

Article

Comparative Modeling of Vanadium Redox Flow Batteries Using Multiple Linear Regression and Random Forest Algorithms

Ammar Ali ^{1,*} , Sohel Anwar ² and Afshin Izadian ^{1,*} ¹ Electrical Engineering Technology, Purdue University, West Lafayette, IN 47907, USA² Mechanical Engineering, Purdue University, Indianapolis, IN 46202, USA; sanwar@purdue.edu

* Correspondence: ali67@purdue.edu (A.A.); aizadian@purdue.edu (A.I.)

Abstract

This paper presents a comparative study of data-driven modeling approaches for vanadium redox flow batteries (VRFBs), utilizing Multiple Linear Regression (MLR) and Random Forest (RF) algorithms. Experimental voltage–capacity datasets from a 1 kW/1 kWh VRFB system were digitized, processed, and used for model training, validation, and testing. The MLR model, built using eight optimized features, achieved a mean error (ME) of 0.0204 V, a residual sum of squares (RSS) of 8.87, and a root mean squared error (RMSE) of 0.1796 V on the test data, demonstrating high predictive performance in stationary operating regions. However, it exhibited limited accuracy during dynamic transitions. Optimized through out-of-bag (OOB) error minimization, the Random Forest model achieved a training RMSE of 0.093 V and a test RMSE of 0.110 V, significantly outperforming MLR in capturing dynamic behavior while maintaining comparable performance in steady-state regions. The accuracy remained high even at lower current densities. Feature importance analysis and partial dependence plots (PDPs) confirmed the dominance of current-related features and SOC dynamics in influencing VRFB terminal voltage. Overall, the Random Forest model offers superior accuracy and robustness, making it highly suitable for real-time VRFB system monitoring, control, and digital twin integration. This study highlights the potential of combining machine learning algorithms with electrochemical domain knowledge to enhance battery system modeling for future energy storage applications.

Keywords: VRFB; machine learning; ECM; linear regression; Lasso; Ridge; MLR

Received: 17 May 2025

Revised: 23 June 2025

Accepted: 30 July 2025

Published: 5 August 2025

Citation: Ali, A.; Anwar, S.; Izadian, A. Comparative Modeling of Vanadium Redox Flow Batteries Using Multiple Linear Regression and Random Forest Algorithms. *Energy Storage Appl.* **2025**, *2*, 11. <https://doi.org/10.3390/esa2030011>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Vanadium redox flow batteries (VRFBs) [1] are advanced energy storage systems designed for large-scale applications due to their flexibility, safety, and long life cycle [2–4]. These benefits have resulted in a 21% yearly increase in VRFBs in China [5]. The principle of operation is based on storing energy in vanadium ions dissolved in electrolyte solutions, which are circulated through electrochemical cells during charging and discharging, as shown in Figure 1 [1]. There are several modeling approaches for VRFBs, including equivalent circuit models. These models provide a discrete set of equations selected to express the operating condition of the battery. The coefficients must also change based on various sizes and operating conditions. Puleston et al. [6] provided a comprehensive review of modeling and estimation methods for vanadium redox flow batteries (VRFBs). The review highlights state observers, parameter estimation techniques, and dynamic models for estimating internal states like state of charge (SOC) and State of Health (SOH), emphasizing the challenges of managing VRFBs' nonlinear behavior.

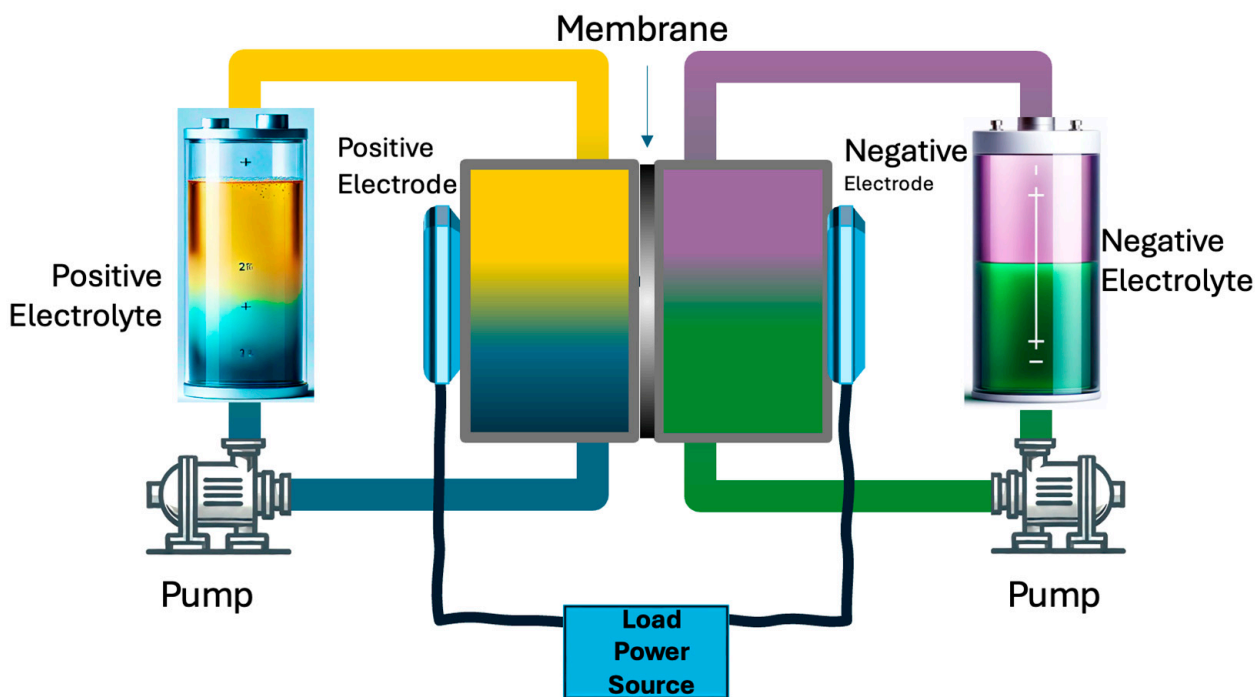


Figure 1. General schematic of a VRFB.

Machine learning (ML) is a transformative tool that enables systems to analyze large datasets, uncover patterns, and model complex, nonlinear relationships with minimal programming [7]. Introduced by Sir Francis Galton in the late 19th century [8], linear regression is a statistical method that, in its modern form, can optimize operating parameters and predict performance metrics. Ridge [9–12] and Lasso [13] regression methods were introduced to shrink or eliminate the coefficients, effectively performing feature selection while maintaining predictive accuracy. Random Forest (RF) algorithms have gained attention for their robustness, ability to model nonlinear relationships, and resistance to overfitting. For instance, in a study focused on electrode optimization, Random Forest and Gradient Boosting methods were applied to predict and enhance the microstructural characteristics of carbon cloth electrodes in VRFBs, ultimately aiming to improve energy efficiency [14]. Another study employed ensemble ML techniques, including Random Forest, to predict power loss in VRFB systems under varying operating conditions, demonstrating the effectiveness of data-driven approaches in capturing complex system behaviors [15].

The modeling of batteries has been advanced by machine learning tools. Operating parameters and performance metrics, like state of charge (SOC) and voltage, have been accurately replicated [16]. Li et al. [17] integrated ML with Absolute Nodal Coordinate Formulation (ANCF-e.), which used machine learning (ML) techniques to estimate nonlinear potentials, such as open-circuit potential, electrolyte potential, and lithium-intercalation overpotential, in lithium-ion batteries (LIBs) with reduced computational complexity for real-time state and parameter estimation. Two hybrid models, “HYBRID-I and HYBRID-II,” were introduced to integrate the Single-Particle Model with Thermal dynamics (SPMT) and feedforward neural networks (FNNs) to improve the prediction of lithium-ion battery behavior [18]. Using linear regression, Support Vector Regression, and XGBoost on data in a 1 kW, 6 kWh VRFB, XGBoost achieved the best performance (R^2 : 0.99, MAE: 0.24, RMSE: 0.32), demonstrating its effectiveness in optimizing VRFB thermal management. The SOC improvement, voltage, and current predictions were developed through a deep reinforcement of the learning framework [19]. Compared to traditional methods like Kalman filters and regression, deep Q-learning and dueling DQN improved accuracy by 10%,

particularly in voltage prediction, demonstrating more accurate model generalization. An optimization framework was developed for VRFBs using ANN and GA to maximize net discharge power [20]. By optimizing the electrode compression ratio, applied current density, and electrolyte inflow rate, the hybrid ANN-GA approach reduced resistances by 66% and improved efficiency, outperforming traditional methods across different SOC levels.

Zheng et al. [21] used a Recurrent Equilibrium Network (REN), a neural network-based approach, to estimate the state of charge (SOC) in vanadium redox flow batteries. The model was trained on real-world data using voltage, current, and temperature inputs. It outperformed BPNN (Backpropagation Neural Network) and LSTM (Long Short-Term Memory), showing high accuracy and robustness. Narayan et al. [22] applied machine learning models (LR, SVR, XGBoost) to predict the temperature rise in vanadium redox flow batteries under varying current conditions. Using experimental data, XGBoost achieved the highest accuracy ($R^2 = 0.99$), showing strong potential for thermal management. Xiong et al. [23] used a GRU-based neural network model to predict the terminal voltage in VRFBs, incorporating the flow rate, current, SOC, and voltage. Trained on experimental data, the model outperformed ECMs, SVM, and BPNN with an MAE < 0.006 V and $R^2 > 0.99$, proving effective for real-time control. Martínez-López et al. [24] used two ANN models trained on CFD-generated data to predict cell voltage and overpotential in VRFBs. Their models achieved high accuracy ($R \approx 0.999$) and drastically reduced the computation time, enabling fast, real-time battery analysis.

