

# Energy Advances

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: A. Berkowitz, A. A. Caiado, S. R. Aravamathan, A. Roy, E. Agar and M. Inalpolat, *Energy Adv.*, 2024, DOI: 10.1039/D4YA00248B.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

June 10, 2024

Manuscript ID: YA-ART-04-2024-000248

TITLE: Optimization Framework for Redox Flow Battery Electrodes with Improved Microstructural Characteristics

### Data Availability Statement

Dear Editor:

We are pleased to submit **a revised version** of our original research article entitled “*Optimization Framework for Redox Flow Battery Electrodes with Improved Microstructural Characteristics*” by Alina Berkowitz, Ashley A. Caiado, Sundar Rajan Aravamuthan, Aaron Roy, Ertan Agar and Murat Inalpolat for your consideration for publication in the Energy Advances Journal. We have provided below a data availability statement addressing all reviewer comments indicating all codes and resources used are available for other researchers. Please let us know if you have further questions.

- (i) The original COMSOL model can be requested from the corresponding author of Knehr, et. al [28].

References:

28. Knehr, K.W., et al., *A Transient Vanadium Flow Battery Model Incorporating Vanadium Crossover and Water Transport through the Membrane*. Journal of The Electrochemical Society, 2012. **159**(9): p. A1446.

- (ii) Additionally, the optimization steps are detailed in the source code for the GBR and RFR Model. The source code can be found using the following links:

- [https://github.com/AlinaBerkowitz/VRFB-Electrode-Optimization/blob/master/2022\\_10\\_23\\_Multioutput\\_GBR\\_Optuna\\_MAPE.ipynb](https://github.com/AlinaBerkowitz/VRFB-Electrode-Optimization/blob/master/2022_10_23_Multioutput_GBR_Optuna_MAPE.ipynb)
- [https://github.com/AlinaBerkowitz/VRFB-Electrode-Optimization/blob/master/2022\\_10\\_23\\_Multioutput\\_RFR\\_Optuna\\_MAPE.ipynb](https://github.com/AlinaBerkowitz/VRFB-Electrode-Optimization/blob/master/2022_10_23_Multioutput_RFR_Optuna_MAPE.ipynb)

- (iii) This computational model, detailed in previous studies [28, 29, 55], has been validated experimentally. The transient, isothermal computation model in COMSOL Multiphysics® simulation software incorporates vanadium crossover and water transport through the membrane, along with all the corresponding losses.

References:

28. Knehr, K.W., et al., *A Transient Vanadium Flow Battery Model Incorporating Vanadium Crossover and Water Transport through the Membrane*. Journal of The Electrochemical Society, 2012. **159**(9): p. A1446.
29. Agar, E., et al., *Species transport mechanisms governing capacity loss in vanadium flow batteries: Comparing Nafion® and sulfonated Radel membranes*. Electrochimica Acta, 2013. **98**: p. 66-74.
55. Agar, E., et al., *Reducing capacity fade in vanadium redox flow batteries by altering charging and discharging currents*. Journal of Power Sources, 2014. **246**: p. 767-774.

Thank you for your consideration!



June 10, 2024

Manuscript ID: YA-ART-04-2024-000248

TITLE: Optimization Framework for Redox Flow Battery Electrodes with Improved Microstructural Characteristics

Sincerely,

Dr.Ertan Agar  
Associate Professor  
Department of Mechanical Engineering and  
University of Massachusetts Lowell  
197 Riverside Street, Dandeneau Hall 212  
Lowell, MA 01854, USA  
Phone: 978-934-5345  
Fax: 978-934-3048

Dr.Murat Inalpolat  
Associate Professor  
Department of Mechanical Engineering and  
Center for Energy Innovation  
University of Massachusetts Lowell  
197 Riverside Street, Dandeneau Hall 216  
Lowell, MA 01854, USA  
Email: Murat\_Inalpolat@uml.edu  
Phone: 978-934-2556  
Fax: 978-934-3048



# 1 Optimization Framework for Redox Flow Battery

## 2 Electrodes with Improved Microstructural

### 3 Characteristics

4  
5 **Alina Berkowitz<sup>1</sup>, Ashley A. Caiado<sup>1</sup>, Sundar Rajan Aravamathan<sup>1</sup>, Aaron Roy<sup>2</sup>, Ertan**  
6 **Agar<sup>1\*</sup>, Murat Inalpolat<sup>1\*</sup>**

7  
8 *<sup>1</sup>Department of Mechanical Engineering, University of Massachusetts Lowell, Lowell, MA 01854*  
9 *USA*

10 *<sup>2</sup>AvCarb Material Solutions, Lowell, MA 01854 USA*

11  
12  
13  
14 Submitted as a Research Paper to

15 *Energy Advances*

16 *2024*

17  
18  
19  
20 \*Corresponding author: Ertan Agar, E-mail: [Ertan\\_Agar@uml.edu](mailto:Ertan_Agar@uml.edu)

21 \*Corresponding author: Murat Inalpolat, E-mail: [Murat\\_Inalpolat@uml.edu](mailto:Murat_Inalpolat@uml.edu)



## 22 Abstract

23 This research aims to advance the field of vanadium redox flow batteries (VRFBs) by  
24 introducing a pioneering approach to optimize the microstructural characteristics of carbon cloth  
25 electrodes. Addressing the traditional challenge of developing high-performance electrode  
26 materials for VRFBs, this study employs a robust, generalizable, and cost-effective data-driven  
27 modeling and optimization framework. A novel sampling strategy using low-discrepancy Latin  
28 Hypercube and quasi-Monte Carlo methods generates a small-scale, high-fidelity dataset with  
29 essential space-filling qualities for training supervised machine learning models. This study goes  
30 beyond conventional methods by constructing two surrogate models: a random forest regressor  
31 and a gradient boosting regressor as objective functions for optimization. The integration of a  
32 non-dominated sorting genetic algorithm II (NSGA-II) for multi-objective optimization  
33 facilitates exhaustive exploration of the surrogate models, leading to the identification of  
34 electrode designs that yield enhanced energy efficiencies (EEs) under specific operating  
35 conditions. The application of NSGA-II in exploring surrogate models not only facilitates the  
36 discovery of realistic design combinations but also adeptly manages trade-offs between features.  
37 The mean pore diameter was reduced compared to the tested carbon cloth electrodes while  
38 maintaining a similar permeability value based on the results obtained using the developed  
39 algorithms. Based on this suggestion, a new type of carbon cloth electrode has been fabricated  
40 by introducing a carbonaceous binder into the woven fabric to make carbon cloths with more  
41 complex pore structures and reduced mean pore diameter. The new electrode demonstrates 24%  
42 and 66% reduction in average ohmic and mass transport resistances, respectively, validating the  
43 machine-learning recommendations. This research highlights the critical role of improved  
44 electrical conductivity and porosity in carbon materials, showing their direct correlation with



45 increased EE. Overall, this study represents a significant step forward in developing more  
46 efficient and practical VRFBs, offering a valuable contribution to the renewable energy storage  
47 landscape.

48 **Keywords:** Multiple-objective optimization, porous carbon electrode, regression, supervised  
49 learning, surrogate model, vanadium redox flow battery



## 50 1. Introduction

51 The current climate crisis has underscored the need for net-zero carbon emission policies,  
52 both in the United States and globally [1]. Following the United States' re-entry into the Paris  
53 Agreement in 2021, a long-term strategy was established with the goal of reaching net-zero carbon  
54 emissions by 2050. A critical milestone of this strategy is the 50-52% reduction in greenhouse gas  
55 emissions by 2030, necessitating a shift away from fossil fuels across all economic sectors. This  
56 decarbonization milestone is expected to increase electricity demand by approximately 50% over  
57 the next 10 years [1]. The surge in electricity demand poses significant challenges due to (i) the  
58 complex and failure-prone architecture of current electrical grid systems and (ii) the fact that 60%  
59 of electrical energy is currently supplied by fossil fuels [2-4]. Therefore, addressing the rise in  
60 electricity demand is crucial for sustaining the energy requirements necessary for a transition to a  
61 cleaner future [5].

62 In recent decades, renewable energy technologies such as wind and solar, have experienced  
63 significant market growth. Despite their increasing popularity, these low-carbon alternatives are  
64 sometimes considered unreliable for long-duration demands due to their intermittent nature [6].  
65 To address this issue and balance the energy supply and demand, cost-effective, large-scale energy  
66 storage capabilities are essential [7, 8].

67 Among the potential candidates for large-scale stationary energy storage are lead-acid  
68 batteries, lithium-ion (Li-ion) batteries, pumped storage hydropower (PSH), compressed air  
69 energy storage (CAES), and redox flow batteries (RFB) [9]. Li-ion batteries, predominant in  
70 consumer electronics and electric vehicles (EVs), face obstacles in grid-scale energy storage  
71 implementation due to their limited natural abundance and high cost for long-duration solutions  
72 [9-12]. PHS and CAES, while effective, require specific conditions for safe operation and are

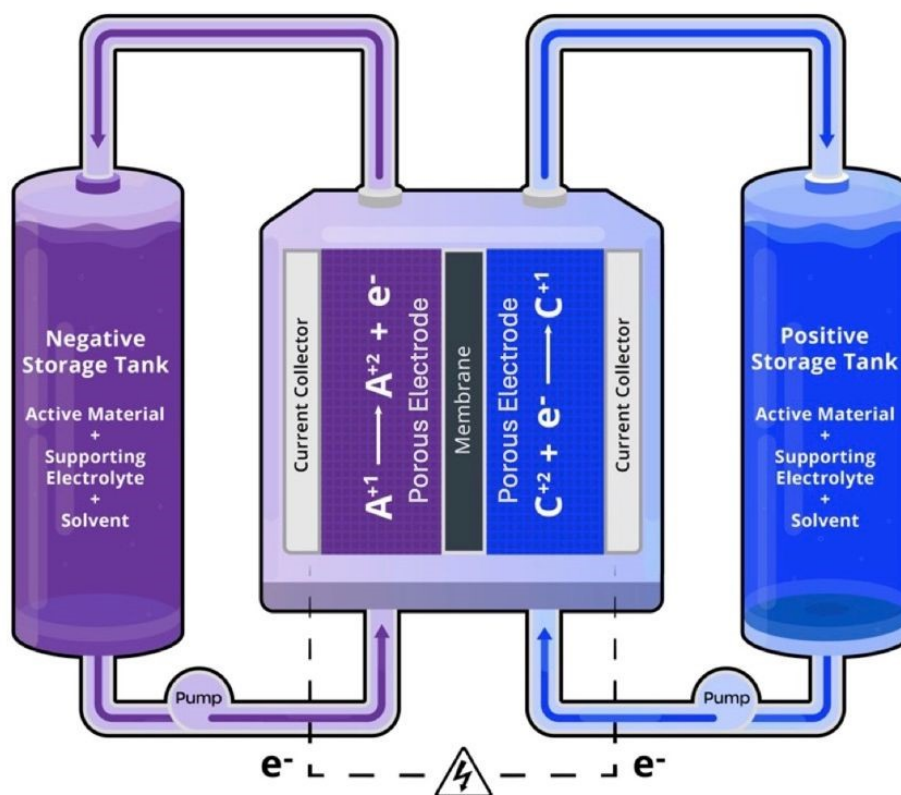


73 geographically restricted due to the necessity for suitable topography. These challenges are  
74 extensively discussed in review studies [4, 6, 8, 13, 14].

75 The search for a highly efficient, reliable, large-scale, and modular energy storage system  
76 continues to be a focus of active research [15]. Among various options, RFB technology has  
77 received considerable attention due to its scalability, efficiency, safety, and cost-effectiveness for  
78 long-duration storage [16-19]. VRFBs, where vanadium serves as the electroactive species that is  
79 dissolved in the electrolyte, are the most common RFB technology [20]. In RFBs, energy is  
80 attributed to the charged active species in the electrolytes; enabling decoupled power generation  
81 and energy storage – a key feature that underscores the promise of RFBs for grid-scale and long-  
82 duration energy storage [18, 21-23]. Figure 1 illustrates the structure of a RFB setup, with the  
83 negative and positive half-cells are separated by an ion exchange membrane. The negative and  
84 positive electrodes, critical for facilitating electrochemical reactions and providing pathways for  
85 reactant/product transport, are shown.







86  
87 Figure 1. Schematic of a RFB – A and C represent the redox active materials in the negative and  
88 positive electrolytes, respectively. In a VRFB, the negative electrolyte has  $V^{2+}/V^{3+}$  and the  
89 positive electrolyte has  $V^{4+}/V^{5+}$  redox couples.

90  
91 The major obstacle to the global implementation of VRFB technology is their high capital  
92 cost. Large-scale commercialization will remain unrealistic until the capital costs of VRFBs are  
93 reduced to meet the DOE's cost target of \$100 per kWh [24]. Performance improvement, achieved  
94 by increasing power-density and reducing resistances, will lead to reduced system costs [25, 26].  
95 Enhancing power density involves research focused on performance diagnostics at the cell level  
96 and improving the functionality and efficiency of components [27].



97 The porous electrode plays a crucial role in key functions such as facilitating ion/charge  
98 transfer, providing reaction sites for electrochemical reactions, and distributing liquid electrolytes  
99 [27-32]. Positioned adjacent to current collectors, which typically have flow channels machined  
100 within, porous electrodes benefit from interdigitated flow channel designs that increase average  
101 velocity and enhance overall battery performance [30, 33, 34]. Amongst other cell-level  
102 components, porous carbon electrodes are yet to be fully customized specifically for RFB  
103 applications. Operating conditions such as current density, flow rate, temperature, and electrolyte  
104 composition heavily impact the functionality of the porous carbon electrode, meaning that there is  
105 no singular optimal electrode design; performance will vary significantly based on operating  
106 conditions. Research aimed at improving the morphology of porous carbon electrodes has focused  
107 on maximizing active surface area for redox reactions and enhancing pathways for effective  
108 electrolyte transport [35-38].

