron migration from the interface to the external circuit. Developing a predictive model for transport and voltage loss breakdown within the SEI is challenging, requiring detailed estimates of the cell's transport and physicochemical properties.

Table 1. Properties and performance of materials for photoelectrochemical cells.

| Material	Properties	Voltage (V)	Thermodynamic Properties	Synthesis Method	Electrolyte	Efficiency (%)	Kinetic Properties
Titanium Dioxide (TiO_2)	- Strong photocatalytic activity under UV light - Chemically stable- High stability and durability - Non - toxic - Wide band gap	2.7–3.2	- High band gap energy - Low electron affinity	Sol-gel method, hydrothermal method	Aqueous electrolyte (e.g., Na_2SO_4, KNO_3)	3–5	- Slow charge carrier dynamics - Low recombination rates under UV light
Doped TiO_2 (with Mo, Cr, Nb, etc.)	- Enhances properties based on dopant type - Improved photocatalytic efficiency - Enhanced light absorption - Varying physico-chemical properties depending on dopant	Varies with dopant	Varies with dopant	Sol-gel method, hydrothermal method	Aqueous electrolyte (e.g., Na_2SO_4, KNO_3)	5–10	- Improved charge carrier dynamics - Reduced electron–hole recombination
Molybdenum (Mo)	- Good electrical conductivity - High stability - Resistant to corrosion	2.5–3.0	- Moderate band gap energy - Moderate electron affinity	Chemical vapor deposition (CVD), sputtering	Aqueous electrolyte (e.g., Na_2SO_4, KNO_3)	6–8	- Faster charge carrier transfer - Reduced recombination rates
Chromium (Cr)	- Good electrical conductivity - High stability- Resistant to corrosion - Resistant to corrosion	2.3–2.8	- Moderate band gap energy - Moderate electron affinity	CVD, sputtering	Aqueous electrolyte (e.g., Na_2SO_4, KNO_3)	7–9	- Intermediate charge carrier dynamics - Moderate recombination rates
Niobium (Nb)	- Good electrical conductivity - High stability - Resistant to corrosion	2.6–3.1	- Moderate band gap energy - Moderate electron affinity	CVD, sputtering	Aqueous electrolyte (e.g., Na_2SO_4, KNO_3)	8–10	- Enhanced charge separation - Reduced recombination rates
Silicon (Si)	- High electrical conductivity - Good thermal conductivity - High mechanical strength	0.8–1.1	- Low band gap energy - High electron affinity	CVD, sputtering	Aqueous electrolyte (e.g., Na_2SO_4, KNO_3)	15–20	- High charge carrier mobility - Reduced recombination rates
Gallium Arsenide (GaAs)	- High electrical conductivity - Good thermal conductivity - High mechanical strength	1.2–1.4	- Low band gap energy - High electron affinity	CVD, molecular beam epitaxy (MBE)	Aqueous electrolyte (e.g., Na_2SO_4, KNO_3)	20–25	- High charge carrier mobility - Low recombination rates
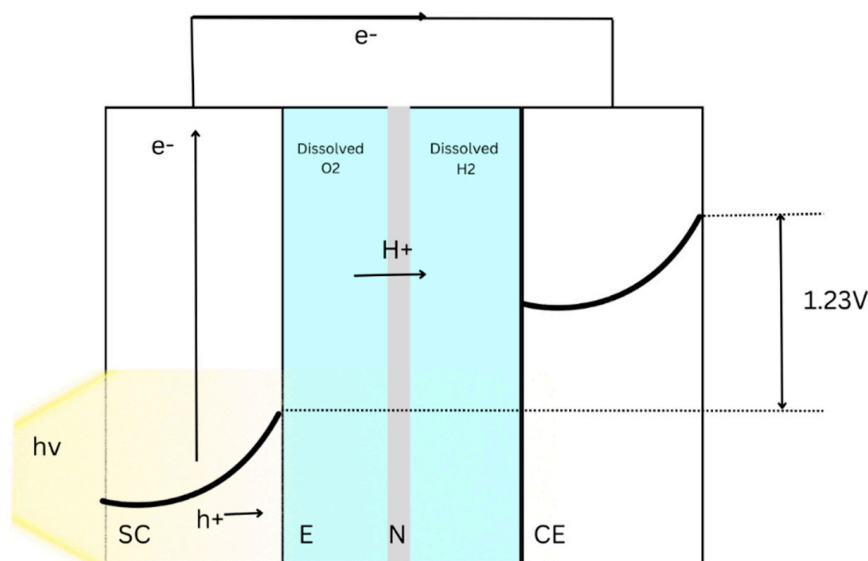
Figure 1. Schematic of a PEC. Figure concept adopted from [2,3] and redrawn.
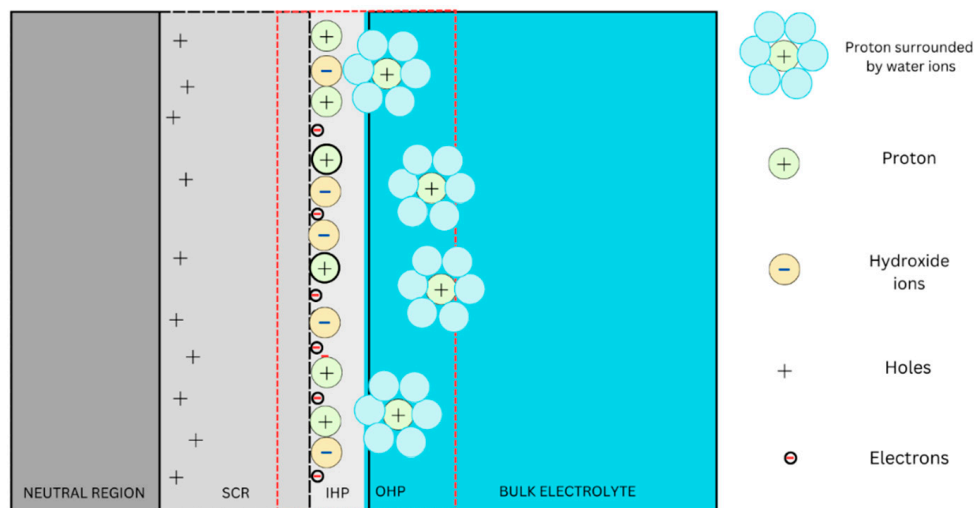
Figure 2. Schematic of the semiconductor–electrolyte interface, SCR: space charge region, IHP: inner Helmholtz plane, OHP: outer Helmholtz plane. Figure concept adopted from [2,3] and redrawn.