Despite these advancements, most studies have focused on optimizing specific subsystems (e.g., electrodes) or predicting individual performance metrics (e.g., power loss, temperature rise). They often focus on optimizing isolated components or evaluating a single algorithm. Few works have conducted a direct, systematic comparison between interpretable models like Multiple Linear Regression (MLR) and more complex ensemble methods such as Random Forest (RF) in the context of full VRFB system modeling. Addressing this gap, the present work evaluates both MLR and RF models using the same experimental dataset to benchmark their predictive accuracy, interpretability, and applicability to dynamic and steady-state battery behavior.

The selection of MLR and Random Forest for this study is motivated by their complementary strengths. MLR is a foundational, interpretable algorithm that provides explicit weight-based relationships between input features and predicted output, making it ideal for deriving analytical insights into battery behavior. In contrast, Random Forest is a powerful ensemble method that captures complex nonlinear dependencies and interactions without requiring detailed domain-specific feature transformations. Compared to more computationally intensive models such as deep neural networks or support vector machines, RF offers strong predictive performance while maintaining relatively low training costs and good generalizability. This combination allows for both model transparency and accuracy in VRFB performance prediction.

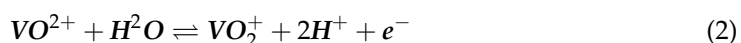
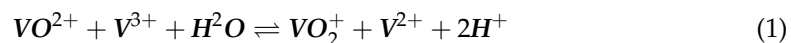
This paper develops MLR and Random Forest models to predict VRFB performance based on experimental datasets. Furthermore, it evaluates both approaches based on their prediction accuracy and computational running times, providing practical insights into their suitability for real-time applications. It presents a novel model of VRFB through a single governing equation utilizing ML techniques, which includes a large set of operating points and size variations. Experimental data is used to train the ML model. Several operating conditions are examined and compared with the existing literature. The Table 1 shows the Nomenclature.

Table 1. Nomenclature.

Symbol	Definition	Symbol	Definition
μ	Viscosity (kg/m·s)	σ	Conductivity (S/cm)
L, D, W, H	Length, diameter, width, height (cm)	V	Volume (L)
CD	Current density (mA/cm ²)	SOC	State of charge
E	Potential (V)	Q	Flow rate (L/min)
T	Temperature (K)	s	Laplace operator
R	Gas constant (8.314 J mol ⁻¹ K ⁻¹)	α	Fix coefficient
F	Faraday constant (96,487 C mol ⁻¹)	Re	Reynolds number
p	Pressure (Pa or psi)	en	Energy
A	Cross-sectional area (cm ²)	opv	Open-circuit voltage
r	Density (kg/m ³)	cha/dis	Charge/discharge
c	Concentration (M)	Man/M	Manifold
P	Power (W)	int	Internal
h	Friction loss/minor loss	A	Anode
~R	Hydraulic resistance (Pa·s·m ⁻³)	C	Cathode
η	Efficiency/diffusion ratio	G	Guide channel
		n	Number of cells

2. Operating Principle of VRFB

The VRFB consists of two tanks (Figure 1). The tanks store the electrolytes with four different oxidization states. These electrolytes circulate by pumps through a cell stack where redox reactions occur, facilitating charge and discharge. The cell stack of a VRFB consists of bipolar plates (graphite or composite), graphite felt electrodes, and an ion-exchange membrane (typically Nafion) that enables proton transfer while preventing electrolyte mixing. Flow frames and gaskets from corrosion-resistant materials like PVC or PTFE ensure proper electrolyte distribution and sealing. These are V^{2+} , V^{3+} , V^{4+} (same as VO^{2+}), and V^{5+} (which is the same as VO_2^+). The chemical cell reactions are as follows:



In this research, the VRFB specifications used were similar to those by Kim et al. [25]. Specifications are shown in Table 2.

Table 2. Specifications of VRFB system.

Specification	Details
Electrolyte Properties	Density: 1400 kg/m ³ , Viscosity: 0.006 kg/m·s, Conductivity: 0.2 S/cm
Cell Active Area	780 cm ²
Flow Field Design	5 guide channels (32 cm effective length, 1 mm width, 3 mm depth) connected to buffer layers
Pump Type	Centrifugal pumps (Finnish Thomson, DB4V-T-M613)

Table 2. Cont.

Specification	Details
Piping	1-inch PVC pipe for tank to pump; 1/2-inch PVC pipe elsewhere
Flow Monitoring	Flow meters (FMG83-FKM for 1 kW system)
Pressure Sensors	Measurement specialties MSP-300-100-P-1-N-1
Temperature Sensors	1/800 Teflon-coated type N thermocouple in tanks
Operating Temperature	35–40 °C (80 mA/cm ²), 45–49 °C (160 mA/cm ²)
Internal Cell Resistance	1.0 Ω·cm ²
Pumping Power Efficiency	0.5
Shunt Current	Less than 0.15 A per stack

3. Dynamics of VRFB

The operation of vanadium redox flow batteries (VRFBs) involves complex dynamic processes governed by electrochemical, electrical, and transport phenomena. Understanding these dynamics is critical for the accurate modeling, efficient control, and real-time state estimation of VRFB systems.

Three key physical phenomena dominate the dynamic behavior of VRFBs:

1. The Nernst equation governs the open-circuit voltage (OCV) as a function of the state of charge (SOC) and reaction kinetics.
2. Shunt currents arise due to parasitic pathways between cells in series, resulting in efficiency loss and imbalances.
3. Diffusion currents reflect concentration gradients within the electrolyte, influencing both voltage response and charge transport.

3.1. Nernst Equation and State of Charge

The open-circuit voltage (OCV) of the vanadium redox flow battery (VRFB) depends on the concentrations of the active species and the operating temperature. The OCV can be expressed using the Nernst equation:

$$E_{ocv} = E_0 + \frac{2RT}{F} \ln \left(\frac{SOC_{cell}}{1 - SOC_{cell}} \right) \quad (4)$$

The electrolyte's state of charge (SOC) represents the overall concentration level of vanadium ions within the battery.

$$SOC_{tank}(t) = SOC_{tank}(0) + \frac{n_{cell}}{FV_{tank}C_v} \int_0^t I_{cell} d\tau \quad (5)$$

$$SOC_{cell} = SOC_{tank}(t) + \frac{n_{cell}I_{cell}}{2FQ_{vrb}C_v} \quad (6)$$

3.2. Shunt Current and Diffusion Current

Shunt currents result from potential gradients across the cell stack, driving ion movement through manifolds and guide channels, reducing coulombic efficiency. Minimizing them requires increasing electrolyte path resistance using narrow channels and common manifolds. During charging, shunt currents lower the available cell current, while during discharge, they increase it, leading to efficiency losses. Though minimal in this study [26] (<0.15 A), their impact grows with more series-connected cells, making them critical in

large-scale VRFBs. Additionally, vanadium ion diffusion causes self-discharge and side reactions, leading to capacity loss over extended cycling.

To simplify the analysis, Zhang [26] proposed an equation to estimate the diffusion effect. The complementary part of the coulombic efficiency η_{CE} is denoted as the diffusion ratio η_{diff} , which represents the self-discharge loss due to ion diffusion in each charge/discharge cycle. The diffusion ratio is defined as follows:

$$I_{diff} = \left(\frac{\eta_{diff}}{2 - \eta_{diff}} \right) |I_{ter}| \quad (7)$$

Based on Kim's report [25], the coulombic efficiency is approximately 95%, resulting in a diffusion ratio η_{diff} of 5%. The term I_{diff} is introduced to estimate the self-discharge current caused by ion diffusion.

4. Machine Learning: Multiple Linear Regression

Machine learning regression (MLR) enhances traditional regression by efficiently capturing nonlinear relationships in high-dimensional data. In battery electrochemical modeling, MLR predicts key parameters like terminal voltage and internal resistance based on factors such as the SOC, current density, and temperature. Unlike computationally intensive models, Multiple Linear Regression (MLR) is a statistical method used to model the relationship between a dependent variable and multiple independent variables by fitting a linear equation to observed data. The general form of MLR is $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon$, where y is the dependent variable, $x_1, x_2, \dots, x_n$ are the independent variables, β_0 is the intercept, $\beta_1, \beta_2, \dots, \beta_n$ are the coefficients (weights), and ϵ represents the error term. The weights β are calculated by minimizing the sum of squared residuals, which is the difference between the observed and predicted values of y . This is typically performed using methods like Ordinary Least Squares (OLS), where the weights are computed as $\{\beta\} = \left(\{X\}^T \{X\} \right)^{-1} \{X\}^T y$, with " $\{X\}$ " representing the matrix of independent variables and y is the vector of observed dependent variable values. This approach ensures that the best-fit line minimizes prediction errors while balancing all contributing factors.

Another technique used with MLR is L_2 regression, also known as Ridge regression, which adds an ℓ_2 -norm penalty to the cost function, which discourages large coefficient values by minimizing their squared magnitudes, reducing model complexity, and preventing overfitting without eliminating any variables.