109 Recent studies have made significant contributions to understanding and improving  
110 electrode materials for VRFBs. For example, Zhou et al. Zhou et al. explored highly permeable  
111 carbon cloth electrode materials for VRFBs, investigating the activation of carbon cloth with KOH  
112 to increase active surface area. This study demonstrated that woven carbon fiber arrangements  
113 enhance mass transport, with the KOH-activated carbon-cloth electrode achieving notable  
114 performance metrics: at a current density of 400 mA/cm<sup>2</sup>, the VRFB displayed an energy  
115 efficiency of 80.1% and electrolyte utilization of 74.6% [39]. The improved performance seen in  
116 the VRFB with carbon cloth electrodes could be attributed to the low tortuosity, low pressure  
117 drops, and high ionic conductivity associated with the larger pore sizes [39]. Furthermore, Forner-  
118 Cuenca et al. conducted a thorough investigation of three commonly used carbon fiber-based  
119 electrode materials: carbon paper, carbon felt, and carbon cloth to understand the influence of



120 carbon cloth microstructure on electrode performance through microscopic, analytical, and  
121 electrochemical methods under fixed operating conditions [40]. The research presented by Nourani  
122 et al. aligns with the conclusions made by Tenny et al., indicating that while all three carbon fiber  
123 materials have benefits and drawbacks, the structured, ordered arrangement of fibers in carbon  
124 cloth can be strategically modified or tuned [41, 42]. Thus, it can be concluded that significant  
125 performance improvements can be achieved with fabric, carbon cloth electrodes due to their  
126 tunable microstructure and ability to create structured woven patterns.

127         Previous investigations have identified key microstructural characteristics that affect the  
128 functionality of porous carbon electrodes, such as porosity, fiber diameter, and active surface area  
129 [27, 43-45]. However, the expenses associated with laboratory-scale testing are often impractical,  
130 leading most studies to include limited experimental results supplemented with synthetic data that  
131 is collected numerically or computationally via zero-to-three-dimensional modeling [46-51]. To  
132 augment sparse datasets, it has become customary to incorporate machine learning (ML)  
133 techniques to aid the data generation process. Wan et al., for instance, proposed a coupled machine  
134 learning and genetic algorithm approach to design porous electrodes for RFBs [52]. By created a  
135 dataset of 2,275 fibrous electrode structures using a stochastic reconstruction method to generate  
136 three-dimensional fibrous structures, and then applying the Lattice Boltzmann method and a  
137 morphological algorithm to calculate specific surface area and hydraulic permeability, the authors  
138 were able to use a genetic algorithm to screen and pinpoint morphological traits of 700 porous  
139 electrode candidates. Results showed that fiber diameter ( $d_f$ ) and porosity ( $\epsilon$ ) are impactful  
140 structural properties, and that tuning these properties can increase hydraulic permeability and  
141 specific surface area by 50% and 80%, respectively, thus improving overall energy efficiency [52].



142 As an emerging technology, much remains to be discovered about the electrochemical and  
143 physical properties of carbon cloth electrodes in VRFBs. This research highlights that improved  
144 electrode designs can be uncovered using interpretable ML methods to develop cost-effective and  
145 generalizable surrogate models. While the methodology is focused on vanadium chemistries, it  
146 can be extended to various flow battery chemistries, offering a versatile approach for researchers  
147 to apply to their specific conditions. This modeling and optimization framework will reveal  
148 improved electrode designs that can be mapped back to the physical domain, providing insight and  
149 quantifiable metrics that can be associated with specific and ordered fiber arrangements. The  
150 sequential steps taken to reach improved electrode properties within the modeling and optimization  
151 framework are outlined below:

- 152 • Baseline experimental microstructure characterization and performance results are obtained  
153 to gain a physical understanding of structure-property-performance linkages.
- 154 • Experimental results are used to enhance a 2D COMSOL Multiphysics® model of a VRFB.  
155 This model is used for data-generation.
- 156 • A high-fidelity sampling plan is designed with Latin Hypercube Sampling (LHS) using Quasi-  
157 Monte-Carlo methods. This modified LHS strategy uses low-discrepancy methods to  
158 uniformly distribute an arbitrarily small number of samples ( $n < 500$ ) throughout the design  
159 domain. The space-filling quality of this plan is not compromised when implemented in high-  
160 dimensions.
- 161 • The data-generation process consists of acquiring responses for each sample (electrode  
162 design) in the modified LHS plan. The charge-discharge curves produced by the  
163 computational model are used to calculate the response information for each sample. Three  
164 response values are calculated: energy efficiency (EE), coulombic efficiency (CE), and

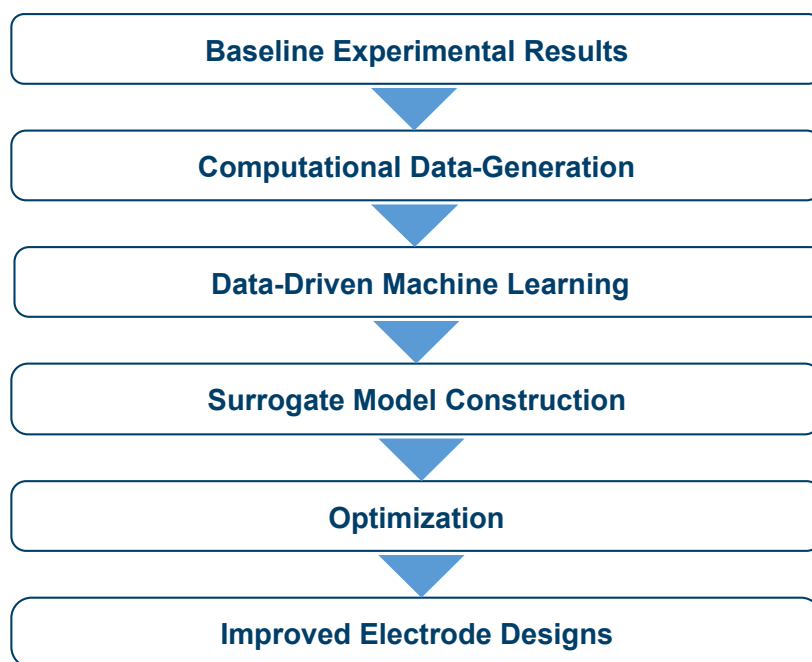


165 voltage efficiency (VE). This computational data-generation step will result in training data  
166 to support the data-driven modeling.

- 167 • Supervised regression techniques are utilized to produce an ML-based surrogate model with  
168 high prediction accuracy. Multi-output gradient boosting regression models and multi-output  
169 random forest regression models result in the lowest prediction error. A multi-output regressor  
170 is crucial to develop a surrogate model that accurately maps the relationships between the  
171 input design variables and the three target values.
- 172 • Multi-objective optimization then explores the surrogate model to obtain a Pareto set of  
173 design solutions. A nondominated genetic sorting algorithm-II (NSGA-II) is an elite multi-  
174 objective optimization algorithm that will maximize the efficiency targets while managing  
175 tradeoffs between the three target efficiencies to produce a set of the most advantageous  
176 designs.
- 177 • Combining the well-defined design constraints, accurate ML based surrogate modeling  
178 process, and optimization with NSGA-II increases likelihood that one of the designs in the  
179 Pareto set will be manufacturable.



180 The overall structure of this study and the elemental steps taken to develop this framework  
181 are highlighted in Figure 2.



182 Figure 2. Workflow diagram illustrating the multi-stage framework development process.

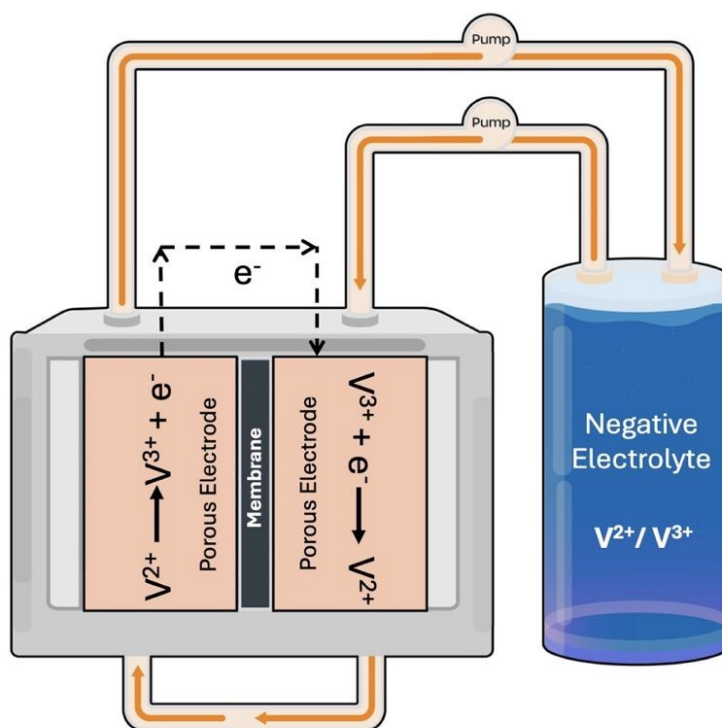
## 183 2. Methodology

### 184 2.1 Experimental Benchmarking

185 Carbon cloth electrode samples with different woven patterns, provided by AvCarb  
186 Material Solutions in Lowell, MA, are tested in the laboratory. The following AvCarb carbon  
187 fabric samples are assessed: 1698, 1615, 7497, 1185, 1698, 1070 [53]. The experimental setup is  
188 a single tank symmetric cell, where the negative electrolyte is circulated through both sides of an  
189 interdigitated flow field at 80 mL/min. The cell is assembled with zero gap architecture and a 5-  
190 cm<sup>2</sup> geometrical area. Nafion 212 is selected as the membrane which separates two layers of carbon  
191 cloth electrodes that are placed on either side of the cell. With the use of a Bio-Logic SP-240  
192 potentiostat coupled with EC-Lab software, electrochemical impedance spectroscopy (EIS) is



193 performed on a symmetric, single tank VRFB cell with electrolyte composition of 1.5M vanadium  
 194 ( $V^{2+}/V^{3+}$ ) and 3M sulfuric acid at 50% SOC. To mitigate potential oxidation of  $V^{2+}$ , nitrogen gas  
 195 is flowed constantly within the electrolyte storage tank. A  $\pm 200$  mV overpotential is applied for  
 196 24 hours with EIS experiments carried out every 4 hours [44]. With the use of data from EIS, the  
 197 resistances associated with the electrodes can be quantified and used as a benchmark for electrode  
 198 performance. Figure 3 depicts the experimental test setup. The insights gained from the baseline  
 199 experimental results are directly or indirectly mapped to global parameters in the computational  
 200 model to support and enhance the data-generation process.



201

202 Figure 1. Schematic of the experimental setup: A single tank symmetric VRFB cell.

### 203 2.2 COMSOL Multiphysics® Model for Computational Data-Generation

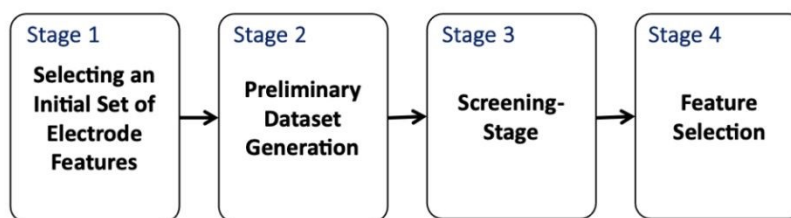
204 Due to the intensive time and resource demands of testing critical structural properties of  
 205 porous carbon electrodes, an experimentally validated computational model supports the data-



206 driven modeling approach. This computational model, detailed in previous studies [28, 29, 54],  
207 was validated experimentally, and the transient, isothermal computation model in COMSOL  
208 Multiphysics® simulation software incorporates vanadium crossover and water transport through  
209 the membrane, along with all the corresponding losses. The baseline experimental microstructure  
210 characterization and performance data enhance the computational model and guide the initial  
211 feature selection process.

### 212 *2.3 Feature Selection Process*

213 Identifying microstructural characteristics that enhance the performance of porous carbon  
214 electrodes requires extensive laboratory-scale testing. However, due to time and resource  
215 constraints, experimental data may be limited, thus serving as benchmark results that guide the  
216 incorporation of a computational model for data generation. These outcomes also play a crucial  
217 role in the feature selection process, where an initial set of design parameters or features  
218 (microstructural traits of porous carbon cloth media) that influence electrode functionality is  
219 identified. The primary stages of this process are illustrated in Figure 4.



220

221 Figure 2. Schematic outlining the four primary stages in the feature selection process.

222





### 223 2.3.1 *Stage 1: Selecting an Initial Set of Electrode Features*

224 Selecting the initial set of electrode features is heavily influenced by the experimental  
225 observations. The initial set of features will be further analyzed in Stage 2. The following  
226 measurements were obtained from the laboratory experiments and used to influence the feature  
227 selection process:

- 228 i. Pore size distribution, tortuosity, specific surface area, and porosity measurements.
- 229 ii. Electrolyte flow resistance measurements.
- 230 iii. Charge transport resistances are measurements.
- 231 iv. Mechanical properties and surface feature characterization is achieved.
- 232 v. Flow cell performance is evaluated by collecting polarization curves, charge/discharge  
233 curves for cycling analysis to determine area specific resistance (ASR) and energy  
234 efficiency (EE).

235 The initial features are displayed in Table 1 along with their units in the computational  
236 model. Each feature has a lower bound, upper bound, and recommended step size that were  
237 defined based on the baseline experimental setup and physical limitations of the materials or  
238 operating conditions that are being used in the lab. The full set of features that were initially  
239 considered and their subsequent ranges are displayed in the table below.



240 Table 1. Initial set of selected electrode features that are defined as global parameters in the  
241 computational model.