Figure 2 presents a schematic of the SEI, highlighting the interactions among electrons, donor ions containing holes, and ions in the electrolyte. Empirical models documented in the literature elucidate the roles of reverse current, bias, and recombination parameters at the SEI [4,5]. The precision of any model designed to analyze voltage loss across the SEI depends on the quality of the photocurrent and potential data. While extensive historical data have been used to develop interface models for forecasting, control, and optimization, these models often suffer from diminished spatial accuracy due to their reduced dimensionality. Additionally, modeling the semiconductor–electrolytic behavior and potential losses in the cell is inherently complex. Zero-dimensional (0-D) models established in the literature outline relationships between cell components and transport barriers at the semiconductor. Gartner's theoretical equations estimate the photocurrent generated by the photoelectrode under reverse bias [4], but this model is constrained by the Schottky barrier and issues related to ion transfer at the SEI, and it does not address recombination challenges. Reichman improved the 0-D model by providing equations to compute the potential drop across the

bulk and SCR recombination regions at the SEI [5], using Equations (5)–(7) to calculate the current due to minority carriers (Jp).

$$J_p = \frac{J_g - J_o e^{\frac{qV}{kT}}}{\left(1 + \frac{J_o e^{\frac{qV}{kT}}}{I_p{}^o}\right)} \tag{5}$$

$$J_o = \frac{qp_o L_p}{\zeta} \tag{6}$$

$$p_o = \frac{n_i{}^2}{N_d} \tag{7}$$

where n_i is the intrinsic carrier density (m^{-3}), N_d is the carrier density (m^{-3}), J_g corresponds to current density in the Gartner model (A m^{-2}), q corresponds to elementary charge (C), I_0 is the photon flux, (A-m^{-2}) $I_n{}^o$ and $I_p{}^o$ are the electron and hole exchange current parameters (A m^{-2}), W is the space charge width (m), L_p is the hole diffusion length (m), α is the optical absorption coefficient of the semiconductor (m^{-1}), J_o is the saturation current density (A m^{-2}), and ζ is the hole lifetime (s).

Unlike zero-dimensional (0-D) models, one-dimensional (1-D) models represent the PEC system along a single spatial axis. Wenger et al. [6] have developed 1-D models for dye-sensitized cells (DSCs), where optical and electrical equations are coupled to capture the behavior of DSCs. Their model provides an accurate depiction of how excited electrons convert incident photons into electricity [7]. In DSCs, a photosensitive dye is absorbed into a wide band gap semiconductor to facilitate photocurrent generation. Consequently, the dye absorption rate, derived from the optical model, is crucial for determining the steady-state dye electron injection efficiency and hole diffusion length [6]. Achieving a balance between model complexity and computational efficiency is challenging for both 0-D and 1-D PEC models. As a result, two-dimensional (2-D) models have been reported by Hernandez et al. [3], who analyzed oxygen bubbles covering the electrode by approximating the inner surface layer as a two-dimensional flat surface. Giacoppo et al. have introduced three-dimensional (3-D) models to provide a more accurate representation of the PEC system. However, their model's convergence is limited by computational complexity [8]. Based on Gartner's equations, their 3-D model demonstrated that the cell achieved 10 mA cm^{-2} at a bias voltage of 1 V, although recombination losses were not included in their model. Despite extensive studies on DSCs and microscopic interfacial behavior in PECs, focused research on modeling the semiconductor–electrolyte interface (SEI) and probing recombination mechanisms in PEC nanomaterials remains limited. Reichman's 0-D models (Equations (5)–(7)) approximate experimental polarization plots well, showing that recombination in the space charge region (SCR) and charge transfer across the SEI depend on current exchange parameters. These parameters are influenced by (a) the overlap in semiconductor bands, (b) interface states with redox ions, and (c) charge transfer transition probabilities [5]. However, exchange current parameters for n- or p-type materials are not readily available for scaling models to new photoelectrode nanomaterials, and efforts to estimate these parameters for nanomaterials are lacking.

Operating strategies for PECs are determined by integrating theoretical models with experimental data. Simplified semi-empirical models in the literature often use linear regression to achieve correlations among variables. This approach is insufficient for diagnosing cause–effect mechanisms within PEC components. Increasing volumes of data generated from PEC cells are not fully utilized to assess cell durability and performance. A model capable of (a) revealing electrochemical parameters, (b) classifying a large number of material properties to enhance learning, (c) performing iterative tasks through repeated interactions with the I-V dataset in a dynamic environment, and (d) exploring extensive I-V datasets to identify features and patterns in PEC electrodes would be invaluable for setting design metrics for next-generation cells. Identifying hidden patterns or features in the I-V dataset

could provide new strategies and operating protocols for PECs. Currently, no such model exists in the literature. Modern quantum chemistry methods also fall short in predicting complex characteristics such as diurnal variations in cell performance. Machine learning (ML) models could address these research needs by resolving performance decay related to diurnal variations and establishing structure–property–performance relationships for the cell. Integrating experimental data with ML techniques could uncover operating and failure modes, providing common patterns, specific predicted electrochemical properties, and strategies for real-world testing.

Predictive maintenance of PECs depends on the durability and I-V datasets. Machine learning and deep learning models are essential for understanding the non-linearity in I-V data collected from the cell. Deep learning models can process a wide range of datasets and predict I-V characteristics more accurately than traditional machine learning or 0-D models. They are capable of ingesting and processing datasets through multiple iterations to learn non-linear features effectively. Advancements in algorithm development, computational efficiency, and database creation of PEC parameters could significantly enhance the application of Deep Learning in Simulation (DLS) for the cell. Deep learning models can handle data noise and make precise predictions based on historical data. They can capture complex trends between photocurrent, voltage, and key intrinsic parameters in PECs across different nanomaterials. Wang et al. [9] developed a machine learning model to predict the effect of dopants on photoelectrode. Oral et al. [10] implemented Random Forest, Decision Tree, and Association Rule Mining to predict band gap and photocurrent density, using thirty-three features including electrode materials, synthesis methods, irradiation properties, and electrolytes. Kharade et al. reported an Artificial Neural Network (ANN) model predicting adsorption efficiency in DSCs with a three-layer approach, considering parameters such as fill factor, short-circuit current, and open-circuit voltage [11]. Despite existing ML models for predicting photocurrent using feature and voltage datasets, they require significant computational power and time. To effectively capture real-time cell events while minimizing the computational demands of 0-D modeling, a robust model is absolutely essential. This method is distinguished by its ability to seamlessly integrate detailed cell dynamics with efficient computational techniques, enabling rapid analysis and valuable insights. When compared with the existing literature, the results highlight the paramount importance of this approach in enhancing the understanding of cell behavior and material interactions. Through the combination of a 0-D model with DLS algorithms, we can significantly expedite the development of new diagnostic tools for PECs, showcasing a more efficient and impactful approach to research and innovation in this field.

Based on the current state of research, this work aims to achieve the following objectives: (a) Develop a theoretical model to assess the current density across the SEI for two different PEC components; (b) Approximate the hole exchange current parameter (I_{po}); (c) Evaluate the photocurrent across the SEI; (d) Demonstrate voltage loss breakdown in relation to cell parameters for gallium arsenide and pristine silicon cells; and (e) Develop an ANN model to predict photocurrent for the cell. This work will provide a novel framework that integrates both analytical and machine learning models to determine parameters such as space charge width and photocurrent.

2. Materials and Methods

In Section 2.1, we focus on formulating equations for the I-V model. We present a derivation to estimate the hole current parameter and the photocurrent density across the SEI. Section 2.2 focuses on developing an ANN-based DLS model to predict the photocurrent density across the SEI. Table 2 provides key properties for photoelectrochemical materials, including FeO, GaAs, Si, and doped TiO_2, used in the model.

Table 2. Comparison of key properties for photoelectrochemical materials: FeO, GaAs, Si, and doped TiO$_2$.