In Figure 2 the first step is Input Data. This data was collected from [25]. Twelve curves of the voltage–capacity experiment results were digitized to obtain the numerical value of the results.

Step 2: Data Cleaning. This step involves removing or handling missing values, outliers, and inconsistencies in the dataset. It ensures the quality of the data for analysis.

Step 3: Normalization, where we scale the features to a standard range, often between 0 and 1. This step ensures that features with more extensive ranges do not dominate those with smaller ranges.

Step 4: Feature engineering and the creation of additional features. In this step, we used Kirchhoff's law and calculated the following features: terminal current, time, I_{diff} , I_{cell} , $\int_0^t I_{cell} d\tau$, $\int_0^t I_{load} d\tau$, SOC_{tank} , SOC_{cell} , E_{opv} , ΔV_{ter} , SOC_{ini} , flow rate, and $R_{soc}(a)$. These features were calculated using the experiment specifications as in Table 2.

Step 5: Data Splitting. Data was split into training and validation data: 80% of the data for training, and 20% for data testing.

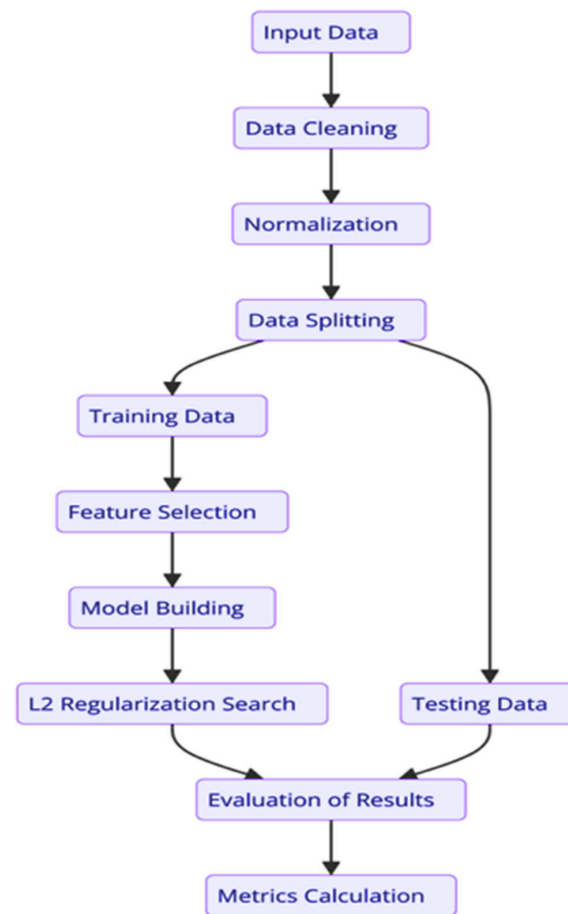


Figure 2. Machine learning process flowchart.

Step 6: Feature Selection on the training data. Here, we used the features as in Equations (9) and (10). This will make the calculated weights lead to the ECM parameters. We also removed the intercept term to model the exact equations.

Step 7: Model Building. Using the training data, the MLR model is built by finding the optimal coefficients for the linear relationship between predictors and the target variable.

Step 8: Ridge Regularization set. A technique to prevent overfitting by penalizing large coefficients in the model. L2 regularization (Ridge regression) introduces a hyperparameter γ to balance the trade-off between minimizing the error and maintaining smaller coefficients. The search of γ was from 10^{-5} to 100. The best value is found and validated from the validation data.

Step 9: Evaluation of Results. The testing data is used here to evaluate and test the model's performance using metrics such as Mean Absolute Error (MAE) and RSS.

The final step is Metric Calculation, where the final performance metrics are calculated to quantify the model's accuracy and reliability, ensuring it meets the desired performance criteria.

Additional steps are used when needed, e.g., Lasso regression, where a linear regression technique is utilized that adds an ℓ_1 -norm penalty to the cost function, shrinking less significant coefficients to zero, effectively performing variable selection and regularization to enhance model simplicity and prevent overfitting.

5. Random Forest Modeling

Random Forest is an ensemble-based machine learning algorithm that builds multiple decision trees and aggregates their outputs to achieve better generalization performance,

as shown in Figures 3 and 4. For regression tasks, such as predicting the terminal voltage of vanadium redox flow batteries (VRFBs), the model averages the output of all trees:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^T h_t(x) \quad (8)$$

where $\hat{y}$ is the predicted voltage, T is the total number of trees, and $h_t(x)$ is the prediction from the t -th decision tree given input vector x .

Random Forest

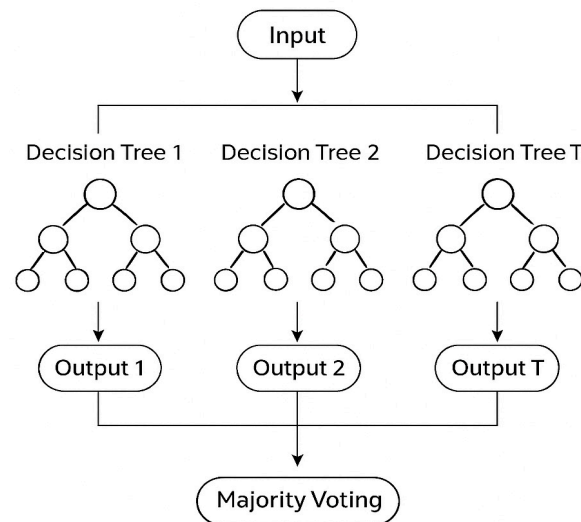


Figure 3. Random Forest overall approach.

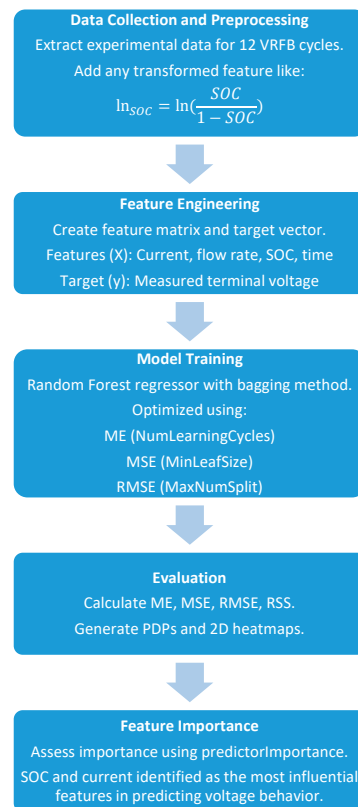


Figure 4. Random Forest flow chart.

Each tree is trained using a different bootstrap sample of the data. To further decorrelate the trees, only a random subset of the input features is considered at each split within each tree. This process reduces overfitting and improves robustness.

5.1. Modeling Workflow

The complete pipeline as shown in Figure 4 used in this study is outlined below as follows:

1. Data Collection and Preprocessing

We extracted experimental data for 12 VRFB charge/discharge cycles from Zhang et al. using digitization techniques. Each dataset includes voltage, capacity, current, SOC, flow rate, and other derived features. We added a transformed feature: $\ln\left(\frac{SOC_{cell}}{1-SOC_{cell}}\right)$. Redundant variables and noise-sensitive columns were removed during preprocessing.

2. Feature Engineering

Among the engineered features, the natural logarithmic transformation $\ln(SOC/(1-SOC))$ was included due to its strong basis in electrochemical theory. This transformation reflects the theoretical structure of the Nernst equation, which governs the open-circuit voltage of redox systems as a logarithmic function of reactant concentrations. In the context of VRFBs, the SOC influences terminal voltage nonlinearly, especially near the limits of charge and discharge. The $\ln(SOC/(1-SOC))$ term helps the model better capture these nonlinear dynamics, improving accuracy in regions where voltage changes rapidly.

Features were divided into dynamic and stationary zones based on observed voltage behavior. We created a feature matrix and target vector. Seventeen features were chosen including current, flow rate, SOC, time, capacity, etc. The target is measured terminal voltage.

3. Data Splitting

The dataset was split into 80% training/validation and 20% test sets using MATLAB R2024a cvpartition. The data size was 1103 for training and 275 for testing.

4. Model Training

A Random Forest regressor was trained using fitrensemble with the bagging method.

5. Evaluation

We calculated the ME, MSE, RSS, and RMSE on both the training and test sets. Partial dependence plots (PDPs) and 2D heatmaps were generated to interpret the influence of top predictors.

6. Feature Importance

Feature importance was assessed via the predictorImportance() function. SOC and current were identified as the most influential features in predicting voltage behavior, aligning with physical intuition.

5.2. Out-of-Bag (OOB) Error for Tree Selection

Out-of-bag error was monitored to determine the optimal number of trees:

$$OOB\ Error = \frac{1}{|O|} \sum_{i \in O} \left(y_i - \hat{y}_{OOB,i} \right)^2 \quad (9)$$

where O is the set of OOB samples and $\hat{y}_{OOB,i}$ is the OOB prediction for the i-th sample. In our model, the OOB error stabilized at 70 trees, which was selected to balance accuracy and computational efficiency.