Parameter Description	Units	Lower Bound	Upper Bound	Step Size
Porosity	%	0.7	0.97	0.03
Electrical Conductivity of the Electrode	S/m	66.7	66.7	---
Current Density	A/m <sup>2</sup>	1,000	1,500	100
Permeability of the Electrode	m <sup>2</sup>	$1.0 \times 10^{-10}$	$5.0 \times 10^{-10}$	$0.01 \times 10^{-10}$
Mean Pore Diameter	m	$1.0 \times 10^{-4}$	$1.2 \times 10^{-4}$	$0.001 \times 10^{-4}$
Average Fiber Diameter	m	$1.0 \times 10^{-5}$	$2.0 \times 10^{-5}$	$1.0 \times 10^{-7}$
Reaction Rate Constant for Reaction 1	m/s	$1.0 \times 10^{-8}$	$9.0 \times 10^{-8}$	$0.1 \times 10^{-8}$
Reaction Rate Constant for Reaction 2	m/s	$1.0 \times 10^{-8}$	$9.0 \times 10^{-8}$	$0.1 \times 10^{-8}$
Flow Rate	m <sup>3</sup> /s	10	200	5
Electrical Conductivity of the Current Collector	S/m	750	1200	50

242

243

### 244 2.3.2 Stage 2: Preliminary Dataset Generation

245 Initially, a random sampling plan is generated to collect a wide range of electrode design  
246 combinations. The responses (predicted outcomes) for these initial design combinations result in  
247 a preliminary dataset with fully labeled *data-pairs*, which is then used to identify a set of critical  
248 electrode design variables and computational limitations of the Multiphysics® model. A  
249 systematic approach to collecting and processing the raw cycling data from the computational  
250 model is established in Stage 2. The computational model supplies cycling data, which refers to  
251 charging and discharging curves. The raw data output by the computational model is in the form  
252 of comma separated values that have electric potential measurements at given timestamps. A semi-  
253 automatic process is used to clean the data-files exported from COMSOL Multiphysics® [55].



254 The semi-automatic cleaning of the raw csv files involves removing unnecessary columns  
255 or default outputs from COMSOL Multiphysics® and renaming headers for integration into  
256 MATLAB® [56]. A custom MATLAB® peak finder algorithm facilitates manual peak selection,  
257 and the charging, discharging, and oscillating peak data are saved as a .mat file. A MATLAB®  
258 function then calculates the coulombic efficiency (CE), voltage efficiency (VE), and energy  
259 efficiency (EE) using the saved peak data. The efficiency values can be obtained from the cycling  
260 data and are good measures of electrode and cell performance, therefore they will be used as the  
261 target or *response* variables in the data-driven modeling process. These efficiencies can be  
262 calculated using the Equations (1) through (3), where *charging* and *discharging* are denoted by  
263 the subscripts *c* and *d*, respectively. For each cycle, the coulombic efficiency (CE) calculation  
264 requires the charging and discharging time are represented as  $t_c$  and  $t_d$ , respectively.

$$265 \quad \text{CE} = \frac{t_d}{t_c}$$

266 The voltage efficiency (VE) calculation requires the average charging voltage ( $V_{ave,c}$ ) and  
267 average discharging voltage ( $V_{ave,d}$ ) for a given cycle.

$$268 \quad \text{VE} = \frac{V_{ave,d}}{V_{ave,c}}(2)$$

269 The overall energy efficiency is represented by EE and calculated using the voltage  
270 efficiency (VE) and coulombic efficiency (CE).

$$271 \quad \text{EE} = (3)$$

### 272 2.3.3 Stage 3: Screening-Stage

273 This stage is essential to eliminate non-active and non-critical electrode properties,  
274 reducing the number of features to avoid *the curse of dimensionality* which refers to the  
275 computational costs and limitations that arise when working with high-dimensional feature spaces.



276 After generating the preliminary dataset (using a random sampling plan), a thorough sensitivity  
 277 analysis is performed to determine the significance of the initial features. Visualization techniques  
 278 such as scatterplots, histograms, kernel density estimates (KDEs) and Pearson correlation  
 279 coefficients help quantify feature-to-feature correlations and feature-to-target correlations, serving  
 280 as a statistical sanity check before deploying the ML models [57, 58].

#### 281 2.3.4 Stage 4: Feature Selection

282 Results from the screening stage quantify the impact of each feature on the voltage,  
 283 coulombic, and energy efficiencies. Operating conditions, such as current density, directly relate  
 284 to these targets; hence, including fixed operating conditions could overshadow microstructure-  
 285 performance relationships. The final set of features is selected by isolating key geometric  
 286 parameters of a porous carbon electrode and fixing the operating conditions, which can be shown  
 287 in Table 2.

288 Table 2. Final selected features and their corresponding ranges.

<b>Design Space</b>			
Fixed Operating Conditions: Current Density = 1000 [A/m <sup>2</sup> ] and Flow Rate = 3.3333E-7 [m <sup>3</sup> /s]			
<b>Index</b>	<b>Parameter Description</b>	<b>Lower Bound</b>	<b>Upper Bound</b>
<b>1</b>	Porosity	0.7	0.97
<b>2</b>	Electric Conductivity of the Electrode [S/m]	60	110
<b>3</b>	Permeability of the Electrode [m <sup>2</sup> ]	1.0 E-10	5.0 E-10
<b>4</b>	Mean Pore Diameter [m]	1.0 E-4	1.2 E-4
<b>5</b>	Average Fiber Diameter [m <sup>2</sup> ]	1.0 E-5	2.0 E-5
<b>6</b>	Cycle Number	2	6

289



290 The mean pore diameter in the Multiphysics model accounts for a 30% compression ratio.  
291 Compression and permeability are the two key components of mass transport in porous carbon  
292 electrodes. Energy efficiency will increase or decrease depending on how well the geometrical  
293 features of the carbon cloth electrode perform.

## 294 *2.4 Sampling Plan Design*

### 295 *2.4.1 Latin Hypercube Sampling using Quasi Monte-Carlo Methods*

296 A common sampling strategy for surrogate modeling is Latin Hypercube Sampling (LHS).  
297 This plan takes an arbitrary number of samples and distributes them uniformly throughout the  
298 design space [59]. The LHS plan proves to be successful for lower dimensional problems. The  
299 LHS plan is expensive and often inefficient for multi-dimensional problems as a minimum number  
300 of samples,  $n^d$ , must be specified for each dimension. As the number of dimensions increases, the  
301 minimum number of required samples will increase to uniformly distribute samples throughout  
302 each dimension of the feature space [59-61]. The optimal space-filling properties that LHS plans  
303 achieve in a single dimension can be maintained in multiple dimensions by combining the LHS  
304 strategy with Quasi-Monte-Carlo methods, also referred to as low-discrepancy sampling methods,  
305 [59, 62]. The minimum number of samples needed for the modified LHS plan will not necessarily  
306 increase if the number of features increases.

307 LHS with Quasi-Monte-Carlo methods is used to create a set of samples that are uniformly  
308 distributed throughout the multi-dimensional feature space. This plan randomly selects  $n$   
309 uniformly distributed points within the constrained feature space. The constraints refer to the lower  
310 and upper bounds for each feature. Reducing the number of samples will reduce computational or  
311 experimental expenses but may lead to a less robust training dataset. The following notation can  
312 be used to represent the sampling plan, where  $m$  is the features and  $n$  is the number of samples.



$$X = \begin{bmatrix} x_1^1 & x_1^2 & \dots & x_1^m \\ x_2^1 & x_2^2 & \dots & x_2^m \\ \vdots & \vdots & \ddots & \vdots \\ x_n^1 & x_n^2 & \dots & x_n^m \end{bmatrix} \quad (4)$$

$$X = [x^1 \quad x^2 \quad \dots \quad x^m] \quad (5)$$

$$X \in \mathbb{R}^{n \times m} \quad (6)$$

$$x^i \in \mathbb{R}^n \quad (7)$$

### 317 2.5 Supervised Machine Learning Techniques

318 Supervised ML strategies, also referred to as instance-based learning, are employed to  
319 model the dynamic behavior of VRFB system. The supervised ML algorithm learns from the data  
320 that is generated from the computational model. The model complexity is then increased to develop  
321 multiple-output regression models that accurately imitate system behavior with respect to three  
322 target values (EE, CE, VE) as opposed to the single output energy efficiency models.

323 All machine learning models aim to learn a function,  $f$ , that maps observed data,  $x$ , to the  
324 corresponding response,  $y$ .

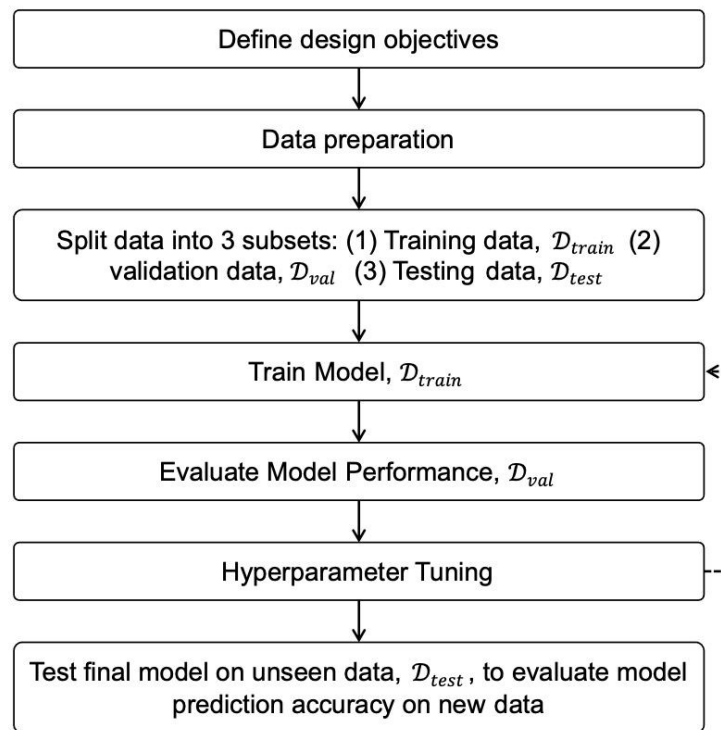
$$325 \quad f: x \rightarrow y$$

326 Typically, engineering design problems are multi-variate, *meaning they contain multiple*  
327 *design variables*. Design variables are also commonly called *features* or *predictors*. This results in  
328 a design variable vector, also called a feature vector, where the number of features is denoted as  
329  $m$ . The number of features also defines the dimensionality of the problem where a  $m$ -dimensional  
330 problems contain  $m$  number of features.

331 Tree-based methods are based on an application called *decision-trees*, which are algorithms  
332 that can solve both classification and regression problems for single output and multiple output  
333 problems [57]. The following characteristics of tree-based methods make them desirable for the  
334 application of this paper; (1) Tree-based methods are interpretable and typically do not require



335 feature standardization since these methods do not weigh the magnitude of feature vector values ,  
 336 (2) outliers are managed well in both the target and the features space, (3) these methods are able  
 337 to be computationally scaled for larger datasets, (4) tree-based methods provide a good balance  
 338 between model complexity and model [63]. Figure 5 illustrates the phases of building a ML model.



339  
 340 Figure 3. Machine learning workflow.

341 The generated data is broken into subsets for training, validating and testing the ML model.  
 342 Figure 6 depicts how the dataset is typically split into the three subsets. Before tuning the ML  
 343 model on all the data, it is customary practice to split the data into training, validation, and testing  
 344 sets (samples of the larger dataset). The model trains on approximately 70% of the data. The model  
 345 is then validated using the validation subset of data that it has never seen before. The process of  
 346 training and validation is repeated for a defined number of iterations.





Figure 4. Training, validation, and testing split.

Occasionally, when ML models learn from small datasets (<1000), hyperparameter tuning can quickly lead to overfitting or underfitting. This is especially true for tree-based methods trained on small datasets.  $k - fold$  cross validation is used in the hyperparameter tuning stages to prevent overfitting.  $k - fold$  cross validation repeats the process of splitting the dataset into training, validation, and testing five times; each iteration uses a different subset of data for training and validation. This method of cross validation assures that your dataset is generalizable. Referring to the ML flow diagram, the dataset is split into a training, testing, and validation data set. The  $k$  in  $k - fold$  cross validation refers to the number of validation folds (typically 5 or 10).

### 2.6 Machine Learning Model Evaluation

The evaluation metric best suited for the applications in this paper is Mean Absolute Percentage Error (MAPE) which is defined in the following equation, where  $y_i$  is the predicted value of the  $i^{\text{th}}$  sample and  $n_{\text{samples}}$  is the number of samples [64].

The *mean absolute percentage error* (MAPE) is another risk metric used to evaluate regression problems. In the Python module scikit-learn, MAPE falls between zero and one. Values outside of this range suggest that the model is overfitting, underfitting, or the selected model may not be appropriate for the dataset and other models should be explored [64].

$$MAPE(y, \hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} \frac{|y_i - \hat{y}_i|}{\max(\epsilon, |y_i|)} \quad (10)$$





366 This equation will be used in the model evaluation process to determine how well the ML  
367 model will respond to new or unseen data. Lower errors mean that it is highly probable that the  
368 model will make good predictions on new data. High error metrics suggest that it is unlikely that  
369 the ML model is making accurate predictions on new data.

## 370 *2.7 Constructing a Machine Learning Based Surrogate Model*

371 Surrogate modeling serves as a vital tool for approximating complex, non-interpretable  
372 (black box) ML or deep learning models, providing an affordable and interpretable alternative,  
373 denoted  $\hat{f}$ . In the realm of engineering surrogate modeling, the strategy involves employing a  
374 comprehensible ML model to approximate an unknown function  $f$ . This approximation is achieved  
375 using a judiciously chosen subset of high-fidelity samples that effectively encapsulate the  
376 intricacies of the design space.

377 The machine learning methods utilized in surrogate modeling are not universally  
378 interpretable, and their complexity tends to escalate with an increasing number of features. Despite  
379 this, the application of surrogate models remains crucial in situations where understanding the  
380 underlying mechanisms is paramount.

381 Akin to the steps involved in developing a conventional ML model, surrogate modeling  
382 comprises several integral stages, each contributing to the overall efficacy of the process.

### 383 *2.7.1 Computational Data Collection Benchmarked with Physical Laboratory Results*

384 The initiation phase involves the collection of computational data, aligning it with physical  
385 laboratory results for benchmarking. This ensures a congruence between simulated and real-world  
386 outcomes, laying a robust foundation for subsequent modeling.



### 387 2.7.2 *Preliminary Data-Generation and Feature-Screening*

388 Following data collection, preliminary steps encompass data generation and feature  
389 screening. This involves generating an initial dataset and screening features to identify those  
390 wielding significant influence on the target function, thereby streamlining subsequent analyses.