Property	FeO	GaAs	Si	Doped TiO$_2$ (with Mo, Cr, Nb)
Band gap (eV)	1.3–1.5	1.43	1.12	Varies with dopant (e.g., 3.0–3.2 for undoped TiO$_2$)
Absorption Coefficient	Moderate	High	High	Improved with doping (e.g., higher for Mo-doped TiO$_2$)
Carrier Mobility	Moderate	High	Very high	Improved with doping (e.g., higher for Nb-doped TiO$_2$)
Carrier Lifetime	Moderate	High	Very high	Improved with doping (e.g., longer for Cr-doped TiO$_2$)
Crystal Structure	Rock salt	Zincblende	Diamond	Varies with dopant (e.g., rutile for undoped TiO$_2$)
Chemical Stability	Moderate	High	High	High (generally)

2.1. I-V Model

The voltage loss in the cell depends on the stability of the band gap, number of photons collected by the semiconductor, recombination occurring in the photoelectrode, and the performance of the individual cell components. The photo-oxidation of the electrode materials and their reaction with water molecules is limited by the presence of the surface-trapped holes at the surface. The voltage loss in the cell is contributed by the rate anodic decomposition and cathodic photo-reduction reactions. Understanding the role of the layers in the SEI such as SCR and inner and outer Helmholtz layers is the first step in identifying the voltage loss contributors. The SCR is the first region in the SEI where the electrons in the semiconductor and the ions in the bulk regions form the electric fields. The optimization of the SCR width and the electric charge in this region is significant to improve overall efficiency of the cell. The region formed between the ions (protons and hydroxide ions) adsorbed in the semiconductor and the solvated ions in the electrolyte is called the Helmholtz region. The adsorption and desorption of protons and hydroxide ions occurs continuously at the interface. The polarization of the SEI depends on the thickness of these three SEI layers [12]. Two main drivers for the photocurrent generation are the width of the space charge recombination and the distance between the inner and outer Helmholtz layers. If the lattice structure of the photoelectrode is abnormally terminated or not optimized, electronic states will be formed within its band gap. These states are filled with the electrons supplied from the bulk region of the semiconductor. The ions originate from the bulk free electrons regions create positive space charge regions [1]. Poisson's law is used to model the charge in the SCR and the space charge width. It correlates the potential and the net amount of charge produced at the interface. The width of the SCR is computed according to Equation (8).

$$W = \sqrt{\frac{2\epsilon_r \epsilon_0 \left(V - \frac{kT}{q} \right)}{q N_d}} \tag{8}$$

where W is the width of the SCR (m), ϵ_r is the relative permittivity of the semiconductor material (unitless), ϵ_0 is the electric constant, V is the voltage across the SCR(V), k is the Boltzmann Constant (JK^{-1}), e is the elementary charge (C), and T is the temperature (K).

The charge accumulated in the SCR is computed using Equation (9).

$$Q_{SC} = \sqrt{2\epsilon_r \epsilon_0 e N_d A^2 \left(V - \frac{kT}{q} \right)} \tag{9}$$

where A is the surface area of the semiconductor (m^2).

The pH value at which the net adsorbed charge is zero is termed as the point of zero charge (PZC). In this region, both the inner and the outer Helmholtz layer have adsorbed

and solvated ions, respectively, at the electrolyte surface (Figure 2). The potential drop across the Helmholtz region is calculated according to Equation (10).

$$V_H = \frac{2.3kT}{q}(PZC - pH) \tag{10}$$

where V_H Helmholtz potential is a key parameter in our analysis. When dopants are used in the photo-electrodes, the charge and the width of the space charge region (SCR) vary. The carrier concentrations of doped species can be found in reference [13]. To evaluate the voltage drops across a TiO_2-doped semiconductor, we use a provided value of the point of zero charge (PZC) as 5.6, which can be found elsewhere [14]. The assumptions employed in the derived model are as follows: (a) The charge transfer across the solid electrolyte interface (SEI) is due to the movement of holes and electrons generating an electric current. (b) The current density across the SEI is a function of the exchange current parameter. (c) The ratio of concentration of holes and electrons across the interface is calculated based on certain assumptions (cf. [4]). (d) The width of the Helmholtz layer is assumed to be small. (e) The voltage across the Helmholtz layer is assumed to be constant. (f) Ion concentrations at the interface are assumed to be constant. (g) No mass transfer effect is considered. (h) No concentration polarization is taken into account. (i) Charge transfer reactions are assumed to be first order with respect to minority and majority carriers at the interface. To calculate the current density due to holes (Jp), we substitute it with the total current density from the performance curves in the semi-empirical formula for the hole exchange current parameter, as isolating the values of Jp requires highly sophisticated equipment. The current density due to minority carriers (holes) [5] is computed using Equation (11).

$$J_p = I_p{}^o\left(\frac{p_s}{P_{so}} - 1\right) \tag{11}$$

where J_p is the current density driven by holes (A m^{-2}), p_s is the hole concentration at the interface (m^{-3}), and p_{so} is the equilibrium hole concentration at the interface (m^{-3}). The hole carrier density across the interface is obtained by solving the diffusion equation for minority carriers in the neutral region of the semiconductor.

$$D\frac{d^2p}{dx^2} - \frac{p - p_o}{\varsigma} + I_o\alpha e^{-\alpha x} = 0 \tag{12}$$

where D is the diffusion coefficient (m^2/s), τ is the hole lifetime (s), p_eq is the equilibrium hole density (m^{-3}), and G is the monochromatic photon flux (A/m^2) incident on the semiconductor after accounting for loss corrections. The correction factor accounts for interface reflections and electrolyte absorption, and α is the optical absorption coefficient (m^{-1}). The first term in Equation (12) represents the change in the hole density across the x-axis inside the semiconductor, which depends on the diffusion of holes at the semiconductor surface. The second term refers to the recombination of holes driven by the completion of their lifetimes. The third term represents the formation of new holes due to incident monochromatic photon flux. The Gartner model assumes that there is no recombination in the SCR or the SEI. Therefore, the hole concentration at the SCR is zero. Equation (13) presents the Gartner model used to evaluate Jg. Equation (13) can be considered as a solution derived from Equation (12).

$$J_g = J_o + qI_o\left(1 - \frac{e^{-\alpha W}}{1 + \alpha L_p}\right) \tag{13}$$

The SCR width can be approximated using Equation (14).

$$W = \sqrt{\frac{2\epsilon_r\epsilon_o\phi_b}{qN_d}} \tag{14}$$

where $\phi_b = \phi_{b0} - V$, ϕ_{b0} is the flat-band potential of the semiconductor, and V is the voltage across the SCR. The current density due to electron flow is given in Equation (15). The Reichman model (Equation (5)) considers the increasing rates of recombination in the neutral region and neglects the recombination effect in the SCR and the interface. Equations (5) and (15) describe the hole carrier current density (Equation (5)) and the current density measured due to flow of electrons, respectively (Equation (15)). By combining these two equations, the polarization plot for neutral recombination regions can be obtained.

$$J_n = -I_n{}^o \left(e^{\frac{qV}{kT}} - 1 \right) \tag{15}$$

The recombination current is expressed using K and the ratio of hole concentration and the equilibrium hole concentration at the SEI. Reichman studied the recombination in the SCR using a method proposed by Sah et al. [15] According to their model, the recombination current depends on the parameters A, B, and K. Equations (16)–(20) summarizes the modified method adopted in this study based on [5,15].