6. Simulation Results

6.1. MLR-Based VFRB Model

In this model, we initially used most of the data features. The initial candidate features were extracted from the experimental data, including voltage, current, flow rate, SOC values, and time-integrated signals. To reduce redundancy and improve model performance, a two-step feature selection process was applied. Firstly, Lasso regression and Ridge regularization were used to penalize less influential coefficients based on their contribution to model accuracy. Then, secondly, features that were functionally dependent or strongly correlated with others were removed based on domain knowledge. For example, SOC_{cell} was excluded because its information is captured in the logarithmic term $\ln\left(\frac{SOC_{cell}}{1-SOC_{cell}}\right)$, and SOC_{tank} was dropped due to its strong correlation with SOC_{cell}. Similarly, current-related variables such as I_{diff}, I_{cell}, and I_{cell-integral} were excluded since they are derivable from I_{load} and its time integral. After these filtering steps, the final MLR model was constructed using eight features that showed both strong physical relevance and statistical significance in predicting terminal voltage, resulting in the following single governing equation for the proposed VRFB model:

$$\alpha_1 CD + \alpha_2 Q + \alpha_3 I_{load} + \alpha_4 I_{shunt} + \alpha_5 E_{opv} + \alpha_6 \text{Charging state} + \alpha_7 \int_0^t I_{load}(\tau) \cdot d\tau + \alpha_8 \ln\left(\frac{SOC_{cell}}{1-SOC_{cell}}\right) = V \quad (10)$$

In Equation (10), the predicted terminal voltage V is modeled as a weighted sum of engineered features, with coefficients α_i learned from the training data. The variables are defined as follows:

- CD : current density (mA/cm²);
- Q : electrolyte flow rate (L/min);
- I_{load} : terminal load current (A);
- I_{shunt} : shunt current due to side paths (A);
- E_{opv} : open-circuit potential (V);
- Charging state: binary indicator (1 for charging, 0 for discharging);
- $\int_0^t I_{load}(\tau) \cdot d\tau$: time-integrated load current (cumulative charge delivered, A·s);
- $\ln\left(\frac{SOC_{cell}}{1-SOC_{cell}}\right)$: log-transformed state of charge (dimensionless), derived from the Nernst equation formulation.

Each term captures a key operational parameter affecting VRFB voltage behavior, with both electrochemical and time-dependent contributions included. The resulting model coefficients are shown in Table 3.

Table 3. MLR equation weights results.

Parameters	Weights(α_i)
CD	0.017418
Q	−0.018771
I_{load}	−0.0044512
I_{shunt}	−1.0962
E_{opvt}	1.1301
$Discharging$	−3.841
$\int_0^t I_{load}(\tau) \cdot d\tau$	-1.29×10^{-5}
$\ln SOC$	1.2873

To evaluate the sensitivity of model performance to different data split ratios, we tested the MLR and RF models using 70/30, 80/20, and 90/10 train–test partitions. As shown in Table 4, both models maintained consistent accuracy across splits, though RF

consistently outperformed MLR in all metrics. The RF model achieved the lowest RMSE of 0.1157 V with a 90/10 split, compared to 0.1833 V for MLR. Even at the lower training ratio (70/30), RF achieved better error values (RMSE = 0.1561 V) than MLR at any ratio. These results indicate that the RF model is robust to variations in training size and generalizes well, while MLR shows slightly more sensitivity to data partitioning, as shown in Table 4 and Figure 5. In the machine learning literature, an 80/20 train–test split is widely accepted as a standard baseline due to its balance between sufficient training data and statistically meaningful evaluation [27]. The findings in this study confirm the appropriateness of that choice while also demonstrating that RF maintains high accuracy even with limited training data, making it a strong candidate for real-time VRFB system modeling. The 80:20 split evaluation metrics for the training data and the test data are in Table 5.

Table 4. Test error comparison across different data splits.

Model	Split	ME	MSE	RSS	RMSE
MLR	90/10	0.01861	0.03359	4.60154	0.18327
MLR	80/20	0.02039	0.03226	8.87039	0.17960
MLR	70/30	0.01775	0.03440	14.20663	0.18547
RF	90/10	−0.00473	0.01339	1.83445	0.11572
RF	80/20	0.00591	0.02854	7.84939	0.16895
RF	70/30	0.00852	0.02437	10.06540	0.15611

Table 5. Evaluation metrics for MLR vs. RF models.

Model	Zone	ME	MSE	RSS	RMSE
Random Forest	Dynamic	0.02846	0.04189	2.42934	0.20466
	Stationary	−0.00012	0.02498	5.42006	0.15804
	Total	0.00591	0.02854	7.84939	0.16895
MLR	Dynamic	0.00726	0.05067	2.93901	0.22511
	Stationary	−0.02778	0.02733	5.93139	0.16533
	Total	0.02039	0.03226	8.87039	0.17960

Table 5 presents a side-by-side comparison of evaluation metrics for the Multiple Linear Regression (MLR) and Random Forest (RF) models based on the test dataset. The RF model consistently achieved lower error values across all metrics: mean error (ME), mean squared error (MSE), residual sum of squares (RSS), and root mean squared error (RMSE). Specifically, RF produced a total RMSE of 0.1689 V compared to 0.1796 V for MLR, and a lower total RSS of 7.85 versus 8.87. These results confirm that the RF model provides more accurate and less biased voltage predictions for VRFB systems than the linear MLR model.

Figure 6 compares the prediction errors of the Random Forest (RF) and Multiple Linear Regression (MLR) models using four metrics—ME, MSE, RSS, and RMSE—across dynamic, stationary, and overall regions. The RF model consistently outperformed MLR, particularly in dynamic conditions, with lower RMSE (0.2047 V vs. 0.2251 V) and RSS (2.43 vs. 2.94). In stationary regions, RF also showed slightly better accuracy, and overall, it achieved the lowest total RMSE (0.1689 V) and near-zero mean error, highlighting its superior robustness and accuracy in modeling VRFB terminal voltage.

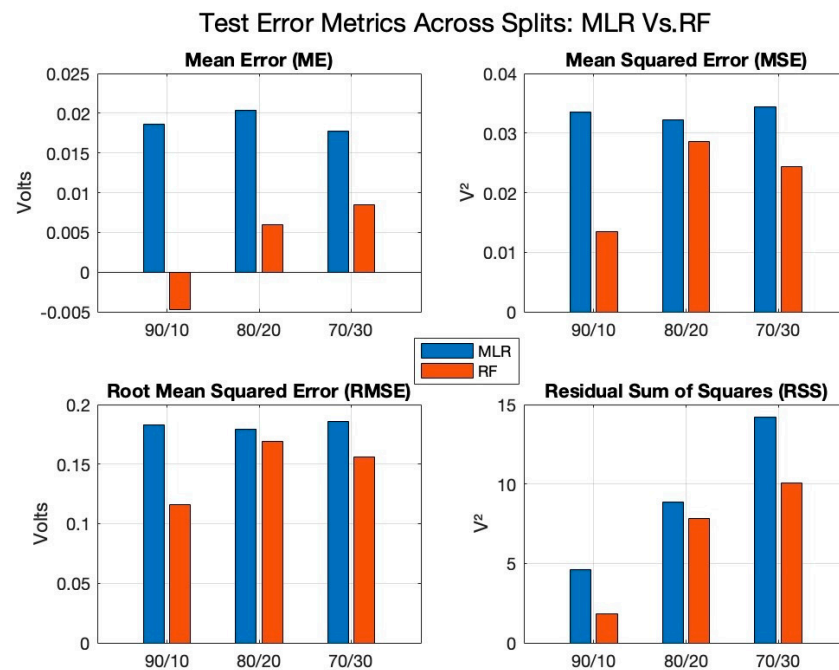


Figure 5. Test error metrics across different data splits.

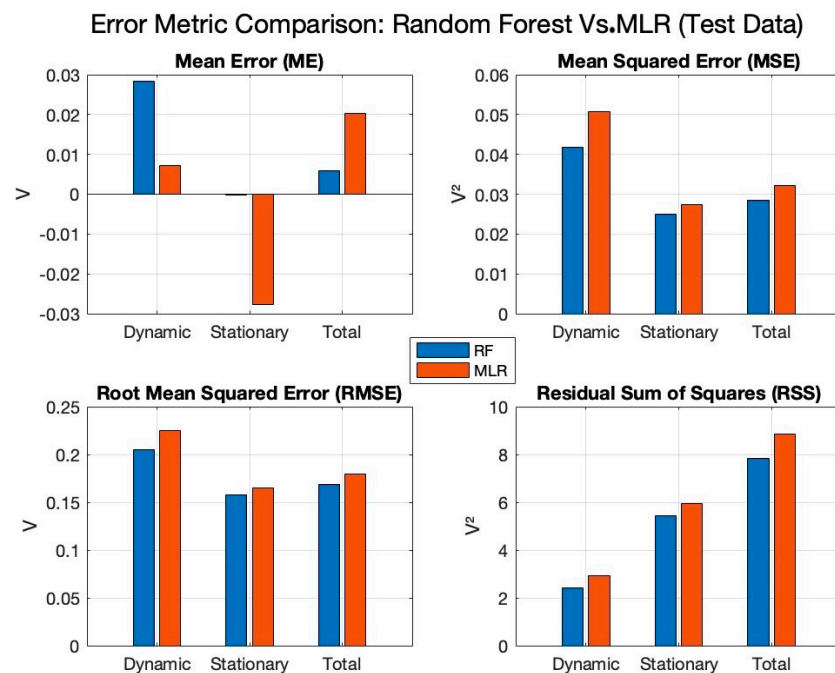


Figure 6. Errors of both RF and MLR models.