### 391 2.7.3 *Data Analysis and Final Feature Selection*

392 A meticulous data analysis procedure is then conducted to further refine the feature set.  
393 This stage aims to discern the most pertinent features, optimizing the model's accuracy and  
394 interpretability.

### 395 2.7.4 *Sampling Plan Design*

396 A critical aspect of the surrogate modeling process involves the design of an effective  
397 sampling plan. This entails planning the selection of data points, ensuring a judicious  
398 representation of the design space while maintaining computational efficiency.

### 399 2.7.5 *Data-Generation*

400 Subsequent to the sampling plan, additional data points are generated to augment the  
401 dataset. This augmentation bolsters the model's capacity to capture complex relationships within  
402 the design space.

### 403 2.7.6 *Machine Learning Modeling and Evaluation*

404 The crux of surrogate modeling lies in the application of ML techniques. Models are trained  
405 using the collected data to approximate the target function. Rigorous evaluation ensures the  
406 resultant model's accuracy and reliability.



### 407 2.7.7 *ML Model Selection and Surrogate Model Construction*

408 The concluding phase involves the judicious selection of a suitable ML model, followed  
409 by the construction of the surrogate model ( $\hat{f}$ ). This step is pivotal in developing an interpretable  
410 model that effectively approximates the complex behavior of the original non-interpretable model  
411  $f$ .

## 412 2.8. *Multi-Objective Optimization to find a Pareto Set of Improved Electrode Designs*

### 413 2.8.1 *Multi-Objective Optimization and Pareto Sets*

414 After constructing an efficient and reliable ML based surrogate model, multi-objective  
415 optimization is employed to explore the surrogate model to find a Pareto set of optimal electrode  
416 designs. As discussed earlier, multi-objective optimization problems often have competing  
417 objectives. This problem maximizes VE, EE, and CE, which are calculated according to Equations  
418 1-3. Next, the reasoning behind why a Pareto set of solutions is necessary for this specific problem  
419 is explained using a select few design parameters. For example, previous studies proved that cell  
420 efficiency can be improved by maximizing porosity and maximizing active surface area. With that  
421 said, increasing porosity inherently decreases active surface. This is due to the competing  
422 properties of the parameters causing a necessary tradeoff between the two. An increased porosity,  
423 while decreasing the mass transport resistance, has an indirect relationship with surface area  
424 causing an increased charge transfer resistance. Multi-objective optimization will account for the  
425 interactions between porosity, energy efficiency, coulombic efficiency, and voltage efficiency and  
426 provide a set of solutions that balances the tradeoffs between porosity and surface area.

### 427 2.8.2 *Non-Dominated Sorting Genetic Algorithm II (NSGA-II)*

428 A non-dominated genetic sorting algorithm II (NSGA-II) is a variation of the genetic  
429 algorithm that is best suited to find a Pareto set of optimal solutions for multi-objective



430 optimization problems. Similar to a traditional genetic algorithm, NSGA-II will begin with an  
431 initial population. The best design combinations in the initial population will move onto the second  
432 generation and this process will repeat until convergence. The main nuance to this approach is that  
433 each design combination is evaluated on its fitness score and the combinations are also ranked  
434 based on their location in the design domain. This eliminates the chance of having repetitive  
435 offspring in future generations as well as assuring that the entirety of the design space is explored.

### 436 *2.9 Fabricated Electrodes and Their Performance Characterization*

437 The microstructure of the base carbon cloth electrode (AvCarb 1071 HCBA) displays a  
438 bi-modal pore size distribution [44], which is a critical feature allowing for lower mass transport  
439 resistances. Because of this, power density is improved, and pumping losses are reduced [40].  
440 There are negligible effects of pumping power losses on the cell, leading to the omission of their  
441 effects in efficiency calculations. Larger pores of the electrode are responsible for delivering the  
442 electrolyte through convection, resulting in lower pumping power losses and the smaller pores  
443 allow for electrolyte diffusion to active sites which enhances reaction kinetics [40, 42]. For this,  
444 AvCarb 1071 HCBA is chosen as the baseline for which machine learning suggestions will be  
445 implemented on. Based on the recommendations from the ML-based surrogate model, the binder-  
446 coated electrode (AvCarb T2314B) is prepared by adding a carbonaceous, porous binder layer to  
447 both sides of the AvCarb 1071 HCBA electrode. The electrodes, initially un-activated, are  
448 activated by heating in a furnace at 425°C for 24 hours.

449 For evaluating the performance of the binder coated electrode (AvCarb T2314B),  
450 electrochemical testing is performed and compared amongst the baseline results for AvCarb 1071.  
451 The experimental setup uses a symmetric RFB cell with a 40 mL single tank of electrolyte which  
452 has been described in detail in the subsection “2.1 Experimental Benchmarking of the



453 Computational Model” of the Methodology section. One experiment performed consists of the  
 454 baseline electrode (AvCarb 1071 HCBA), and the second experiment utilizes a binder-coated  
 455 electrode (AvCarb T2314B). The overall compression ratio of the cell is around 41% for the  
 456 experiment consisting of 1071 HCBA and around 49.7% when T2314B electrodes are used. EIS  
 457 results are analyzed to quantify the resistance for direct comparison of electrode performance  
 458 within a VRFB.

### 459 3. Results and Discussion

#### 460 3.1 Selected Features

461 After identifying the initial set of features and completing the preliminary dataset  
 462 generation, the final set of features is selected based on their impact on electrode functionality as  
 463 well as the computational feasibility. The final set of features along with their lower and upper  
 464 bounds are displayed in Table 3. Note that the fixed operating conditions in this study are current  
 465 density set to be 1000 A/m<sup>2</sup> and flow rate set at 3.33E-7 m<sup>3</sup>/s.

466 Table 3. Final set of six selected features and their corresponding bounds.

Parameter Description	Lower Bound	Upper Bound
Porosity	0.7	0.97
Electric Conductivity of the Electrode (S/m)	60	110
Permeability of the Electrode (m <sup>2</sup> )	1 E-10	5 E-10
Mean Pore Diameter (m)	1 E-4	1.2 E-4
Average Fiber Diameter (m <sup>2</sup> )	1 E-5	2 E-5
Cycle Number	2	6

467

468 The bounds can also be written as shown in Equation 13 using porosity as an example.

469 
$$\sigma^e \in [0.7, 0.97] \quad (13)$$



470 The six features and their bounds shown in Table 3 describe the design domain. Please note  
471 that cycle number is an output of the computational model and may not be directly perceived as a  
472 statistical feature. However, it was used in training the ML algorithms and was deemed useful.  
473 Recalling that each feature,  $x^i$ , typically has lower and an upper bound constraints that needs to be  
474 specified, the feature vector,  $x$ , must be within the ML domain, represented by  $\mathcal{D}$ , which is a subset  
475 of all real numbers.  $\mathcal{D}$  is also a vector with  $m$  number of elements (features). This explanation is  
476 clearly summarized in Equation 14 [65].

$$477 \quad x^i \in \mathcal{D} \subset \mathbb{R}^n (14)$$

478 There are six selected features, but permeability is also not included in the sampling plan  
479 design since the permeability is calculated for each sample using the Carman-Kozeny equation.  
480 This equation relates the morphological parameters of porosity and average fiber diameter for each  
481 sample to calculate the permeability and can be shown below in Equation 15 [66].

$$482 \quad \kappa = \frac{d_f^2 \varepsilon^3}{K_{ck}(1 - \varepsilon)^2} (15)$$

483 The response value of cycle number for each electrode design is recorded although it is not  
484 included in the sampling plan since it is technically a response that is output by the computational  
485 model. The porosity can be raised by the mean pore diameter depending on the pore sizes and the  
486 pore distribution in the material. Higher porosity can also be achieved by decreasing the fiber  
487 diameter to increase active surface area.

### 488 *3.2 Latin Hypercube Sampling Plan using Quasi Monte-Carlo Methods*

489 The final statistical sampling plan consists of two hundred samples. This space filling  
490 sampling plan evenly distributes the two hundred samples throughout the design space. There are  
491 six selected features, but permeability is excluded from the sampling plan design as it is calculated



492 using the other two features. Referring back to Equation 4, the sampling plan can be described  
 493 using the matrix below where  $m = 5$  and  $n = 200$ .  $m$  refers to each sample (observation) in the  
 494 sampling plan.

$$495 \quad X = \begin{bmatrix} x_1^1 & x_1^2 & x_1^3 & x_1^4 & x_1^{m=5} \\ x_2^1 & x_2^2 & x_2^3 & x_2^4 & x_2^{m=5} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n=200}^1 & x_{n=200}^2 & x_{n=200}^3 & x_{n=200}^4 & x_{n=200}^{m=5} \end{bmatrix} \quad (16)$$

496 Each sample in the LHS plan is an electrode design. Table 4 clearly outlines the first four  
 497 electrode designs. For data visualization and ML model interpretability purposes, the mathematical  
 498 notation displayed in Table 5 is used to describe the features and targets.

499 Table 4. The first four electrode designs created from the LHS sampling plan.

Sample	$\sigma^e$	$\varepsilon$	$\kappa$	$d_f$	$d_p$
$m = 1$	67.3	0.93	1.7E-10	1.4E-5	1.4E-4
$m = 2$	86.1	0.82	3.6E-11	1.9E-5	1.2E-4
$m = 3$	61.3	0.88	7.7E-11	1.8E-5	1.0E-4
$m = 4$	107.5	0.77	1.4E-11	1.7E-5	1.2E-4
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$m = 200$	103.9	0.95	3.3E-10	1.4E-5	1.3E-4

500  
 501 Table 4 provides clear examples of what each electrode design (*sample*) from the LHS plan  
 502 will look like. Each sample,  $n$ , has a selected value for electrical conductivity, porosity,  
 503 permeability, average fiber diameter, and mean pore diameter.



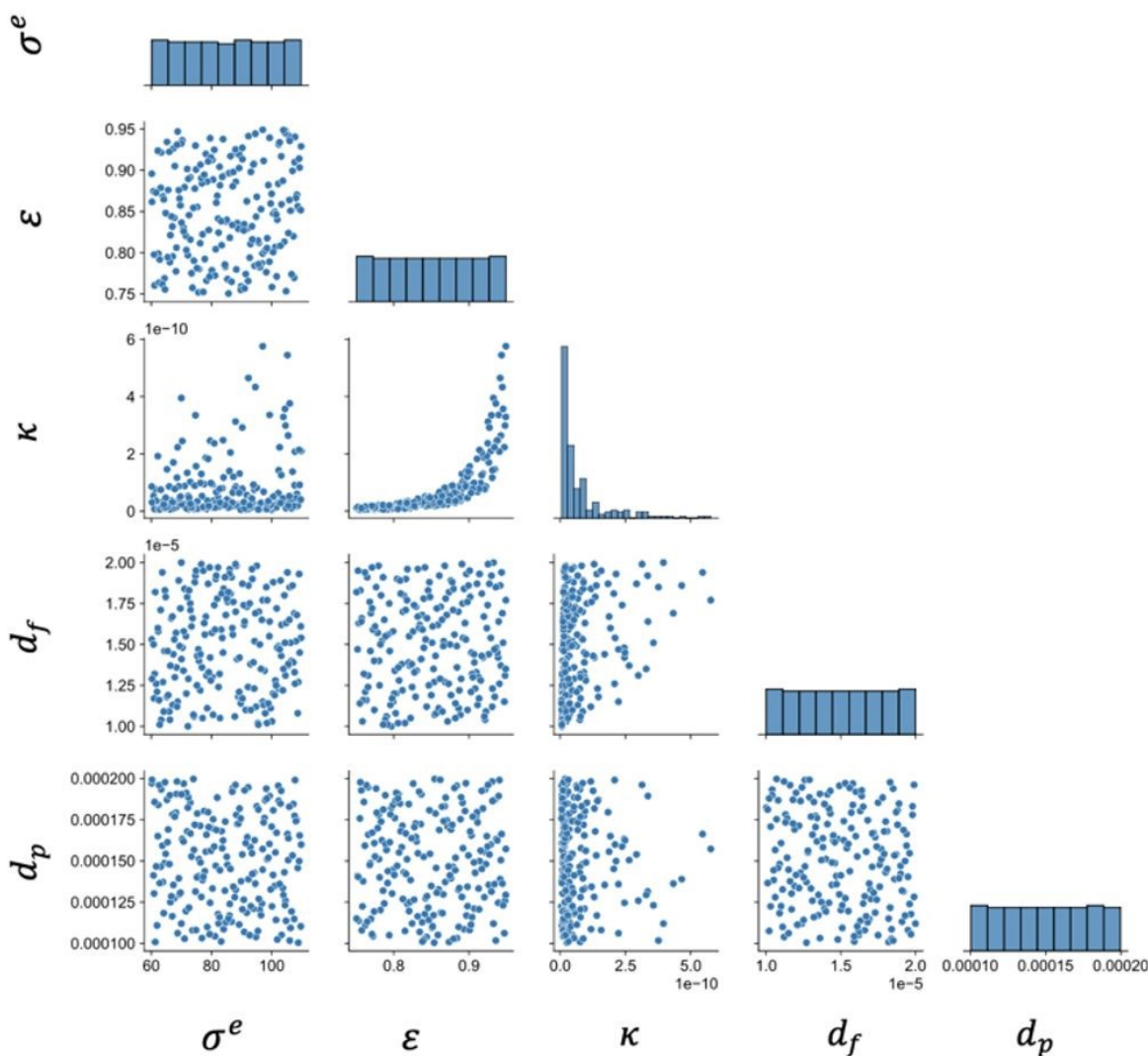
504 Table 5. The notation used to define the electrode features and targets.