$$J_{SCR}{}^R = K \sqrt{\left(\frac{p_s}{p_{so}} \right)} \tag{16}$$

$$A = I_p{}^o + J_o e^{\frac{qV}{kT}} \tag{17}$$

$$\frac{p_s}{p_{s0}} = \left(\frac{\left\{ -K + \sqrt{(K^2 + 4AB)} \right\}}{2A} \right)^2 \tag{18}$$

$$B = I_p{}^o + J_g \tag{19}$$

$$K = \pi k T n_i W e^{\frac{qV}{2kT}} \tag{20}$$

where $J_{SCR}{}^R$ is the recombination current (A m^{-2}), T is the temperature of operation (K), and W is the SCR width (m) (Equation (14)). Since no hole exchange current parameter values $I_p{}^o$ are available for the PEC components, we derive an analytical expression to estimate the $I_p{}^o$ by solving Equations (9) and (16). The assumptions considered in the derivation are given below. $I_p{}^o >> J_o e^{\frac{qV}{kT}}$, $J_g >> I_p{}^o$, and $I_p{}^o >>$ z

The current exchange parameter values found in the literature are significantly lower than the current density according to the Gartner model and higher than the saturation current density with the multiplied exponential, as assumed initially. The hole exchange current parameter values are selected from the J-V curves, meeting all the assumptions for further analysis in the model, justifying the assumptions made. The data points from performance curves, which yield total photocurrent density, are used to calculate the value of all instances. Obtaining the values of hole current density requires highly sophisticated equipment. The mean of all values satisfying the above assumptions is further used for calculation using Equations (5) and (15). The derivation for the hole exchange current parameter (Equation (21)) is provided in Appendix A.

$$I_p{}^o = \frac{-K\sqrt{J_g} + \sqrt{J_g K^2 + 8(J_p - J_g) K^2}}{4(J_p - J_g)} \tag{21}$$

2.2. DLS Model

The DLS approach is used to predict the photocurrent as it reduces the complexity in numerical calculations and provide the I-V curve for the PEC. It needs an Artificial Neural Network (ANN) algorithm which includes the data extraction, data cleaning, and pre-processing. The data used for training the model are obtained from [16]. The data were split into 80% for training and 20% for testing and are used in the DLS models. To prepare the dataset, we applied standard scaling to ensure that all features have a mean of zero and

unit variance. To focus on the most impactful features, feature selection is performed to eliminate those that were less significant. This process not only reduces the dataset size but also enhances model interpretability and efficiency, leading to shorter training times and a lower risk of overfitting. Following the methodology of Wang et al., we used Extra Trees for feature selection, given the large number of input features. This ensemble method creates multiple randomized decision trees from various data subsamples, each built using a random subset of features. The best feature for splitting the data is chosen based on criteria like the Gini Index. This randomness results in a collection of uncorrelated decision trees, each providing an importance score for the input features.

The klib library is employed to expunge and clean the data and assess its data quality. Repeated rows and empty columns are dropped in the data cleaning process. It reduces dimensionality and the memory required for further processing. We used the Standard Scalar method to normalize the data points. The treated datasets are normalized to enable the faster convergence. This is achieved by assessing the values of features to a common scale and an algorithm is developed that is less sensitive to the feature sizes. The model used to treat the dataset is given in Equation (22).

$$Z = \frac{X - \mu}{\sigma} \tag{22}$$

The normalized dataset is split into training and testing sets in an 80–20 ratio before invoking the regressor for training the model. The regressor is initialized with the number of nodes, type of kernel initializer, and activation functions. The ANN model used in DLS is given in Equation (23).

$$y = f(w_1 x_1 + w_2 x_2 + w_3 x_3 + \ldots + ..w_n x_n + b) \tag{23}$$

where x_1, x_2, …, x_n represent the inputs to the ANN model, w_1, w_2, …, w_n are the corresponding weights for each input, b represents the bias term, and f represents the activation function. The perceptron algorithm applies a weighted sum to the given inputs and a bias term is added. It passes the result through an activation function to produce an output. The ANN model used for this research consists of 2 hidden layers with 8 and 2 nodes in the first and second hidden layers, respectively. The model uses random uniform as the kernel initializer to ensure that the weights are initialized with a uniform range of values, preventing vanishing or exploding gradients during model training. If the neurons are initialized with weights having very small or large values, it makes it difficult for the network to learn and update the weights properly, leading to poor performance. Each neuron works in cohesion with a weight initialized by the initializer, a bias value, and the activation function. The activation function used for the ANN model is the rectified linear function (ReL), which is a piecewise linear function that ensures the qualities of linear regression. The model uses a stochastic gradient descent model (Adam) based on the adaptive estimations of lower order moments. This model calculates the error between the predicted and the training data and adjusts the values of the bias and the weight in the neurons to increase the model's efficiency. A regressor is used to predict the output features based on the training dataset, and the efficiency of the model is calculated. Mean squared error is the metric used for error estimation and model validation.

3. Results and Discussions

3.1. I-V Model

The I-V experimental plots for several semiconductor materials are widely recorded. However, quantifying the voltage loss contributed by the oxidation half-reaction on metal oxide surfaces poses challenges in predicting photocurrent-voltage characteristics for the cathodic and anodic reactions. An empirical model is required to compute the I-V simulation. The data points from experimental characteristics and the resulting photocurrent density are used to calculate the values for all materials assessed in this study (Table 3).