Figure 7 shows the predicted values of the trained model compared to the actual experimental training data. The best model will have all points converted onto the 45-degree line. Figure 8 shows the terminal voltage of the data recorded from experiments and error residuals for each voltage. This graph demonstrates the error variations with respect to the voltage. The predicted voltage and measured voltages that are not used in the training of ML are shown in Figure 9. Figure 10 shows the performance of the Random Forest model training data. Figure 11 is the model voltage prediction compared to the experimental voltage at various charging capacities. As the figure shows, the prediction and the measurements are in good agreement, as evidenced by the errors shown in Figure 12.

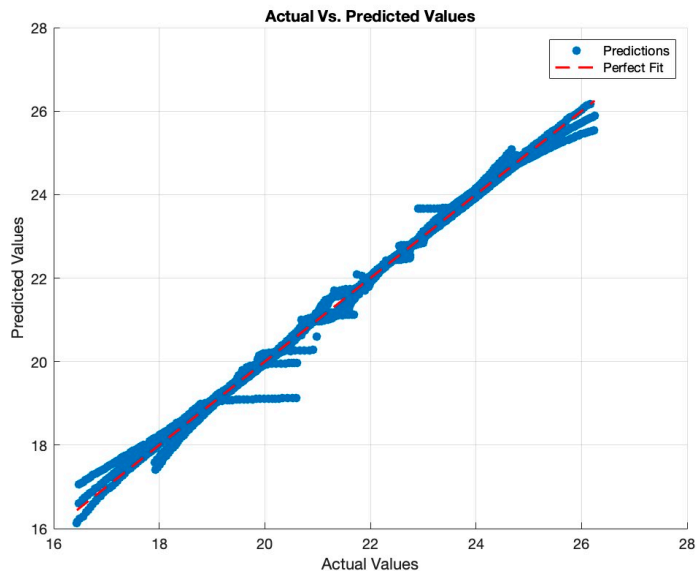


Figure 7. Actual vs. predicted voltage for the 80% training data under MLR equation.

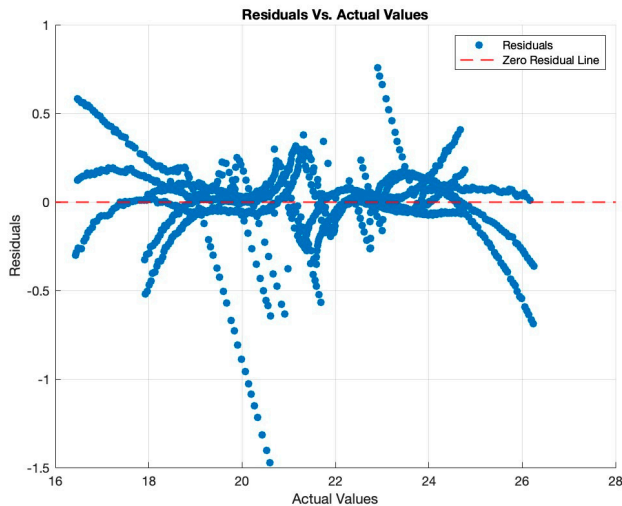


Figure 8. Residuals vs. actual values of voltage for the trained data.

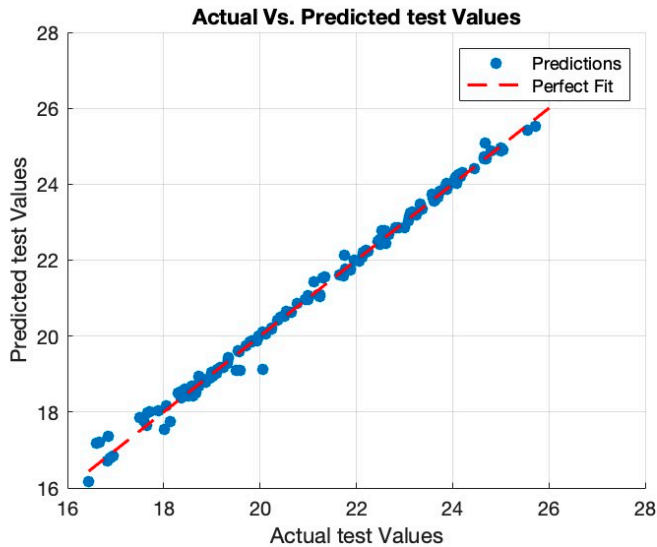


Figure 9. Actual vs. predicted voltage for the 20% test data.

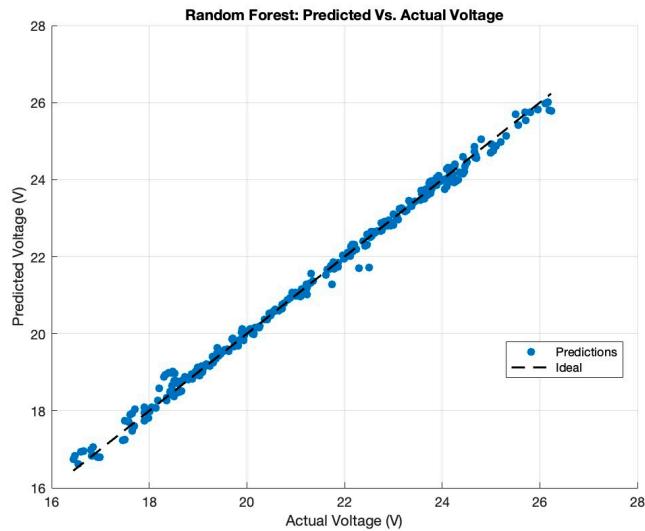


Figure 10. Predicted vs. actual voltages for RF.

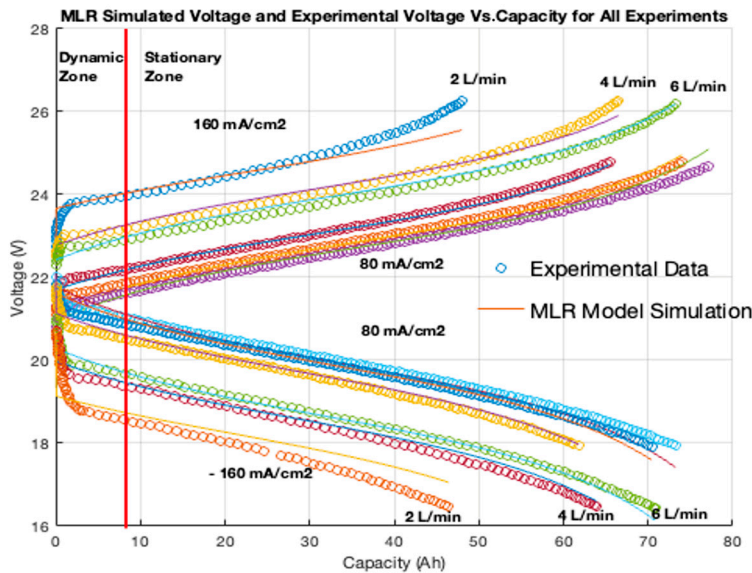


Figure 11. Measured and simulated voltage. Zhang [22] MLR.

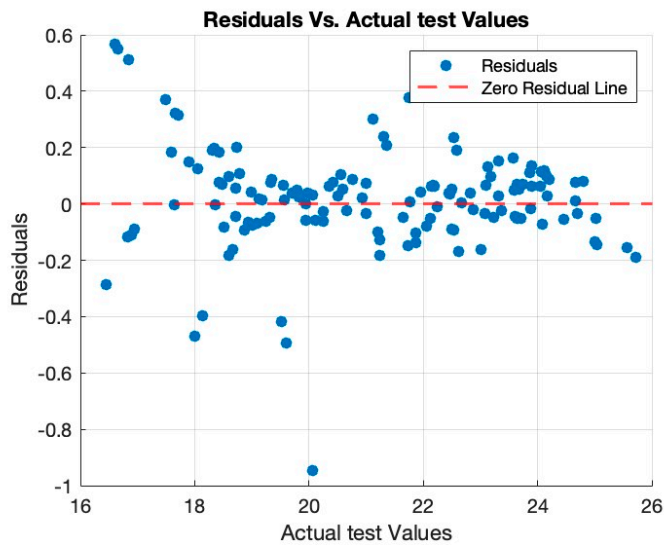


Figure 12. Residuals vs. test data voltage MLR.

To ensure a fair comparison between the MLR and RF models, we computed the mean error (ME) and residual sum of squares (RSS) for both methods. As shown in Table 5, the RF model test data achieved a lower RMSE (0.1699 V vs. 0.1796 V for MLR) and a smaller RSS (7.94 vs. 8.87), indicating a better overall fit to the test data. The ME for RF was also closer to zero (-0.0019 V) compared to MLR (0.0204 V), demonstrating reduced systematic bias. These metrics highlight the RF model's improved predictive accuracy and robustness, particularly under complex operating conditions.

Figure 13 shows the out-of-bag error compared to the number of trees. As the figure shows, the curves stabilize after 20 trees. We have selected 50 trees to achieve better evaluation metrics. The out-of-bag (OOB) error plot provides an internal validation method for Random Forest models without requiring a separate validation dataset. In Random Forest, each decision tree is trained on a bootstrap sample, leaving approximately one-third of the data unused (the OOB samples) for that tree. The OOB error at each stage is computed as the mean squared prediction error for these unseen samples, averaged across all trees grown so far. In the OOB error curve (Figure 13), the horizontal axis represents the number of trees included in the ensemble, while the vertical axis shows the corresponding OOB mean squared error. The curve typically decreases sharply at first as more trees are added, then gradually levels off once additional trees provide diminishing improvements in prediction accuracy. Compared to traditional k-fold cross-validation, OOB evaluation is faster and more efficient, as it leverages the bootstrap sampling inherent to the Random Forest algorithm. It also provides an unbiased performance estimate without additional data splitting. This makes OOB optimization particularly useful for iterative model tuning, such as selecting the ideal number of trees while avoiding overfitting.