Feature and Target Names	Symbol
Electrical Conductivity of the Electrode	$\sigma^e$
Porosity	$\varepsilon$
Permeability	$\kappa$
Average Fiber Diameter	$d_f$
Mean Pore Diameter,	$d_p$
Voltage Efficiency	VE
Coulombic Efficiency	CE
Energy Efficiency	EE

505  
506 The selected values fall between the lower and upper bounds assigned to each feature  
507 (shown in Table 3). The resulting distribution of values that the sampling plan created for each  
508 feature is shown in the Pairplot in Figure 7. A Pairplot, or matrix of scatterplots, is used to show  
509 the distribution of samples for the features. The LHS plan using quasi-Monte-Carlo methods  
510 ensures that a representative subset of values is selected for each feature. The limited white space  
511 in each scatterplot in Figure 7 shows that the sampling plan selected a representative subset of  
512 values for each feature. The permeability is calculated from  $d_f$  and  $\varepsilon$ . The script to generate the  
513 LHS plan with QMC methods considered four features; permeability is calculated using the  
514 Carman-Kozeny equation [66]. Therefore, sparse scatterplots in Figure 7 can be attributed to  
515 permeability being a function of porosity and average fiber diameter.







516

517 Figure 5. Feature distribution of the 200-point Latin hypercube sampling plan generated using

518

QMC methods.

519 Table 4 displays the design combinations from the LHS sampling plan, which are displayed

520 in Figure 7. The numerical values for each of the five features for the first four electrode design

521 combinations are displayed.

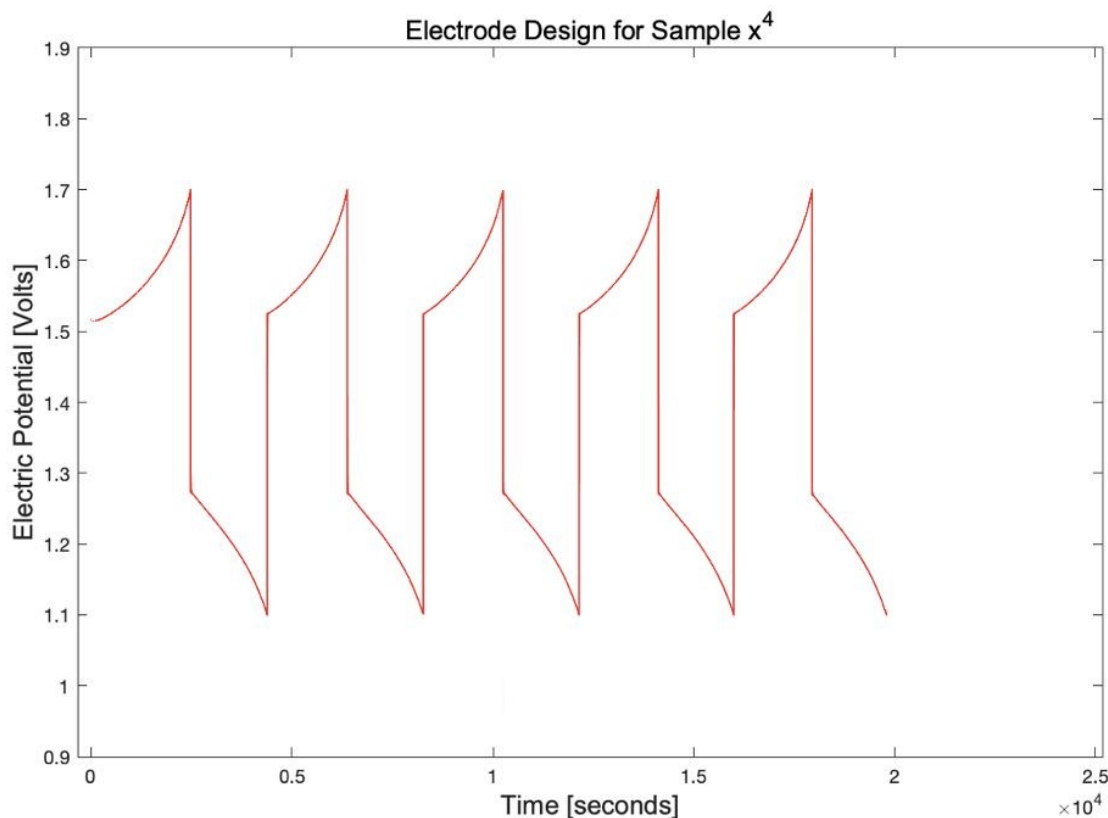


## 522 3.3 Dataset Generation

### 523 3.3.1 Computational Data-Generation, Results and Charge-Discharge Curves

524 The computational time required to obtain cycling data for a single electrode design can  
525 range from 60 to 180 minutes. Simulating 200 samples would take over 300 hours to complete.  
526 An ample amount of time has been invested into collecting response results for all two hundred  
527 electrode designs. Due to the time-consuming nature of computational data-generation, an *active*  
528 *learning* approach is taken as data is collected. *Active learning* refers to re-training the ML models  
529 as the dataset is enriched with more samples [57, 63, 67-69]. Since each sample has between 2 and  
530 6 cycles and each cycle has three target values (VE, CE, EE), the final database has 387 fully  
531 labeled examples to support the data-driven modeling approaches. For each sample, the raw  
532 cycling data produced by the computational model is cleaned, renamed, and imported into  
533 MATLAB for plotting. Figure 8 displays the charge-discharge curve produced when the  
534 computational model parameters are modified to match the electrode design specification of  
535 sample 4 (electrode design for sample 4 is shown in Table 4). The charging, discharging, and  
536 oscillating peaks are selected in MATLAB and the target values (EE, VE, CE) are calculated for  
537 each cycle.





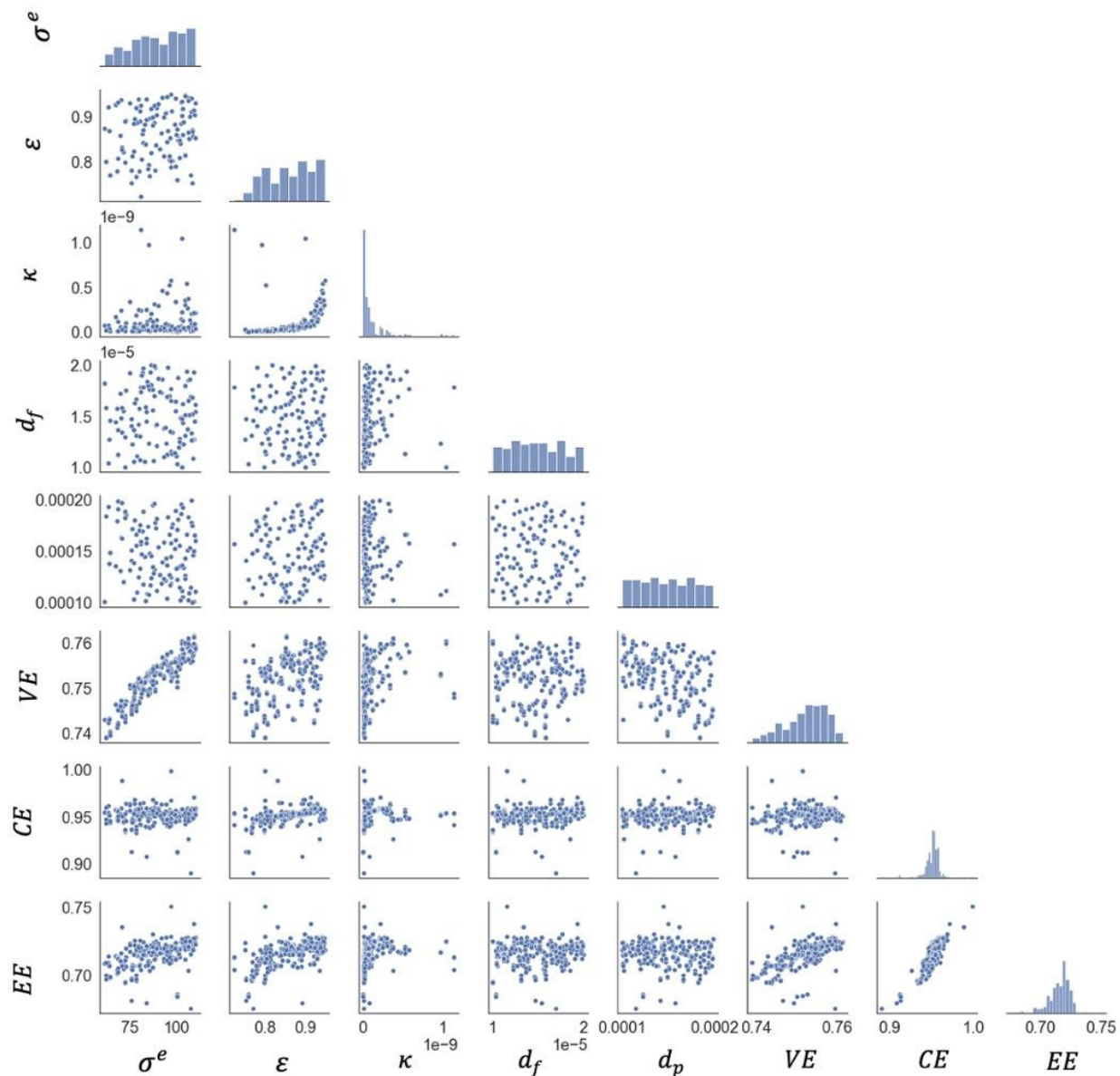
538

539 Figure 6. Charge-discharge curve plotted in MATLAB (refer to Table 4 for the electrode design  
540 details for sample 4 that produced this cycling curve).

### 541 3.3.2 Statistical Analysis and Data Visualization

542 A Pairplot of the 387 fully labeled examples is provided in Figure 9 which also includes  
543 cycle number, and the distribution of each target efficiency. The diagonal of the Pairplot contains  
544 histograms showing the distribution of collected values for each feature. Similar to Figure 7, the  
545 axes labels are based on the mathematical notation displayed in Table 5.





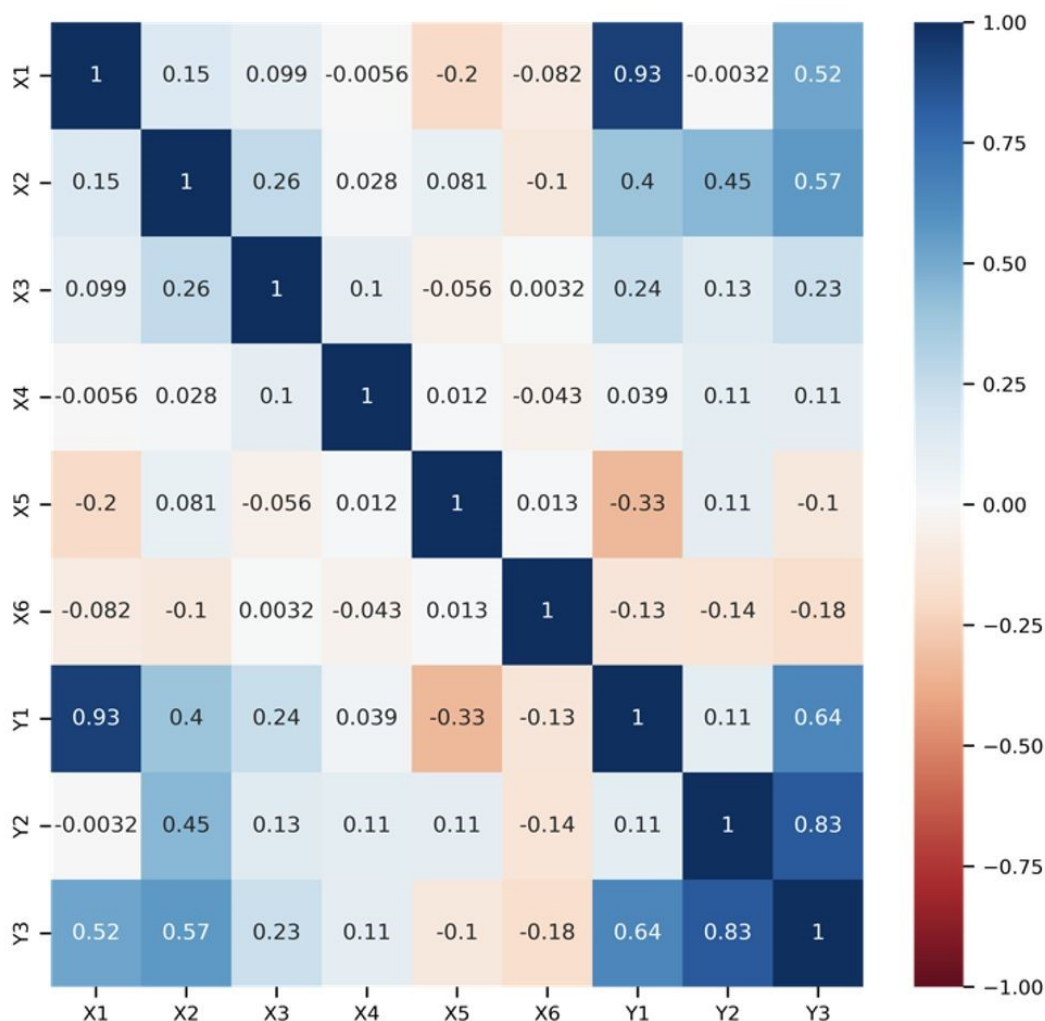
546

547 Figure 7. Pairplot (matrix of scatterplots) showing the feature and target distributions for the  
 548 collected data from the sampling plan.

549 The Pearson correlation heatmap show that  $CE$  and  $VE$  are positively linearly correlated to  
 550  $EE$  with a correlation coefficient of  $r = 0.85$  and  $r = 0.56$ , respectively. All three efficiency values  
 551 are linearly related to porosity. The voltage and coulombic efficiency trends can be summarized  
 552 by the energy efficiency target. The one exception is that  $VE$  is linearly related to  $\sigma^e$  with  $r = 0.93$ .



553 The Pearson correlation coefficients summarized in Figure 10 offer a  
 554 thorough understanding of the design space and will guide machine learning model selection. The  
 555 lack of linear feature-target correlations indicates that simple linear regression techniques are  
 556 unable to capture the complex non-linear relationships.



557

558

Figure 8. Pearson correlation coefficient heatmap.