Since isolating the values of hole current density is not straightforward, the mean of all values calculated from the total current density is taken. Solving Equations (5) and (15) provides the magnitude of the photocurrent generated due to migration of holes across the interface, and the electron migration, respectively. The total photocurrent across the SEI is computed by summing the individual photocurrent densities evaluated from the equations. The hole current density across the SEI is compared with the experimental performance curve of pristine silicon to identify and validate trends between them. The variation of SCR width against the potential drop across the SEI between TiO_2 with different dopants is plotted in Figure 3a. The SCR width of Mo-doped TiO_2 is widest among the dopants, while Cr-doped TiO_2 has the lowest SCR width. The availability of excess valence electrons in the case of Cr-doped TiO_2 increases the carrier density, leading to an increase in charge and width of the SCR. The reduced electron availability in the case of Mo-doped TiO_2 leads to a weaker electric field, thus providing a larger SCR width. Figure 3b shows the variation in space charge width of hematite with changes in carrier concentration. The carrier density depends on the dopant and the amount of doping performed. Therefore, doping level is a deciding parameter in determining the SCR width. Transfer of free electrons from the bulk is the other factor that determines the formation of the SCR. Any increase in the carrier density results in a dramatic increase in the charge in the SCR (Figure 3c,d). The charges in the SCR and the width of the SCR show an inverse relation with each other. The changes in voltage across the Helmholtz layer as a function of the pH value of the electrolyte for anatase TiO_2 is analyzed in Figure 3e. A positive to negative shift in voltage is observed when the pH reaches 5.9. The shift in Helmholtz layer voltage occurs due to the decreasing number of protons (or increasing amounts of hydroxide ions) with increasing pH and the number of electrons available at the surface states of the semiconductor. An increase in pH leads to a subsequent increase in hydroxide ion concentration (or reduction in proton concentration) in the electrolyte. The increased concentration causes more adsorption of hydroxide ions than protons, causing a reversal of Helmholtz voltage. The performance of the cell is given in Figure 4. The details related to the energy conversion efficiency as a function of potential created by the illumination conditions and intrinsic properties of the cell, sign of the photocurrent (anodic or cathodic), onset potential, the limitation related to the electron–hole transport, and the transient effects can be inferred using the cell performance data. The changes observed in photocurrent density, after accounting for bulk recombination and the recombination current, are plotted against voltage in Figure 4a. The model prediction is validated with the data collected from the literature [5]. The Reichman equation does not emphasize the recombination current separately, but the 0-D model proposed in this study can provide the loss contributed by the recombination effect. We observe that the current density is decreasing due to an increase in bulk recombination, as inferred by following the decreasing voltage trend in the characteristic curve. The onset of bulk recombination is observed at 0.4V. SCR recombination current starts rising from 0.3V and reaches a maximum near 0.45V. The values of the low diffusion coefficient and the defects in the semiconductor lead to bulk recombination, hindering the flow of carriers between the electrodes and SEI. The minority carrier is building up at the SEI due to the rate-limiting effect and affects interface charge at higher voltage regions [1]. The total current density across the SEI obtained by summing the SCR and bulk recombination currents is given in Figure 4b. The effect of SCR recombination is more pronounced at higher voltage regions where the SCR recombination current opposes the generation of the photocurrent. The effect of SCR is consistent with the similar observations reported in Figure 4a. At 0.4 V, we observe an 80.78% reduction in the photocurrent density between the effects of bulk and the combined effects of bulk and SCR recombination. This reduction is caused by the reduction in SCR width, consequently causing the holes and electrons to come close together, leading to a higher probability of recombination in the SCR region.
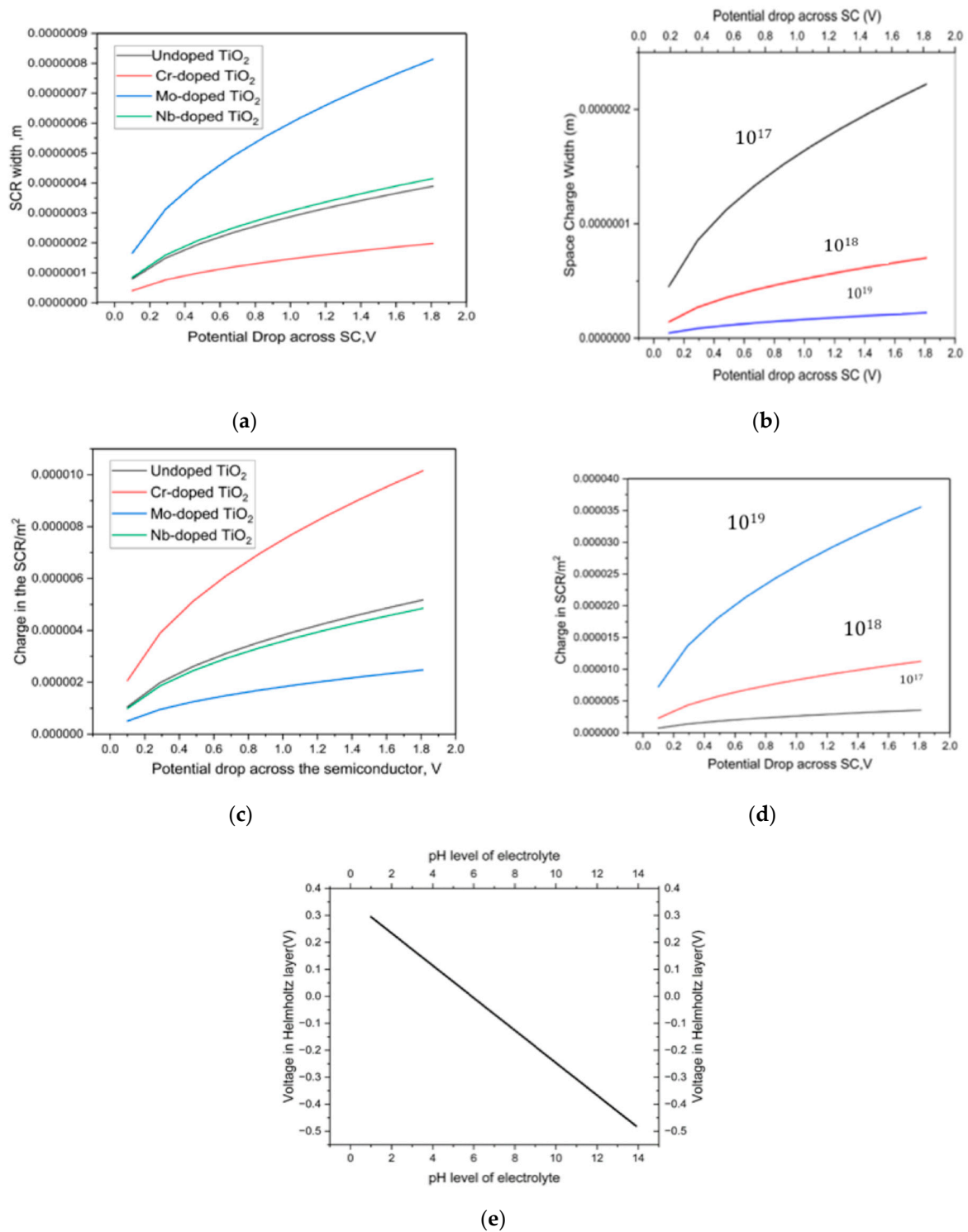
(a)

(b)

(c)

(d)

(e)

Figure 3. (**a**) Variation in space charge width between undoped and doped TiO$_2$-based photo-electrodes. (**b**) Variation of space charge width in hematite with different carrier concentrations. (**c**) Charge in the SCR region of hematite with different carrier concentrations. (**d**) Charge in the SCR of undoped and doped TiO$_2$ photoelectrodes. (**e**) Variation of voltage across the Helmholtz layer vs pH.

Table 3. Parameters used in the model [1,3–12,17–19].