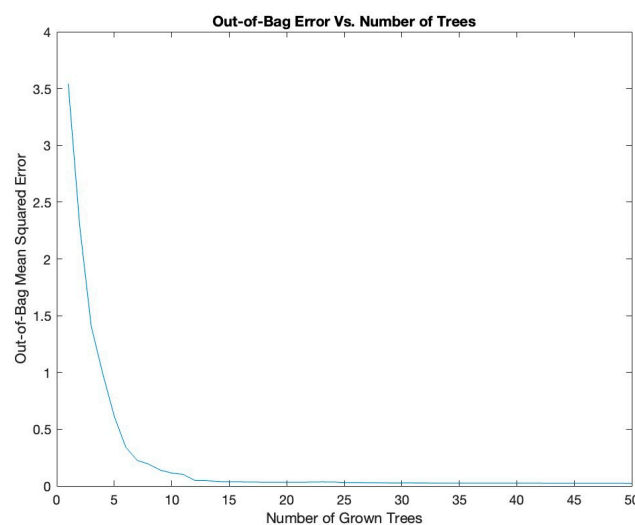


Figure 13. Out-of-bag error vs. number of trees for RF.

Figure 14 shows the RF model's error residual compared to the actual data training set. Figure 15 shows the same residual error for the test dataset. Figure 16 shows the RF modeled voltage compared to the measured voltages. Figure 17 shows the feature importance plot and summarizes the contribution of each input variable to the Random Forest model's predictive accuracy. In this figure, the horizontal axis represents the normalized importance scores, while each bar corresponds to a specific feature used in the model. Features with longer bars indicate a greater overall reduction in prediction error across the forest when that feature was used for node splits. In contrast, features with smaller bars had minimal impact on the model's decision-making process. This visualization provides a clear hierarchy of which variables were most influential in driving the model's predictions,

offering a powerful interpretability tool for analyzing complex machine learning models like Random Forest.

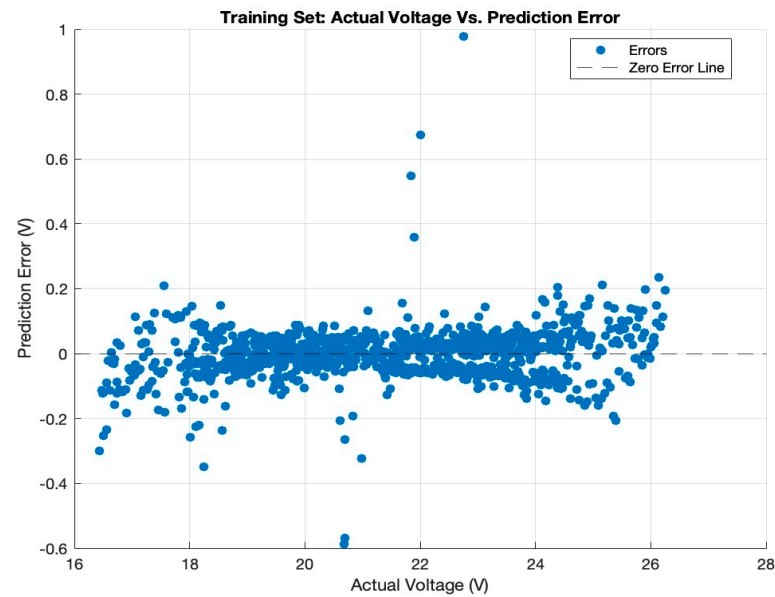


Figure 14. Actual voltage vs. error for training set for RF.

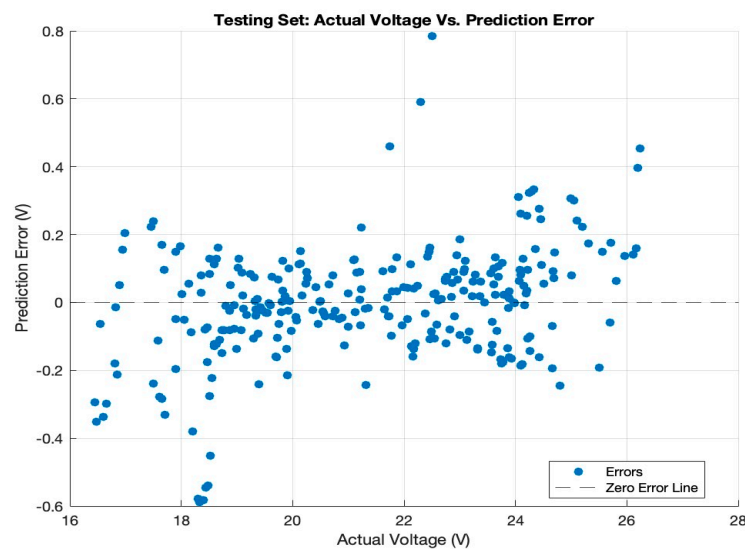


Figure 15. Actual voltage vs. error for testing set for RF.

Partial dependence plots (PDPs) in Figure 18 visualize the marginal effect of individual input features on the predicted output of a machine learning model, in this case, the Random Forest regressor. In each PDP, the horizontal axis represents the range of values for a specific input feature, while the vertical axis shows the average predicted terminal voltage as that feature varies, holding all other features constant. The curve illustrates how changes in a single feature influence the model's output independently of other variables. PDPs provide a valuable tool for interpreting complex, nonlinear models, helping to reveal whether the relationship between a feature and the target is linear, monotonic, or more complex. The model assigned the highest importance to current-related features, particularly time-integrated load current, cell current, and instantaneous load current. These variables directly impact voltage through resistive and electrochemical effects. This highlights the critical role of current dynamics in VRFB voltage prediction and control.

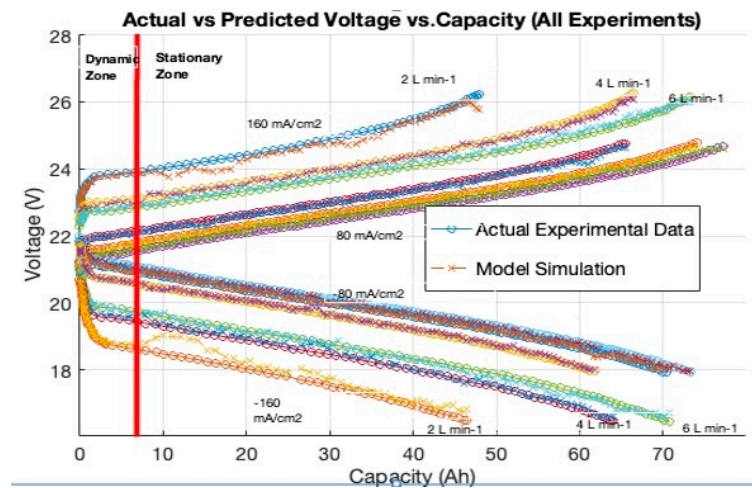


Figure 16. RF model vs. experimental data. Zhang [22].

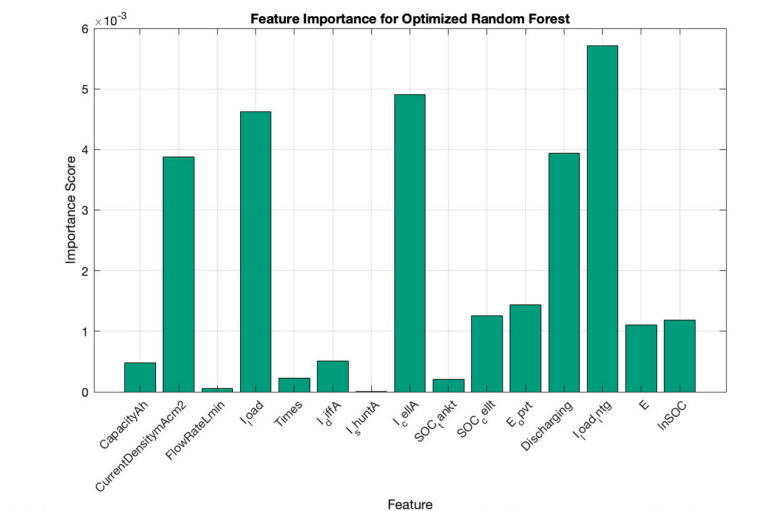


Figure 17. Feature importance on RF.