559



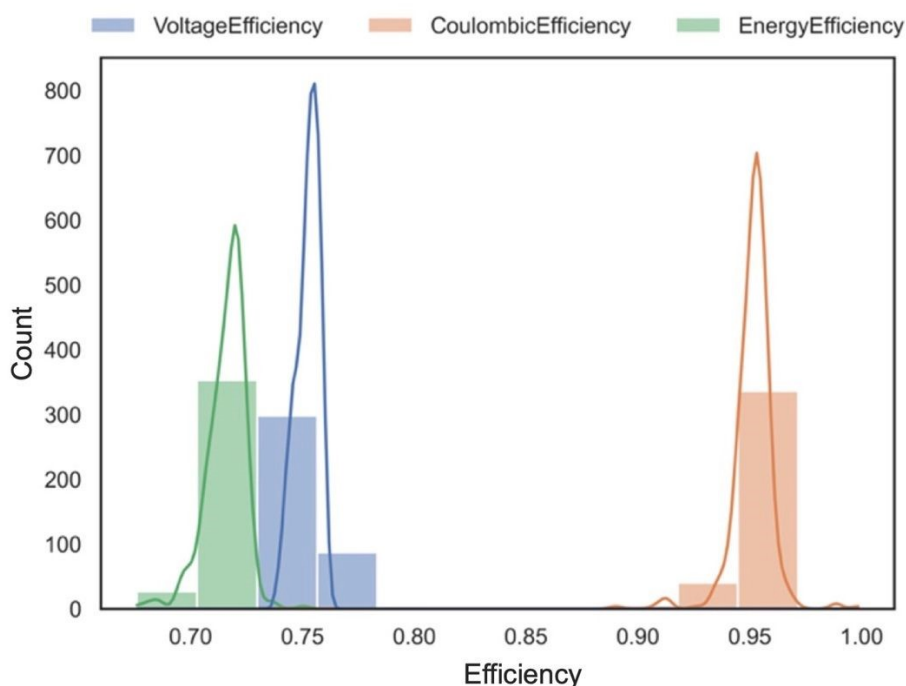
## 560 3.3.3 Understanding the Generated Response Data (EE, CE, VE)

561 Generated response data, shown in Figure 11, highlights the similarities and differences  
 562 between the ranges of values for each response variable. The range of values obtained for CE is  
 563 between 90-98%, which is comparable to the experimentally obtained values. The minimum and  
 564 maximum efficiency values for the three target variables is also outlined in Table 6.

565 Table 6. Minimum and maximum efficiency values for each target.

	VE	CE	EE
Minimum	78.94 %	89.02 %	67.61 %
Maximum	76.15 %	99.85 %	75.04 %

566

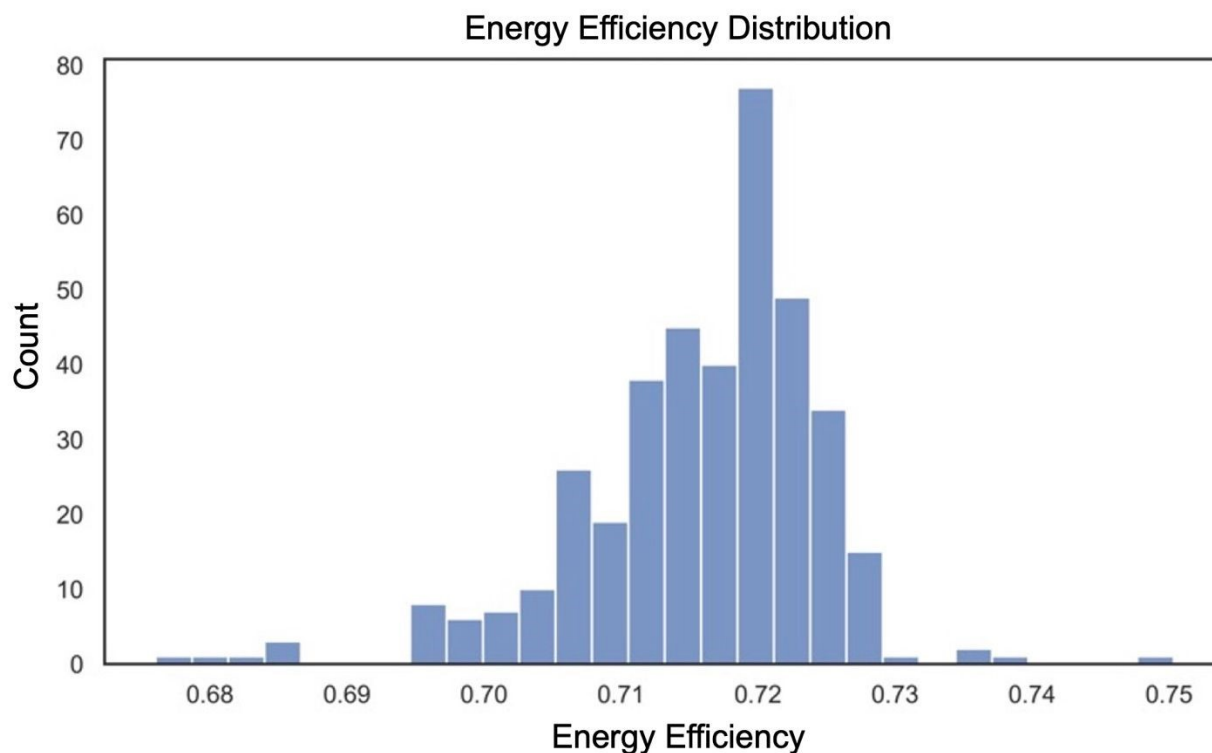


567 Figure 9. Histogram and kernel density estimates (KDEs) containing the distribution of values  
 568 collected for the three response variables, VE, CE, and EE.  
 569

570



571 A more refined, higher resolution histogram for the EE has been provided below in Figure  
572 12. The relatively wide range of values (ranges from 0.68 to 0.75) obtained is an indication of the  
573 relatively large potential improvements on the energy efficiency that can be obtained with an  
574 optimized electrode design.



575  
576 Figure 10. Energy efficiency distribution emphasizing the percentage range for improvement.

### 577 3.4 Machine Learning Model Development

#### 578 3.4.1 Machine Learning Model Selection

579 Initially, since the EE target contains the VE and CE information, single output machine  
580 learning models were trained to determine what models are suitable for this problem. This  
581 approach also reduces the complexity of the model which in turn reduces the computational power  
582 necessary to train, validate, and test each model. A preliminary test was performed using the  
583 *Automated Regression Model Selection with Bayesian Optimization* tool in MATLAB, fitting the





584 regression models to the single response value of energy efficiency. This tool automatically trains  
 585 and evaluates several regression models with various hyperparameters and returns corresponding  
 586 models and hyperparameters with the highest prediction accuracy. The computing time is  
 587 approximately 45 minutes. This process pinpoints appropriate regression models to fit this dataset  
 588 as opposed to manually evaluating every regression algorithm. Although the automated regression  
 589 model selection with Bayesian optimization is performed as a multivariate regression problem with  
 590 a single output, the single output of EE encompasses the CE and VE information therefore no  
 591 information is lost. The results suggested that tree-based ensemble methods, specifically random  
 592 forests, would be the most suitable for this dataset. Therefore, the ML models selected for further  
 593 investigation are random forest regressors and gradient boosting regressors, both of which are tree-  
 594 based ensemble methods.

#### 595 3.4.2 Comparing Feature Importance Scores for Single and Multiple Output Random Forest 596 Regression Models

597 Once the single output and multiple output random forest regressors (RFRs) are trained  
 598 and evaluated, the feature importance scores are found. Table 7 outlines which target variables  
 599 each ML model was trained on. For example, ML Model 1 is trained to predict VE. Model 4 is the  
 600 multiple output model which is trained on all three target variables (VE, CE, and EE).

601 Table 7. Using the mathematical notation to define the target variable for each model.

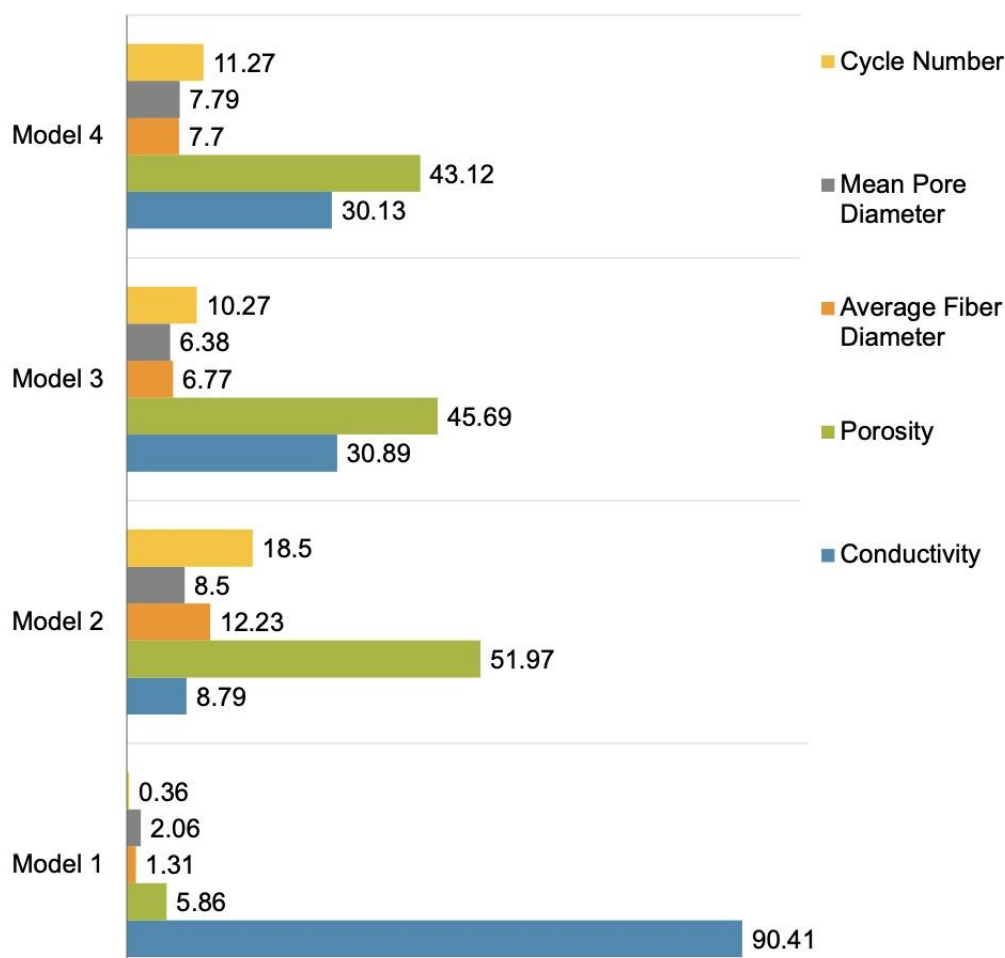
Model	Target Values
Model 1	VE
Model 2	CE
Model 3	EE
Model 4	VE, CE, EE

602





603 The feature importance analysis conducted for all the baseline RFR models reveal that the  
 604 features in Model 3 and Model 4 have approximately the same importance scores. Model 2 follows  
 605 similar trends when compared to Model 3 and Model 4. Model 1, where the target value is VE,  
 606 has a noticeably different distribution of feature importance scores. Model 1 heavily relies on  
 607 conductivity, whereas the other models rely more so on porosity. The comparisons of the four  
 608 models can be seen in Figure 13 and Table 8.



609  
 610 Figure 11. Feature importance scores for single and multiple output random forest regression  
 611 models (Models 1, 2, 3, and 4).  
 612



613 Table 8. Feature importance scores for single and multiple output random forest regression  
614 models.

Feature Importance Scores				
	Model 1	Model 2	Model 3	Model 4
Conductivity	90.41	8.79	30.89	30.13
Porosity	5.86	51.97	45.69	43.12
Average Fiber Diameter	1.31	12.23	6.77	7.7
Mean Pore Diameter	2.06	8.5	6.38	7.79
Cycle Number	0.36	18.5	10.27	11.27

615  
616 The single output models are prone to overfitting, a tell-tale sign of overfitting is if the  
617 testing error is larger than the training error [70-72]. The single output models also did not account  
618 for certain inherent physical limitations that can be accounted for when using a multiple objective  
619 model. The best performing ML models that will be used as surrogate models are a multiple output  
620 gradient boosting regressor and a multi-output RFR.