Parameters	Silicon	Gallium Arsenide
Hole diffusion length	2×10^{-6} (m)	0.5×10^{-6} (m)
Carrier Concentration	10^{25} (m^{-3})	10^{22} (m^{-3})
Optical Absorption Coefficient	3.105×10^5 (m^{-1})	3×10^6 (m^{-1})
Intrinsic carrier density	9.65×10^{15} (m^{-3})	10^{13} (m^{-3})
Relative Permittivity	11.7	12
Flat-band Potential	-0.4 (V)	0.7 (V)
Hole Lifetime	1.3×10^{-4} (sec)	10^{-9} (sec)
Monochromatic illumination flux	1000 (Am^{-2})	10 (Am^{-2})
Hole exchange current parameter	Not available	10^{-4} (Am^{-2})
Electron exchange current parameter	Not available	10^{-9} (Am^{-2})

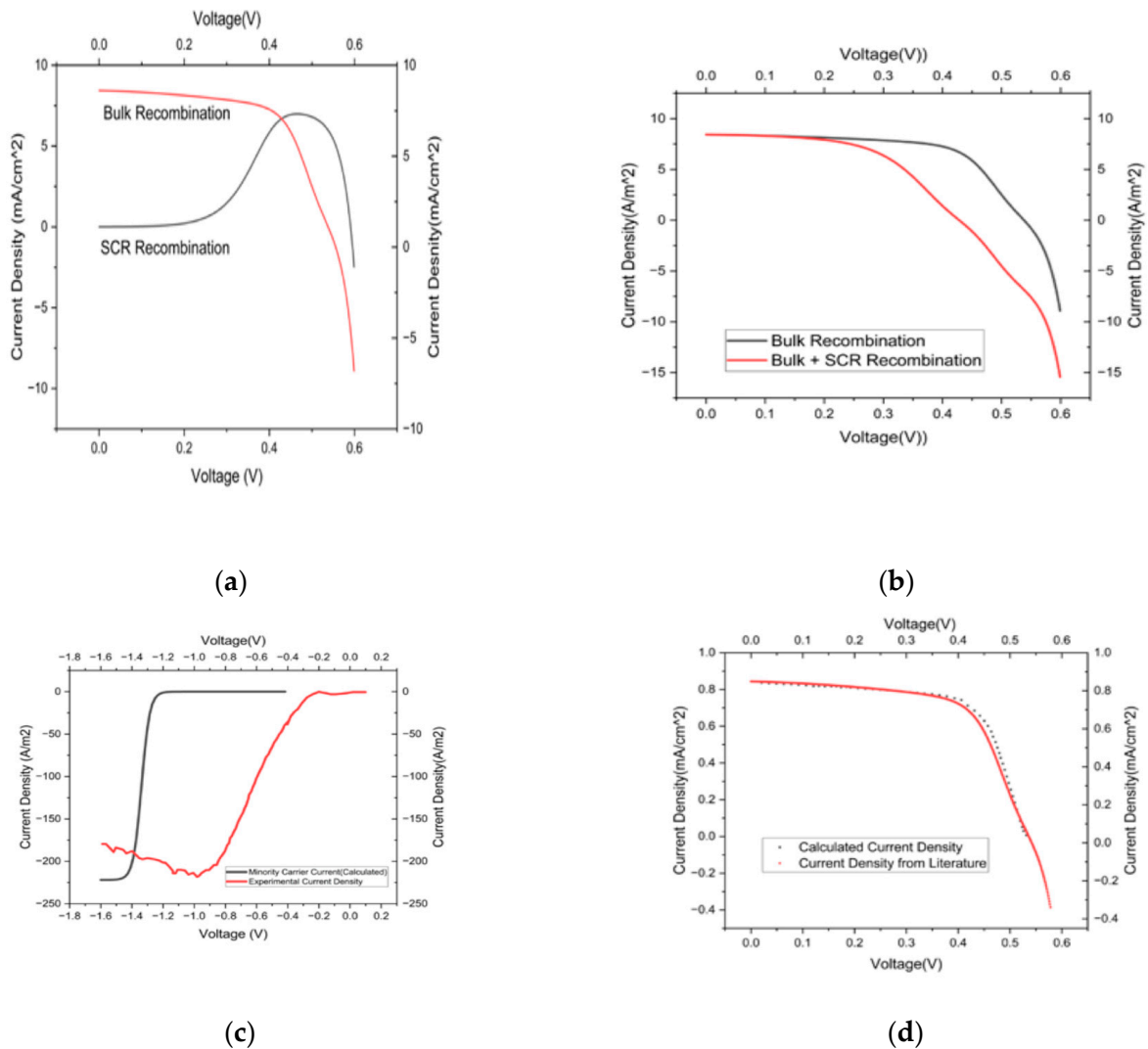

(**a**)

(**b**)

(**c**)

(**d**)

Figure 4. (**a**) Calculated current density with bulk recombination vs. voltage and SCR recombination current density vs. voltage across the interface for gallium arsenide. (**b**) Photocurrent density vs. voltage with bulk recombination and a combination of bulk and SCR recombination. (**c**) Comparison of calculated current density and experimental current density of pristine silicon with $I_p{}^0$ value as 1.2563×10^{-20} Am^{-2}. (**d**) Comparison of calculated current density and current density for gallium arsenide with $I_p{}^0$ value as 8.7321×10^{-5} Am^{-2}.

3.2. DLS Model

The DLS simulations for the PEC study utilized a high-performance computing setup with an Intel i7 processor, 16 GB of DDR4 RAM on Windows 12. A total of 300 samples, each with an average data size of 10 MB, were processed and split into a training set of 250 samples, a validation set of 25 samples, and a test set of 25 samples. The model training lasted approximately 1 h, with validation and testing taking about 10 min and 5 min, respectively. The training involved 300 iterations with a learning rate of 0.001 and a batch size of 32. Python v 3.12.7 was used for the simulations, and output metrics included accuracy, precision, recall, and F1-score, with data visualization conducted using Matplotlib v 3.9.2. This thorough approach ensured a robust evaluation of model performance for the DLS simulation in PECs. A visual representation of the ANN model with the input, output, and hidden layers is depicted in Figure 5a. It simplifies the visualization of the flow of the input feature through the network of neurons and helps understanding the complex network of neurons present in the ANN model. The efficiency and R^2 value of the model is 98.937% and 0.98937, respectively. A mean squared error of 0.0107 is observed. Figure 5b shows a comparison between the actual and predicted data for a change in normalized photocurrent density with normalized voltage. The predicted data closely correspond to the trend of the original dataset with a slight deviation at low voltage. The low mean squared error value is a clear indication that the model is able to capture the complex relationship between the parameters in the dataset and thus is an excellent fit for the data.
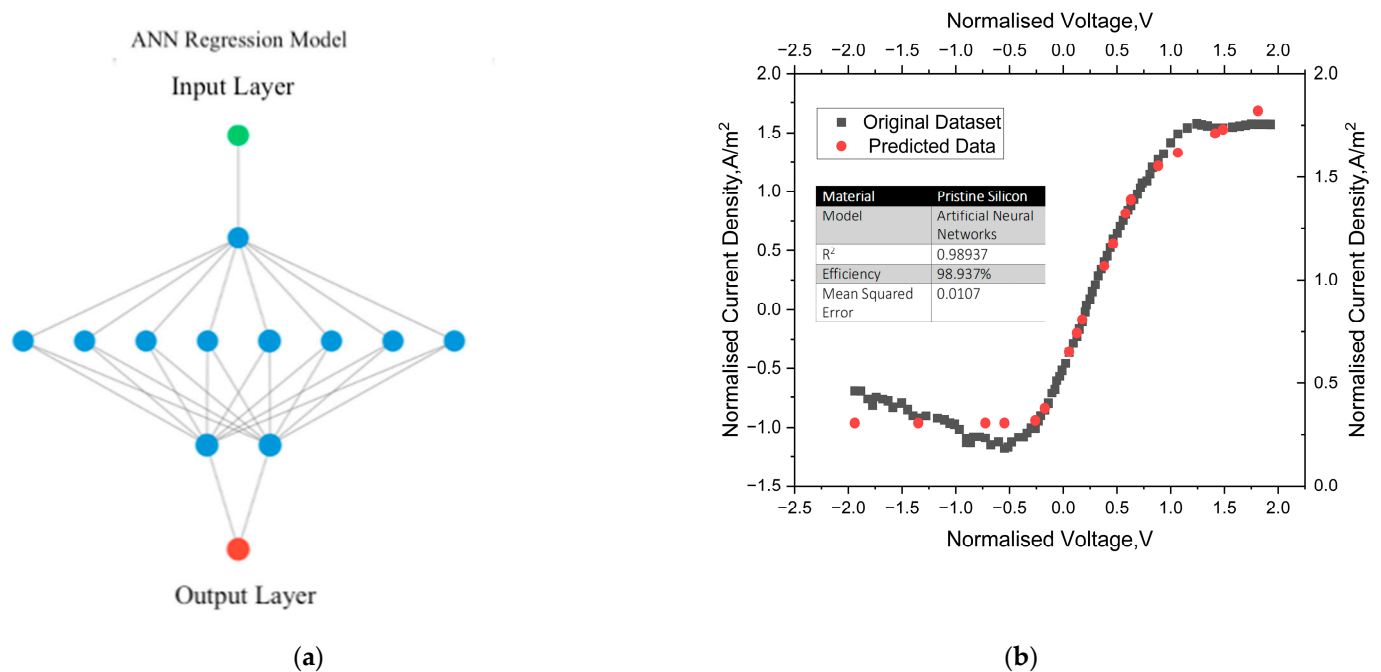