Partial Dependence Plots for Top 8 Features (Optimized Random Forest)

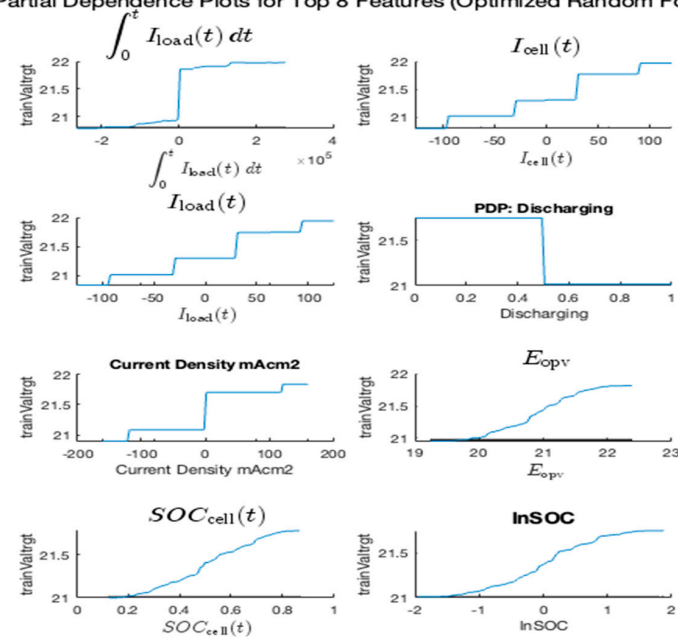


Figure 18. Partial dependence plot for top 8 features on RF.

6.2. Model Validation

Following the model construction in the previous section, several loading conditions were executed to validate the model's response. The model was set to charge with constant current to reach 85% of the tank SOC and then to discharge at constant current discharge (CCCD) to 15% of the tank SOC. The following current densities (60, 100, 200 mA/cm²) and flow rates (3, 5, 8 L/min) were used. Figures 19 and 20 show the results of the nine curves. Since the tank SOC does not depend on the flow rate, the SOC shows only the effect of the current density. It was observed that at a fixed current density under various flow rates, the reach of the maximum SOC was not a function of the flow rate. Figures 21 and 22 show this effect regardless of the NL technique used. The governing equations used to build models also confirm that the flow rates affect the charging and discharging potential voltage, but do not influence the tank SOC.

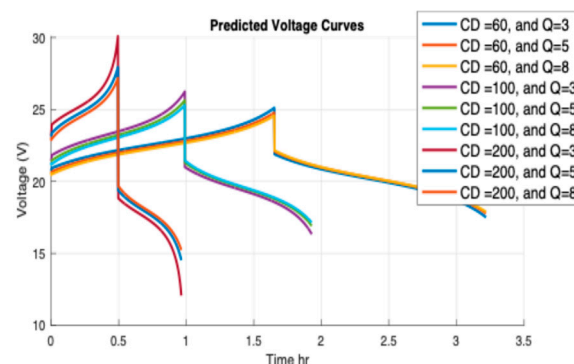


Figure 19. Voltage predictions at various CD and Q for MLR.

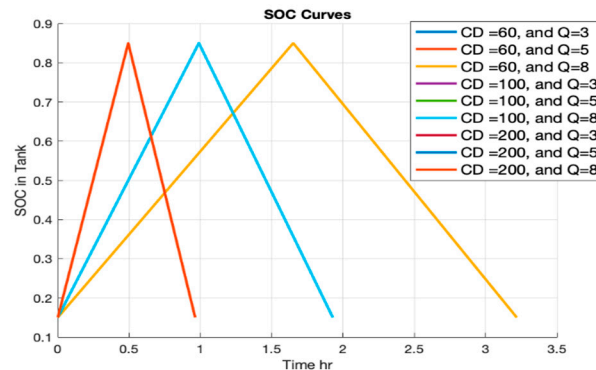


Figure 20. SOC in the tank compared to time for different current densities and flow rates MLR. The 3 curves represent the different current densities since all flow rates collapse on the same curve.

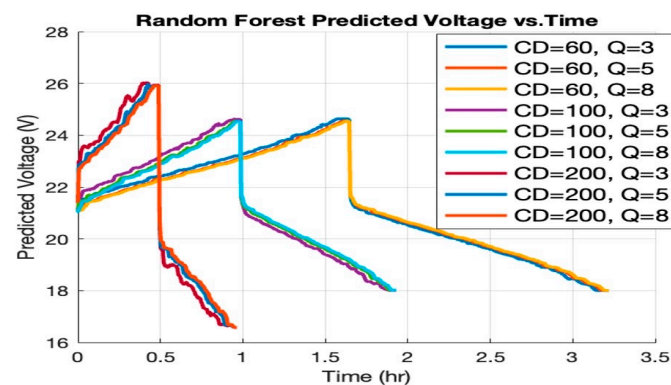


Figure 21. Voltage predictions at various CD and Q for RF.

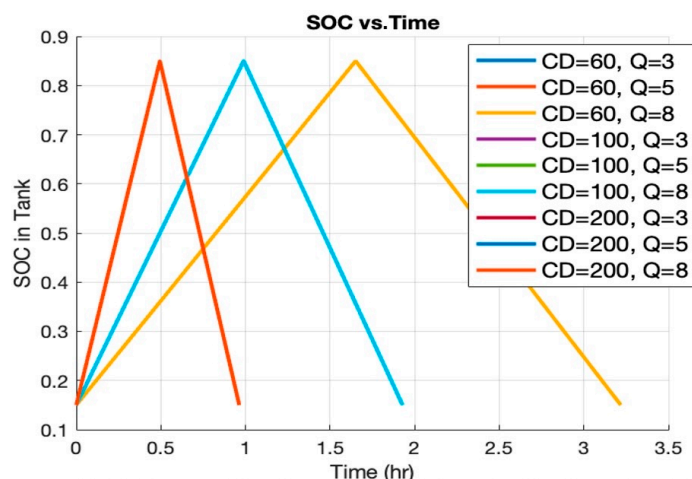


Figure 22. SOC in the tank vs. time for different current densities and flow rates for RF. The 3 curves represent the different current densities since all flow rates collapse on the same curve.

7. Analysis and Discussion

This section presents a comparative analysis of the Multiple Linear Regression (MLR) model and the Random Forest (RF) model, focusing on their performance in the VRFB system's dynamic and stationary operating regions.

ML predicted a simple equation of features and weights like Equation (10) and Table 3 with a small mean error. It requires evaluation of all features, which might become a lengthy process. The model started with 23 features, and through the process, only 8 features were not eliminated. Penalty factors L2 and L1 were used to determine and eliminate less important features.

As seen in Figure 11 in this study, the OOB error curve for the Random Forest model exhibited a rapid decline initially, stabilizing after approximately 50 trees. Beyond this point, adding more trees yielded negligible improvement in prediction performance. Based on this behavior, the number of trees was set to 50 for the final model to achieve an optimal balance between predictive accuracy.

The stabilization of the OOB error indicates that the model has effectively captured the underlying patterns in the VRFB data without overfitting. This result confirms that the Random Forest was appropriately tuned and supports the reliability of the subsequent performance evaluations and feature interpretation analyses.

In the present study, the feature importance analysis in Figure 17 identified the integrated load current ($\int_0^t I_{load}(\tau) \cdot d\tau$) as the most influential predictor, followed closely by the instantaneous cell current (I_{Cell}) and terminal load current (I_{load}). These results are physically intuitive, as both instantaneous and cumulative current directly impact voltage through resistive losses and charge depletion effects. The binary discharging state also showed high importance, indicating that charging and discharging phases significantly modulate the terminal voltage behavior. Current density (CurrentDensity mA/cm²) contributed moderately, while open-circuit voltage (E_{opvt}) and state of charge in the cell (SOC_{cellt}) captured the electrochemical aspects of the system. The natural logarithm of SOC (lnSOC) further reflected the theoretical relationship between SOC and potential, consistent with the Nernst equation. These findings confirm that the Random Forest model captured both the electrical and electrochemical dynamics that govern VRFB voltage behavior.

The PDPs generated in Figure 18 for the top eight features revealed important patterns consistent with VRFB behavior. As the integrated load current ($\int_0^t I_{load}(\tau) \cdot d\tau$) and instantaneous load current (I_{load}) increased, the predicted voltage exhibited a noticeable

decline, reflecting internal resistive losses. The PDPs for cell current (I_{Cell}) and current density (CurrentDensity mA cm²) showed similar step-like reductions in voltage with increasing current, indicating discrete operating regimes possibly corresponding to different flow rates or SOC levels. The discharging state induced a sharp voltage transition between charging and discharging phases, highlighting its critical control role. Meanwhile, SOC_{cellt} , $lnSOC$, and E_{opvt} all showed positive monotonic relationships with terminal voltage, consistent with the expected Nernstian behavior where voltage increases with SOC. These observations demonstrate that the Random Forest model effectively captured both the electrochemical and operational dependencies governing VRFB voltage dynamics.

Figures 10, 15, 18, 20 and Table 5 provided a comparative evaluation between the Multiple Linear Regression (MLR) model and the Random Forest (RF) model across different operating conditions.

A comprehensive comparison was conducted between the Multiple Linear Regression (MLR) and Random Forest (RF) models to evaluate their accuracy across different operating conditions of the VRFB system. The test dataset was segmented based on capacity, with the dynamic zone defined as capacity less than 8 A·h, and the stationary zone defined as capacity greater than or equal to 8 A·h. Error metrics including mean error (ME), mean squared error (MSE), residual sum of squares (RSS), and root mean squared error (RMSE) were computed for each region, as summarized in Table 4.

In the dynamic region, where system behavior changes rapidly, the RF model achieved better performance with an RMSE of 0.2047 V, compared to 0.2251 V for MLR, and a lower RSS (2.43 vs. 2.94). This indicates that RF more effectively captures nonlinear voltage transitions under fluctuating loads.

In the stationary region (capacity ≥ 8 A·h), both models performed similarly, but RF maintained a slight edge with an RMSE of 0.1580 V versus 0.1653 V for MLR, along with lower MSE and RSS values.