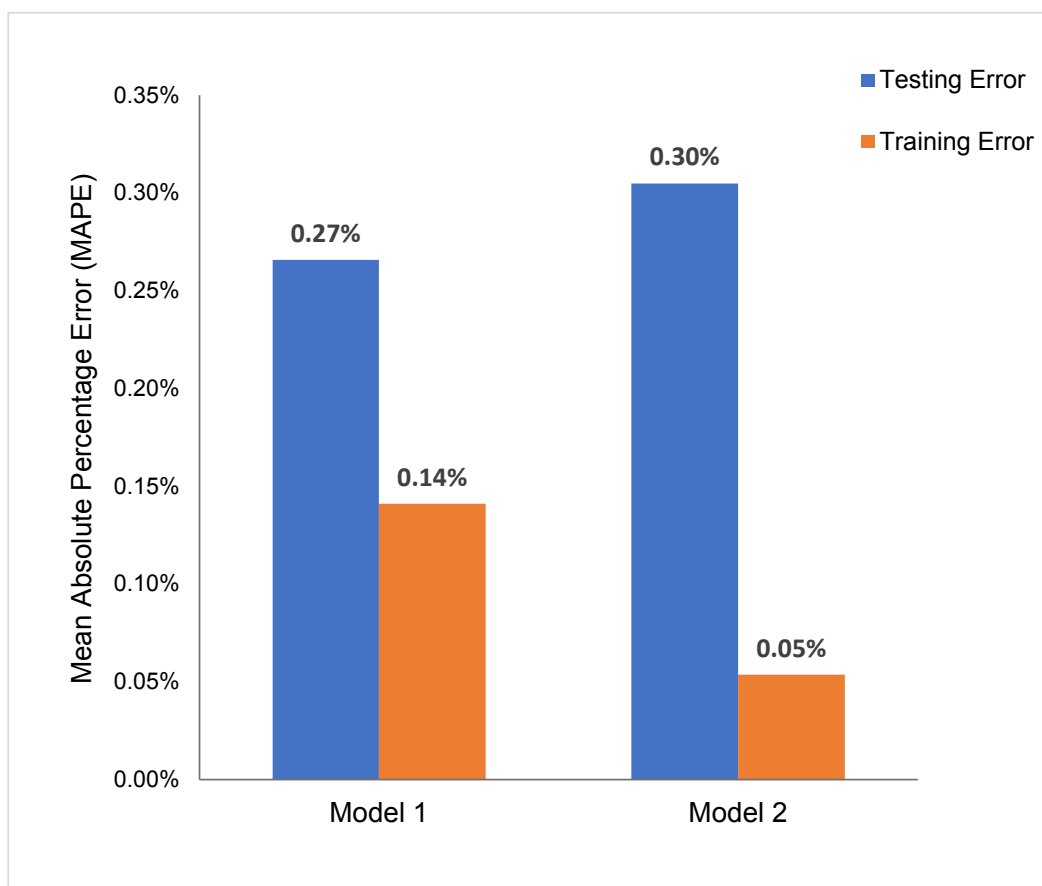
### 621 3.5 ML Based Surrogate Models

622 The best performing ML models are then used to construct the surrogate models. The top  
623 two ML models along with their training and testing error are shown in this section. Two ML  
624 methods to support surrogate modeling were selected as opposed to one method considering that  
625 as the database expands, RFR will perform slower while the GBRs will maintain fast training and  
626 evaluation times. The best performing models will be referred to as *Model 1* and *Model 2*, where  
627 *Model 1* is the multi-output RFR and *Model 2* is the multi-output GBR. RFRs are less complex  
628 than GBRs and therefore more prone to overfitting during the hyperparameter tuning process. The  
629 following hyperparameter tuning methods were performed on *Model 1* and *Model 2* to achieve  
630 maximum model performance: exhaustive grid search over all specified parameters, randomized



631 grid search, and hyperparameter tuning using Bayesian optimization. (These hyperparameter  
632 tuning techniques were performed in the python software's scikit-learn and optuna). The process  
633 of K-fold cross validation was performed with five folds to determine whether the hyperparameters  
634 were causing over or under fitting. Model 1 performed the best with the default scikit-learn  
635 hyperparameters. Model 2 performance increased when implementing hyperparameter tuning  
636 strategy using Bayesian optimization. Figure 14 displays the resulting training and testing error for  
637 the tuned surrogate models. The MAPE scoring metric is used as it is the most interpretable.

638



639

640 Figure 12. Multi-Output RFR; Model 2: Multi-Output GBR - Training and testing scores using  
641 mean absolute percentage error (MAPE) scoring metric.

642



643 The MAPE values in Figure 14 show that the surrogate models prediction errors are less  
644 than 0.15% on the training dataset. The testing error is slightly higher, though still less than 0.3%.  
645 When training error is lower than testing error, it is a sign that that the model is not overfitting.

646 To further emphasize the validity using  $k$ -fold cross validation, the final  
647 hyperparameters for the multi-output random forest regressor are shown in Table 9 where MAPE  
648 remains low for all five folds.

649 Table 9. Hyperparameter Tuning Results for the Multi-Output Random Forest Regressor.

Hyperparameter Description	Hyperparameter Value
mean_fit_time	0.429506
std_fit_time	0.009786
mean_score_time	0.028945
std_score_time	0.001787
param_estimator_max_depth	33
param_estimator_max_features	None
param_estimator_min_samples_leaf	2
param_estimator_min_samples_split	7
split0_test_score	0.480926
split1_test_score	-0.064334
split2_test_score	0.678854
split3_test_score	0.39821
split4_test_score	0.331645
mean_test_score	0.36506
std_test_score	0.24433
rank_test_score	1

650

### 651 3.6 Multi-Objective Optimization with NSGA-II Results

652 A non-dominated genetic sorting algorithm II (NSGA-II) is a variation of the genetic  
653 algorithm that is best suited to find a Pareto set of optimal solutions for multi-objective  
654 optimization problems. Like a traditional genetic algorithm, NSGA-II will begin with an initial  
655 population,  $P$ . The best design combinations in the initial population will move onto the second  
656 generation and this process will repeat until convergence.



657 The main nuance to this approach is that each design combination is evaluated on its fitness  
 658 score and the combinations are also ranked based on their location in the design domain. This will  
 659 eliminate the chance of having repetitive offspring in future generations as well as assuring that  
 660 the entirety of the design space is explored. The final electrode design parameters for surrogate  
 661 Model 1 and 2 using NSGA-II are listed in Table 10. The multiple objective optimization with 5  
 662 inputs  $(x^1, x^2, x^3, x^4, x^5)$  and 3 outputs  $(f_1, f_2, f_3) = (CE, VE, EE)$  using the NSGA-II, the  
 663 optimization problem can be represented as follows: the objective function is represented by  
 664 Equation 17 and the decision variables are  $\sigma^e, \kappa, \varepsilon, d_f, d_p$  shown as  $x$ .

$$\underset{x}{\text{maximize}} f_1(x)$$

$$\underset{x}{\text{maximize}} f_2(x)$$

$$\underset{x}{\text{maximize}} f_3(x) \quad (17)$$

$$\text{where } x = [\sigma^e \quad \kappa \quad \varepsilon \quad d_f \quad d_p]^T$$

665  
 666  
 667  
 668 The objective functions from Equation 17 are then evaluated for each solution  $P$ . The  
 669 solutions are ranked based on non-domination, each solution is assigned to a front, the crowding  
 670 distance for solutions in each front is found. The parents for the next generation are selected abased  
 671 on the non-dominated fronts and crowding distance. Generic operations are applied to create  
 672 offspring solutions.  
 673

674 Table 10. Resulting electrode design parameters for surrogate Model 1 and surrogate Model 2  
 675 using NSGA-II for multi-objective optimization.

	Surrogate Model 1	Surrogate Model 2
Iteration Number	227	212



Electrical Conductivity (S/m)	106.4	107.4
Porosity	0.799	0.900
Permeability (m <sup>2</sup> )	8.1E-10	5.71E-10
Average Fiber Diameter (m)	1.2E-5	1.4E-5
Mean Pore Diameter (m)	1.11E-4	1.85E-4
Predicted Voltage Efficiency	75.75%	75.70%
Predicted Coulombic Efficiency	96.10%	95.72%
Predicted Energy Efficiency	73.12%	72.52%

676

677

678

679

680

681

682

683

684

685

686

687

688

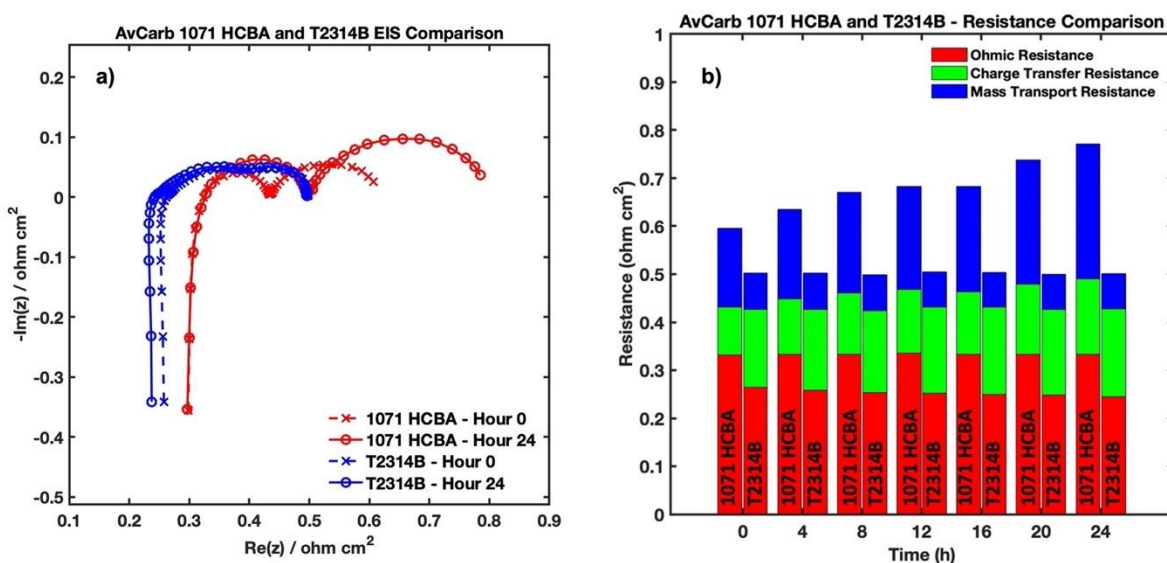
689

The parents of the offspring form a new population. This process continues to repeat until termination criteria is met [73]. The general trend obtained using the ML-based screening and optimization tool suggests that mean pore diameter should be reduced compared to the tested carbon cloth electrodes while maintaining a similar permeability value. Based on this suggestion, a new type of carbon cloth electrode has been fabricated by introducing a carbonaceous binder into woven fabric to make hydrophilic cloths with more complex pore structure and reduced mean pore diameter.

To evaluate the performance of the VRFB with each electrode, ASR values were quantified and compared to visualize the effects of adding a binder to the carbon cloth electrode. Ohmic, charge transfer, and mass transport resistances are determined through curve fitting of the EIS plots, which can be seen in Figure 15a. It is known that the left-most intersection point on the x-axis demonstrates the ohmic resistance for the recorded cycle, the diameter of the first semi-circle of an EIS plot represents charge transfer resistance, and the diameter of the second semi-



690 circle corresponds to mass transport resistance when reading the plot from left to right. Using a Z-  
 691 fit curve fitting analysis within EC-Lab software, the Randles equation ( $R1 + \frac{Q^2}{R2+Wd^2}$ ) is utilized  
 692 which represents the circuit of the physical system. This equation is commonly used to interpret  
 693 impedance data and confirm the values of corresponding resistances obtained from the semi-circle  
 694 intersection points [74]. Figure 15b below displays the comparison of associated resistance values  
 695 throughout the duration of the symmetric cell experiments.



696  
 697 Figure 15. a) EIS data from the beginning and end of each experiment and b) comparison of total  
 698 resistance values of the VRFB with AvCarb 1071 HCBA and AvCarb T2314B electrodes.

699 Figure 15b illustrates the comparative analysis of electrode resistances, showcasing the superior  
 700 performance of the novel binder-coated electrode over the standard 1071 HCBA electrode.  
 701 Symmetric cell cycling coupled with EIS provides a direct correlation of the performance  
 702 enhancement of the electrode. A constant SOC symmetric cell experiment is advantageous for  
 703 multiple reasons, such as the mitigation of cross-over of the active species and the absence of  
 704 chemical or electrical potential gradients which makes the effects of side reactions negligible [44,  
 705 75]. Resistance data from the analysis of EIS experiments can then be used to quantify the



706 performance of the electrode itself without concern for the effects of electrolyte degradation. The  
707 performance enhancement of the VRFB with the new electrodes is evidenced by the reduction in  
708 both ohmic and mass transport resistances by 24% and 66% respectively, attributed to  
709 modifications in the electrode's microstructural parameters induced by the binder coating.  
710 However, it is critical to note the observed increase in charge transfer resistance, which can be  
711 attributed to the suboptimal activation conditions for the newly fabricated electrodes, underscoring  
712 the preliminary nature of these findings. The AvCarb T2314B electrode underwent 24 hours of  
713 thermal activation in a furnace at a temperature of 425°C as an initial activating condition. An in-  
714 depth investigation focused on refining these thermal activation conditions is currently underway,  
715 promising to address this limitation and reduce charge transfer resistance.

716 The aforementioned enhancements in mass transport, ohmic, and total resistance values  
717 signify a marked improvement in carbon cloth electrode performance within VRFB applications.  
718 EIS experiments, performed to compare the base electrode, AvCarb 1071 HCBA, and the electrode  
719 with the addition of a porous binder, AvCarb T2314B, display promising results utilizing the newly  
720 fabricated electrode in terms of reduced total ASRs. These findings corroborate the hypothesis that  
721 integrating a carbonaceous, porous binder layer— as recommended by our optimization analysis—  
722 substantially benefits VRFB performance. Such findings not only highlight the critical role of  
723 electrode composition and structure in optimizing battery performance but also open avenues for  
724 future research to unlock the full potential of VRFB technologies.

#### 725 4. Conclusion

726 In summary, this research makes a substantial contribution to the field by introducing a  
727 cost-effective modeling strategy aimed at optimizing the design of porous carbon cloth electrodes  
728 for VRFB technology. The key innovation lies in the development of a versatile framework that





729 allows for the selection and application of optimal machine learning techniques tailored to the  
730 unique challenges of the design problem. With operating conditions in RFB systems being user-  
731 defined and varying case by case, the behavior of porous carbon electrodes exhibits significant  
732 complexity contingent on specific operational scenarios. Given the impracticality of creating an  
733 exhaustive model for every operating condition, our proposed cost-effective framework offers a  
734 customizable surrogate modeling solution, maintaining high prediction accuracy while ensuring  
735 computational efficiency.

736 Crucially, the adaptability of our framework positions it as a valuable tool for both single-  
737 and multi-objective optimization problems, enabling the discovery of improved electrode design  
738 combinations under the specified operating conditions outlined in the case study. The novel  
739 electrode design not only reduces average ohmic and mass transport resistances but also results in  
740 a reduction to the overall increase of total resistances from 29% to 0.4% during the 24-hour  
741 constant SOC symmetric cycling experiment. It is noteworthy that ongoing experimental results,  
742 set to be disclosed soon, will provide additional empirical insights, further validating the  
743 robustness and applicability of our proposed framework. This study not only represents a  
744 significant step forward but also lays the groundwork for future investigations, offering a platform  
745 for discovering enhanced electrode combinations tailored to specific operating conditions, thereby  
746 eliminating the need for extensive laboratory testing or substantial computational resources. By  
747 addressing the nuanced challenges of electrode design and optimization, this work paves the way  
748 for significant advancements in energy storage solutions, catering to the growing global demand  
749 for renewable energy integration and grid stabilization.