(**a**)

(**b**)

Figure 5. (**a**) Visualization of ANN model and (**b**) overlay of predicted data from testing dataset against the actual trend of dataset with ANN model results.

4. Limitations and Applicability of the Semi-Empirical and ANN Models

The empirical model presented in this study adopts a common derivative methodology based on Gartner et al. A direct correlation between space charge width, I_p^o and voltage for materials such as gallium arsenide and pristine silicon in a PEC system is established. This exercise is not a straight-forward approach as it was impossible to obtain a correlation between hole exchange current parameter and current in a mathematical model. The framework uses a regression approach to solve the model and captures all experimental measured data points. However, the accuracy of the prediction of voltage or current drop relies on the extrapolation attempt and error minimization methods. The fundamental transport mechanisms underlying the solid electrolyte interface is considered for the first

time and correlated to obtain specific parameters. Therefore, it is extremely useful to improvise Gartner's model when a framework is unable to identify or differentiate I_p^0 to perform the analysis. The trend of the experimental data has accurately predicted the behavior of the PEC. However, the model derivation and approach are lengthy. Since the number of variables are more and data analysis methods presented here are cumbersome, the model could yield unpredictable results when the size of the experimental data file is large. To overcome this issue, data mining is needed to make a universal format to present the data obtained from the model and the experiments. The predictability of the empirical model is severely limited by the particular operating range or initial design of the experiment carried out by the observer. However, the model can be directly used to calibrate and predict preliminary experimental data [5,6] in the PEC.

We have used the Artificial Neural Network (ANN) toolbox models to improve the prediction and assess the efficacy of the empirical model although this approach requires a large base at specific operating conditions. Using the database, the polarization patterns are recognized in the big dataset. The ANN model consists of an interconnected collection of artificial neurons and connected input/output parameters where each connection has been associated with the weight. It involves long training times because we have several parameters involved in the hydrogen conversion process. Furthermore, these parameters are determined empirically in this study. The advantage of the proposed ANN model is (a) its ability to handle many tasks at the same time, (b) its predictability improved by number of cells and networks, (c) its ability to function even if the neural network is gradually broken down, and (d) its functionality to learn from past events to make right decisions. However, solving the ANN model and performing data mining are computationally expensive.

5. Conclusions

We report a semi-empirical and deep learning simulation model to assess the performance of a PEC. The I-V model accurately computes the photocurrent density for n- and p-type semiconductor nanoscaled materials using the modified equations provided in this study. Hence, the performance of the cell as a function of voltage, which was difficult to assess so far, can be accurately computed. The polarization curves for two different nanomaterials, gallium arsenide and pristine silicon, were calculated with an R^2 value of 0.997 (GaAs) between the calculated curve and the curve found in the literature. Similar correlations are observed for FeO (0.92) and doped TiO_2 (0.95). The ANN model predicts the photocurrent values with voltage as the input feature with an efficiency of 98.937% and a mean squared error of 0.0107. The polarization analysis and performance prediction pave the way for model-assisted analysis and optimization of the SEI in PECs. The present model can be extended to the analytical approximation for the electron exchange current parameter and calculate the electron's contribution to the current density across the SEI. Nevertheless, the model as presented here is a valuable tool for assessing the performance across the SEI, identifying loss channels, and optimizing the interface parameters for the highest efficiency.

Author Contributions: Conceptualization, S.R.D., R.K.C. and N.S.; methodology, N.S. and S.R.D.; software, R.K.C.; validation, S.R.D., M.W.F., M.S., B.W. and D.S.; formal analysis, N.S. and R.K.C.; investigation, M.W.F., S.R.D. and D.S.; resources, S.R.D.; data curation, N.S.; writing—original draft preparation, S.R.D. and N.S.; writing—review and editing, B.W., S.R.D. and M.S.; visualization, S.R.D. and N.S.; supervision, S.R.D.; project administration, M.W.F., S.R.D. and B.W. All authors have read and agreed to the published version of the manuscript.

Appendix A

The hole exchange current parameter is derived using the following equations (Equations (A1)–(A5)).

$$A = I_p{}^o + J_o e^{\frac{qV}{kT}} \tag{A1}$$

$$\frac{p_s}{p_{s0}} = \left(\frac{\left\{ -K + \sqrt{(K^2 + 4AB)} \right\}}{2A} \right)^2 \tag{A2}$$

$$B = I_p{}^o + J_g \tag{A3}$$

$$K = \Pi k T n_i W e^{\frac{qV}{2kT}} \tag{A4}$$

$$J_p = I_p{}^o \left(\frac{p_s}{p_{so}} - 1 \right) \tag{A5}$$

Assuming $I_p{}^o >> J_o e^{\frac{qV}{kT}}$ and $J_g >> I_p{}^o$, the parameters A and B are reduced to simpler expressions (Equation (A6)).

$$A = I_p{}^o \text{ and } B = J_g \tag{A6}$$

On substituting $\frac{p_s}{p_{s0}}$, A, and B from Equations (A2) and (A6), respectively, in Equation (A5), we obtain the following expression on simplification (Equation (A7)).

$$J_p = \left(\frac{\left(2K^2 - 2(K)\sqrt{K^2 + 4I_p{}^o J_g} + 4I_p{}^o J_g \right) - 4\left(I_p{}^o\right)^2}{4\left(I_p{}^o\right)} \right) \tag{A7}$$

$$J_p = \left(\frac{K^2}{2I_p{}^o} - \left(\frac{K\sqrt{(K^2 + I_p{}^o J_g)}}{2I_p{}^o} \right) + J_g - I_p{}^o \right) \tag{A8}$$

We can further simplify the expression by extracting K from the component under the square root of the expression (Equation (A9)).

$$J_p = \left(\frac{K^2}{2I_p{}^o} \left(\frac{K - \sqrt{K^2 + I_p{}^o J_g}}{K} \right) + J_g - I_p{}^o \right) \tag{A9}$$