Overall, across all capacity values, the RF model demonstrated consistently better accuracy, with a total RMSE of 0.1689 V compared to 0.1796 V for MLR, and a total RSS of 7.85 versus 8.87. The ME for RF was also closer to zero, indicating minimal prediction bias.

While MLR offers the benefits of interpretability and simplicity, its linear formulation limits its ability to model nonlinearities inherent in VRFB voltage dynamics. In contrast, the Random Forest model provides superior robustness and adaptability, making it a strong candidate for real-time predictive control in flow battery systems.

These results demonstrate that the Random Forest model is well-suited for predicting VRFB terminal voltage across both dynamic and stationary operating regions. Its consistent accuracy and low error metrics highlight its value as a standalone tool for real-time system modeling. However, future work should explore the potential of deep learning models, such as recurrent neural networks (RNNs) or physics-informed neural networks (PINNs), which may offer enhanced performance in capturing long-range dependencies and complex nonlinear behavior. A comparative analysis between Random Forest and deep learning approaches would help determine the most effective modeling framework for deployment in real-time VRFB control and digital twin systems.

While temperature significantly affects VRFB behavior, this study was conducted at a constant temperature, and thermal effects were not modeled. Future work will incorporate temperature as an input feature to assess model robustness under varying thermal conditions.

8. Conclusions

This study presented a comprehensive investigation into the modeling of vanadium redox flow batteries (VRFBs) using two machine learning techniques: Multiple Linear Regression (MLR) and Random Forest (RF). Experimental data from various operating

conditions was used to train, validate, and evaluate both models. While the MLR model demonstrated reasonable predictive performance in stationary regions, it consistently underperformed in dynamic zones and showed higher overall error metrics. In contrast, the Random Forest model achieved superior accuracy across all operational regimes, including dynamic and steady-state regions, with lower mean error, residual sum of squares, and RMSE values.

These findings confirm the robustness and versatility of the Random Forest model for real-time voltage prediction in VRFB systems. Its ability to capture nonlinear system dynamics makes it a strong candidate for deployment in control systems and digital twin applications. Future research should focus on benchmarking Random Forest against advanced deep learning models, such as recurrent neural networks (RNNs) and physics-informed neural networks (PINNs), to determine the most effective framework for high-fidelity, real-time modeling of flow batteries.

Author Contributions: Conceptualization, A.A., S.A. and A.I.; methodology, S.A. and A.I.; software, A.A. validation, A.A.; formal analysis, A.A., S.A. and A.I.; investigation, A.A., S.A. and A.I.; resources, A.A., S.A. and A.I.; data curation, A.A.; writing—original draft preparation, A.A.; writing—review and editing, A.A., S.A. and A.I.; visualization, A.A.; supervision, S.A. and A.I.; project administration, S.A. and A.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data can be available upon request.

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

References

1. Skyllas-Kazacos, M.; Rychcik, M.; Robins, R.G.; Fane, A.G.; Green, M.A. New All-Vanadium Redox Flow Cell. *J. Electrochem. Soc.* **1986**, *133*, 1057–1058. [\[CrossRef\]](#)
2. Zhao, M.; Cheng, T.; Li, T.; Xie, C.; Yin, Y.; Li, X. A Long-Life Zinc-Bromine Single-Flow Battery Utilizing Trimethylsulfoxonium Bromide as Complexing Agent. *Small Methods* **2025**, *9*, e2401434. [\[CrossRef\]](#)
3. Xie, C.; Zhang, H.; Xu, W.; Wang, W.; Li, X. A Long Cycle Life, Self-Healing Zinc-Iodine Flow Battery with High Power Density. *Angew. Chem. Int. Ed. Engl.* **2018**, *57*, 11171–11176. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Weber, S.; Peters, J.F.; Baumann, M.; Weil, M. Life Cycle Assessment of a Vanadium Redox Flow Battery. *Environ. Sci. Technol.* **2018**, *52*, 10864–10873. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Wu, D.; Geng, Y.; Gao, Z.; Wu, Y. Uncovering the evolution of vanadium cycle in China during 2000–2022. *Resour. Environ. Sustain.* **2025**, *19*, 100188. [\[CrossRef\]](#)
6. Puleston, T.; Clemente, A.; Costa-Castelló, R.; Serra, M. Modelling and Estimation of Vanadium Redox Flow Batteries: A Review. *Batteries* **2022**, *8*, 121. [\[CrossRef\]](#)
7. Zhou, L.; Pan, S.; Wang, J.; Vasilakos, A.V. Machine learning on big data: Opportunities and challenges. *Neurocomputing* **2017**, *237*, 350–361. [\[CrossRef\]](#)
8. Stanton, J.M. Galton, Pearson, and the peas: A brief history of linear regression for statistics instructors. *J. Stat. Educ.* **2001**, *9*. [\[CrossRef\]](#)
9. Hoerl, A.E.; Kennard, R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12*, 55–67. [\[CrossRef\]](#)
10. Hoerl, A.E.; Kennard, R.W. Ridge regression—1980: Advances, algorithms, and applications. *Am. J. Math. Manag. Sci.* **1981**, *1*, 5–83. [\[CrossRef\]](#)
11. Hoerl, A.E.; Kennard, R.W.; Hoerl, R.W. Practical Use of Ridge Regression: A Challenge Met. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2018**, *34*, 114–120. [\[CrossRef\]](#)
12. Feig, d.g. Ridge regression: When biased estimation is better. *Soc. Sci. Q.* **1978**, *58*, 708–716.
13. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1996**, *58*, 267–288. [\[CrossRef\]](#)
14. Berkowitz, A.; Caiado, A.A.; Aravamuthan, S.R.; Roy, A.; Agar, E.; Inalpolat, M. Optimization framework for redox flow battery electrodes with improved microstructural characteristics. *Energy Adv.* **2024**, *3*, 2220–2237. [\[CrossRef\]](#)
15. Ra, N.; Bhattacharjee, A. Prediction of vanadium redox flow battery storage system power loss under different operating conditions: Machine learning based approach. *Int. J. Energy Res.* **2022**, *46*, 24441–24453. [\[CrossRef\]](#)

16. Wei, Z.; He, Q.; Zhao, Y. Machine learning for battery research. *J. Power Sources* **2022**, *549*, 232125. [[CrossRef](#)]
17. Li, W.; Limoge, D.W.; Zhang, J.; Sauer, D.U.; Annaswamy, A.M. Estimation of Potentials in Lithium-Ion Batteries Using Machine Learning Models. *IEEE Trans. Control. Syst. Technol.* **2022**, *30*, 680–695. [[CrossRef](#)]
18. Tu, H.; Moura, S.; Fang, H. Integrating Electrochemical Modeling with Machine Learning for Lithium-Ion Batteries. In Proceedings of the 2021 American Control Conference (ACC), New Orleans, LA, USA, 25–28 May 2021.
19. Ahmed, M.B.; Hassen, W.F. Optimization of a Redox-Flow Battery Simulation Model Based on a Deep Reinforcement Learning Approach. *Batteries* **2023**, *10*, 8. [[CrossRef](#)]
20. Wang, Q.; Yan, R.; Ren, L.; Qu, Z.; Jiang, Z.; Wang, Z.; Zhang, C.; Wang, J. Optimizing of working conditions of vanadium redox flow battery based on artificial neural network and genetic algorithms. *J. Energy Storage* **2024**, *100*, 113501. [[CrossRef](#)]
21. Zheng, C.; Feng, W.; Wei, Z.; Li, Y.; Iu, H.H.C.; Fernando, T.; Zhang, X. A robust machine learning-based SOC estimation approach for vanadium redox flow battery. *J. Power Sources* **2025**, *645*, 237087. [[CrossRef](#)]
22. Anirudh Narayan, D.; Johar, A.; Kalra, D.; Ardeshta, B.; Bhattacharjee, A. Machine-Learning-Based Accurate Prediction of Vanadium Redox Flow Battery Temperature Rise Under Different Charge–Discharge Conditions. *Energy Storage* **2024**, *6*, e70087. [[CrossRef](#)]
23. Xiong, B.; Tang, J.; Li, Y.; Zhou, P.; Zhang, S.; Zhang, X.; Dong, C.; Gooi, H.B. A flow-rate-aware data-driven model of vanadium redox flow battery based on gated recurrent unit neural network. *J. Energy Storage* **2023**, *74*, 109537. [[CrossRef](#)]
24. Martínez-López, J.; Portal-Porras, K.; Fernández-Gamiz, U.; Sánchez-Díez, E.; Olarte, J.; Jonsson, I. Voltage and Overpotential Prediction of Vanadium Redox Flow Batteries with Artificial Neural Networks. *Batteries* **2024**, *10*, 23. [[CrossRef](#)]
25. Kim, S.; Thomsen, E.; Xia, G.; Nie, Z.; Bao, J.; Recknagle, K.; Wang, W.; Viswanathan, V.; Luo, Q.; Wei, X.; et al. 1 kW/1 kWh advanced vanadium redox flow battery utilizing mixed acid electrolytes. *J. Power Sources* **2013**, *237*, 300–309. [[CrossRef](#)]
26. Zhang, Y.; Zhao, J.; Wang, P.; Skyllas-Kazacos, M.; Xiong, B.; Badrinarayanan, R. A comprehensive equivalent circuit model of all-vanadium redox flow battery for power system analysis. *J. Power Sources* **2015**, *290*, 14–24. [[CrossRef](#)]
27. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 26.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.