The terms J_g and J_p both represent the current across the SEI and thus have been brought to the same side of the equation (Equation (A10)).

$$J_p - J_g = \left(\frac{K}{2I_p{}^o} \left(K - \sqrt{J_g}\sqrt{\frac{K^2}{J_g} + I_p{}^o} \right) - I_p{}^o \right) \tag{A10}$$

$\frac{K^2}{J_g}$ is assumed to be equal to a new term z to simplify the expression (Equation (A11)).

$$J_p - J_g = \left(\frac{K}{2I_p{}^o} \left(K - \sqrt{J_g}\sqrt{z + I_p{}^o} \right) - I_p{}^o \right) \tag{A11}$$

$2(J_p - J_g)$ is substituted with a new term m to obtain Equation (A11).

$$m = \left(\frac{K\left(K - \sqrt{J_g}\sqrt{z + I_p{}^o} \right) - 2\left(I_p{}^o\right)^2}{I_p{}^o} \right) \tag{A12}$$

$$I_p{}^o = \left(\frac{K\left(K - \sqrt{J_g}\sqrt{z + I_p{}^o}\right) - 2\left(I_p{}^o\right)^2}{m} \right) \tag{A13}$$

Assuming $I_p{}^o \gg z$, Equation (A13) is re-written as Equation (A14).

$$I_p{}^o = \frac{K^2}{m} - \frac{K}{m}\sqrt{J_g}\sqrt{I_p{}^o} - 2\frac{K}{m}\left(I_p{}^o\right)^2 \tag{A14}$$

$$mI_p{}^o = K^2 - K\sqrt{J_g}\sqrt{I_p{}^o} - 2K\left(I_p{}^o\right)^2 \tag{A15}$$

Assuming $\sqrt{I_p{}^o}$ as x,

$$2Kx^4 + mx^2 + K\sqrt{J_g}x - K^2 = 0 \tag{A16}$$

Since $I_p{}^o$ is a very small number, $\left(I_p{}^o\right)^2 \ll I_p{}^o$ which implies that $x^4 \ll x^2$. Thus the fourth order term of the equation is neglected to obtain a quadratic equation.

$$mx^2 + K\sqrt{J_g}x - K^2 = 0 \tag{A17}$$

The root of the quadratic equation has been calculated and the negative root has been neglected.

The final expression for the hole exchange current parameter is as follows in Equation (A18).

$$I_p{}^o = \frac{-K\sqrt{J_g} + \sqrt{J_g K^2 + 8(J_p - J_g)K^2}}{4(J_p - J_g)} \tag{A18}$$

The above expression (Equation (A18)) is the derived analytical equation for the hole exchange current parameter.

References

1. Van de Krol, R.; Grätzel, M. *Photoelectrochemical Hydrogen Production*; Springer: New York, NY, USA, 2014.
2. IEA. *Global Energy Review 2*; IEA: Paris, France, 2021; Available online: https://www.iea.org/reports/global-energy-review-2021 (accessed on 31 January 2022).
3. Hernández, S.; Barbero, G.; Saracco, G.; Alexe-Ionescu, A.L. Considerations on Oxygen Bubble Formation and Evolution on BiVO4Porous Anodes Used in Water Splitting Photoelectrochemical Cells. *J. Phys. Chem. C* **2015**, *119*, 9916–9925. [CrossRef]
4. Gärtner, W.W. Depletion-layer photoeffects in semiconductors. *Phys. Rev.* **1959**, *116*, 84. [CrossRef]
5. Reichman, J. The current–voltage characteristics of semiconductor-electrolyte junction photo-voltaic cells. *Appl. Phys. Lett.* **1980**, *36*, 574–577. [CrossRef]
6. Wenger, S.; Schmid, M.; Rothenberger, G.; Gentsch, A.; Grätzel, M.; Schumacher, J.O. Coupled Optical and Electronic Modelling of Dye-Sensitized Solar Cells for Steady-State Parameter Extraction. *J. Phys. Chem. C* **2011**, *115*, 10218–10229. [CrossRef]
7. Sharma, K.; Sharma, V.; Sharma, S.S. Dye-Sensitized Solar Cells: Fundamentals and Current Status. *Nanoscale Res. Lett.* **2018**, *13*, 381. [CrossRef] [PubMed]
8. Giacoppo, G.; Trocino, S.; Lo Vecchio, C.; Baglio, V.; Díez-García, M.I.; Aricò, A.S.; Barbera, O. Numerical 3D Model of a Novel Photoelectrolysis Tandem Cell with Solid Electrolyte for Green Hydrogen Production. *Energies* **2023**, *16*, 1953. [CrossRef]
9. Wang, Z.; Gu, Y.; Zheng, L.; Hou, J.; Zheng, H.; Sun, S.; Wang, L. Machine Learning Guided Dopant Selection for Metal Oxide-Based Photoelectrochemical Water Splitting: The Case Study of Fe_2O_3 and CuO. *Adv. Mater.* **2022**, *34*, 2106776. [CrossRef] [PubMed]
10. Oral, B.; Can, E.; Yildirim, R. Analysis of photoelectrochemical water splitting using machine learning. *Int. J. Hydrogen Energy* **2022**, *47*, 19633–19654. [CrossRef]
11. Kharade, S.K.; Kamat, R.K.; Kharade, K.G. Simulation of Dye Synthesized Solar Cell using Artificial Neural Network. *Int. J. Eng. Adv. Technol.* **2019**, *9*, 1316–1322. [CrossRef]
12. Konorov, P.P.; Yafyasov, A.M.; Bogevolnov, V.B. *Field Effect in Semiconductor-Electrolyte Interfaces*; Princeton University Press: Princeton, NJ, USA, 2006.
13. Sellers MC, K.; Seebauer, E.G. Measurement method for carrier concentration in TiO_2 via the Mott–Schottky approach. *Thin Solid Film.* **2011**, *519*, 2103–2110. [CrossRef]

14.   Kosmulski, M. The significance of the difference in the point of zero charge between rutile and anatase. *Adv. Colloid Interface Sci.* **2002**, *99*, 255–264. [CrossRef] [PubMed]

15.   Sah, C.T.; Noyce, R.N.; Shockley, W. Carrier generation and recombination in p-n junctions and p-n junction characteristics. *Proc. IRE* **1957**, *45*, 1228–1243. [CrossRef]

16.   Xing, Z.; Ren, F.; Wu, H.; Wu, L.; Wang, X.; Wang, J.; Wan, D.; Zhang, G. Enhanced PEC performance of nano porous Si photoelectrodes by covering HfO$_2$ and TiO$_2$ passivation layers. *Sci. Rep.* **2017**, *7*, 43901. [CrossRef] [PubMed]

17.   Geist, J.; Schaefer, A.R.; Song, J.F.; Wang, Y.H.; Zalewski, E.F. An accurate value for the absorption coefficient of silicon at 633 nm. *J. Res. Natl. Inst. Stand. Technol.* **1990**, *95*, 549. [CrossRef] [PubMed]

18.   Couderc, R.; Amara, M.; Lemiti, M. Reassessment of the intrinsic carrier density temperature dependence in crystalline silicon. *J. Appl. Phys.* **2014**, *115*, 093705. [CrossRef]

19.   Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.