## 750 5. Acknowledgements



751 This material is based upon work supported by the U.S. Department of Energy's Office of  
 752 Energy Efficiency and Renewable Energy (EERE) under the Advanced Manufacturing Office,  
 753 award number DE-EE0009102. Any opinions, findings, and conclusions or recommendations  
 754 expressed in this material are those of the author(s) and do not necessarily reflect the views of the  
 755 Department of Energy. The authors are also indebted to AvCarb Material Solutions for providing  
 756 carbon electrodes and insightful discussions.

## 757 6. Author Contributions

758 **Alina Berkowitz**: data curation, formal analysis, writing original draft. **Ashley A. Caiado**:  
 759 data curation, formal analysis, writing - original draft. **Sundar Rajan Aravamuthan**: formal  
 760 analysis, writing - review & editing. **Aaron Roy**: conceptualization, resources, supervision,  
 761 writing - review & editing. **Ertan Agar**: conceptualization, supervision, funding acquisition,  
 762 writing - review & editing. **Murat Inalpolat**: conceptualization, supervision, funding acquisition,  
 763 writing - review & editing.

## 764 7. Nomenclature

$\varepsilon$	=	porosity
$d_f$	=	average fiber diameter, m <sup>2</sup>
$\kappa$	=	permeability, m <sup>2</sup>
$d_p$	=	mean pore diameter, m
$K_{CK}$	=	Kozeny-Carman coefficient
$\sigma^e$	=	electrical conductivity of porous carbon electrode, S/m
$I$	=	current density, A/m <sup>2</sup>
$\Phi$	=	potential, V
$V^{3+}$	=	$V(III)$
$VO^{2+}$	=	$V(IV)$
$VO_2^+$	=	$V(V)$
kWh	=	Kilowatt hour
anode	=	positive electrode
cathode	=	negative electrode
$R^2$	=	coefficient of determination
$y$	=	data label (response)



$n$	= discrete number of observations
$\mathcal{D}$	= domain (machine learning)
$f$	= expensive “black-box” function
$\hat{f}$	= surrogate model ( <i>emulator</i> or <i>meta-model</i> )
$X$	= data matrix
$m$	= number of samples
$n$	= number of design variables (features)
$x^i$	= $m - \text{dimensional}$ feature vector
$\{x_i, y_i\}$	= data pairs
$\mathcal{D}_{train}$	= training dataset
$\mathcal{D}_{validation}$	= validation dataset
$\mathcal{D}_{test}$	= testing dataset
$\sigma$	= standard deviation
$\sigma^2$	= variance
$\mu$	= mean
$\sigma$	= standard deviation
$t_d$	= charging time, s
$t_c$	= discharging time, s
$V_{ave,d}$	= Average discharging voltage, V
$V_{ave,c}$	= Average charging voltage, V
$K$	= number of folds when using k-fold cross validation
$R_{ohmic}$	= ohmic resistances
$R_{ct}$	= charge transfer resistances
$R_{mt}$	= mass transport resistances
ML	= machine learning
VE	= voltage efficiency
EE	= energy efficiency
CE	= coulombic efficiency
MAPE	= mean absolute percentage error
MAPD	= mean absolute percentage deviation (same as MAPE)
GBR	= gradient boosting regressor
RFR	= random forest regressor
LHS	= Latin hypercube sampling
QMC	= Quasi Monte-Carlo
KDE	= kernel density estimation
OCV	= open circuit voltage



MSE = mean squared error  
MAE = mean absolute error  
RMSE = root mean squared error  
 $r$  = Pearson correlation coefficient  
(between -1 and +1)

765



766 **References**

- 767 1. Kerry, J., *The Long-Term Strategy of the United States, Pathways to Net-Zero*  
768 *Greenhouse Gas Emissions by 2050*. 2021. p. 65.
- 769 2. *What is U.S. electricity generation by energy source?* 2023; Available from:  
770 <https://www.eia.gov/tools/faqs/faq.php?id=427&t=3>.
- 771 3. Chang, L. and Z. Wu, *Performance and Reliability of Electrical Power Grids under*  
772 *Cascading Failures*. International Journal of Electrical Power & Energy Systems, 2011.  
773 **33**: p. 1410-1419.
- 774 4. Dey, P., et al., *Impact of Topology on the Propagation of Cascading Failure in Power*  
775 *Grid*. IEEE Transactions on Smart Grid, 2016. **7**(4): p. 1970-1978.
- 776 5. van Ruijven, B.J., E. De Cian, and I. Sue Wing, *Amplification of future energy demand*  
777 *growth due to climate change*. Nature Communications, 2019. **10**(1): p. 2762.
- 778 6. Jimoh, A.-G. and J.L. Munda, *Challenges of Grid Integration of Wind Power on Power*  
779 *System Grid Integrity: A Review*. International Journal of Renewable Energy Research,  
780 2013. **2**.
- 781 7. Pahari, S.K., et al., *Designing high energy density flow batteries by tuning active-*  
782 *material thermodynamics*. RSC Advances, 2021. **11**(10): p. 5432-5443.
- 783 8. Ziegler, M.S., et al., *Storage Requirements and Costs of Shaping Renewable Energy*  
784 *Toward Grid Decarbonization*. Joule, 2019. **3**(9): p. 2134-2153.
- 785 9. Mann, M., et al., *Energy Storage Grand Challenge Energy Storage Market Report*, D.o.  
786 Energy, Editor. 2020. p. 65.
- 787 10. Wheeler, G.P., L. Wang, and A.C. Marschilok, *Beyond Li-ion Batteries for Grid-Scale*  
788 *Energy Storage*. 2022, Cambridge University Press: Cambridge.
- 789 11. Arteaga, J., H. Zareipour, and V. Thangadurai, *Overview of Lithium-Ion Grid-Scale*  
790 *Energy Storage Systems*. Current Sustainable/Renewable Energy Reports, 2017. **4**: p. 1-  
791 12.
- 792 12. Collath, N., et al., *Aging aware operation of lithium-ion battery energy storage systems:*  
793 *A review*. Journal of Energy Storage, 2022. **55**: p. 105634.
- 794 13. Bragard, M., et al., *The Balance of Renewable Sources and User Demands in Grids:*  
795 *Power Electronics for Modular Battery Energy Storage Systems*. Power Electronics,  
796 IEEE Transactions on, 2011. **25**: p. 3049-3056.
- 797 14. Yang, Z., et al., *Electrochemical Energy Storage for Green Grid*. Chemical Reviews,  
798 2011. **111**(5): p. 3577-3613.
- 799 15. Mitali, J., S. Dhinakaran, and A.A. Mohamad, *Energy storage systems: a review*. Energy  
800 Storage and Saving, 2022. **1**(3): p. 166-216.
- 801 16. Weber, A.Z., et al., *Redox flow batteries: a review*. Journal of Applied Electrochemistry,  
802 2011. **41**(10): p. 1137-1164.
- 803 17. Sánchez-Díez, E., et al., *Redox flow batteries: Status and perspective towards sustainable*  
804 *stationary energy storage*. Journal of Power Sources, 2021. **481**: p. 228804.
- 805 18. Soloveichik, G.L., *Flow Batteries: Current Status and Trends*. Chemical Reviews, 2015.  
806 **115**(20): p. 11533-11558.
- 807 19. Li, Z., et al., *Air-Breathing Aqueous Sulfur Flow Battery for Ultralow-Cost Long-*  
808 *Duration Electrical Storage*. Joule, 2017. **1**(2): p. 306-327.
- 809 20. Nourani, M., et al., *Elucidating Effects of Faradaic Imbalance on Vanadium Redox Flow*  
810 *Battery Performance: Experimental Characterization*. Journal of The Electrochemical  
811 Society, 2019. **166**(15): p. A3844-A3851.



- 812 21. Viswanathan, V., et al., *Cost and performance model for redox flow batteries*. Journal of  
813 Power Sources, 2014. **247**: p. 1040-1051.
- 814 22. Darling, R.M., et al., *Pathways to low-cost electrochemical energy storage: a*  
815 *comparison of aqueous and nonaqueous flow batteries*. Energy & Environmental  
816 Science, 2014. **7**(11): p. 3459-3477.
- 817 23. Houser, J., et al., *Architecture for improved mass transport and system performance in*  
818 *redox flow batteries*. Journal of Power Sources, 2017. **351**: p. 96-105.
- 819 24. *Energy Storage Grand Challenge Roadmap*, D.o. Energy, Editor. 2020.
- 820 25. Gyuk, I., et al., *Grid energy storage*. US Department of Energy, 2013.
- 821 26. Skyllas-Kazacos, M., et al., *Recent advances with UNSW vanadium-based redox flow*  
822 *batteries*. International Journal of Energy Research, 2010. **34**(2): p. 182-189.
- 823 27. Dennison, C.R., et al., *Enhancing Mass Transport in Redox Flow Batteries by Tailoring*  
824 *Flow Field and Electrode Design*. Journal of The Electrochemical Society, 2016. **163**(1):  
825 p. A5163.
- 826 28. Knehr, K., et al., *A Transient Vanadium Flow Battery Model Incorporating Vanadium*  
827 *Crossover and Water Transport through the Membrane*. Journal of the Electrochemical  
828 Society, 2012. **159**: p. A1446-A1459.
- 829 29. Agar, E., et al., *Species transport mechanisms governing capacity loss in vanadium flow*  
830 *batteries: Comparing Nafion® and sulfonated Radel membranes*. Electrochimica Acta,  
831 2013. **98**: p. 66-74.
- 832 30. He, Q., et al., *Modeling of Vanadium Redox Flow Battery and Electrode Optimization*  
833 *with Different Flow Fields*. e-Prime, 2021. **1**: p. 100001.
- 834 31. Kim, K.J., et al., *A technology review of electrodes and reaction mechanisms in*  
835 *vanadium redox flow batteries*. Journal of Materials Chemistry A, 2015. **3**(33): p. 16913-  
836 16933.
- 837 32. Forner-Cuenca, A., et al., *Exploring the Role of Electrode Microstructure on the*  
838 *Performance of Non-Aqueous Redox Flow Batteries*. Journal of The Electrochemical  
839 Society, 2019. **166**(10): p. A2230.
- 840 33. Chu, F., et al., *Novel Interdigitated Flow Field with a Separated Inlet and Outlet for the*  
841 *Vanadium Redox Flow Battery*. Energy & Fuels, 2023. **37**(16): p. 12166-12177.
- 842 34. Lu, M.-Y., et al., *A novel rotary serpentine flow field with improved electrolyte*  
843 *penetration and species distribution for vanadium redox flow battery*. Electrochimica  
844 Acta, 2020. **361**: p. 137089.
- 845 35. Jang, J., et al., *Carbon cloth modified by direct growth of nitrogen-doped carbon*  
846 *nanofibers and its utilization as electrode for zero gap flow batteries*. Chemical  
847 Engineering Journal, 2024. **481**: p. 148644.
- 848 36. Jiang, H.R., et al., *A uniformly distributed bismuth nanoparticle-modified carbon cloth*  
849 *electrode for vanadium redox flow batteries*. Applied Energy, 2019. **240**: p. 226-235.
- 850 37. He, Z., et al., *Modified carbon cloth as positive electrode with high electrochemical*  
851 *performance for vanadium redox flow batteries*. Journal of Energy Chemistry, 2016.  
852 **25**(4): p. 720-725.
- 853 38. Zhang, Z., et al., *A composite electrode with gradient pores for high-performance*  
854 *aqueous redox flow batteries*. Journal of Energy Storage, 2023. **61**: p. 106755.
- 855 39. Zhou, X.L., et al., *A highly permeable and enhanced surface area carbon-cloth electrode*  
856 *for vanadium redox flow batteries*. Journal of Power Sources, 2016. **329**: p. 247-254.





- 857 40. Forner-Cuenca, A., et al., *Exploring the Role of Electrode Microstructure on the*  
858 *Performance of Non-Aqueous Redox Flow Batteries*. Journal of The Electrochemical  
859 Society, 2019. **166**(10): p. A2230-A2241.
- 860 41. Nourani, M., et al., *Exploring the Structure-Function-Performance Relationship of*  
861 *Carbon Electrodes Toward Rational Design of High-Performance Redox Flow Cells*.  
862 ECS Meeting Abstracts, 2021. **MA2021-01**(3): p. 215.
- 863 42. Tenny, K.M., et al., *Comparing Physical and Electrochemical Properties of Different*  
864 *Weave Patterns for Carbon Cloth Electrodes in Redox Flow Batteries*. Journal of  
865 Electrochemical Energy Conversion and Storage, 2020. **17**(4).
- 866 43. Nourani, M., et al., *Impact of Corrosion Conditions on Carbon Paper Electrode*  
867 *Morphology and the Performance of a Vanadium Redox Flow Battery*. Journal of The  
868 Electrochemical Society, 2019. **166**(2): p. A353-A363.
- 869 44. Caiado, A.A., et al., *Exploring the Effectiveness of Carbon Cloth Electrodes for All-*  
870 *Vanadium Redox Flow Batteries*. Journal of The Electrochemical Society, 2023. **170**(11):  
871 p. 110525.
- 872 45. Simon, B.A., et al., *Combining electrochemical and imaging analyses to understand the*  
873 *effect of electrode microstructure and electrolyte properties on redox flow batteries*.  
874 Applied Energy, 2022. **306**: p. 117678.
- 875 46. Ma, X. and F. Xing, *A three-dimensional model for negative half cell of the vanadium*  
876 *redox flow battery*. Electrochimica Acta - ELECTROCHIM ACTA, 2011. **58**.
- 877 47. Ali, E., et al., *A numerical study of electrode thickness and porosity effects in all*  
878 *vanadium redox flow batteries*. Journal of Energy Storage, 2020. **28**: p. 101208.
- 879 48. Yin, C., et al., *Three dimensional multi-physical modeling study of interdigitated flow*  
880 *field in porous electrode for vanadium redox flow battery*. Journal of Power Sources,  
881 2019. **438**: p. 227023.
- 882 49. You, D., H. Zhang, and J. Chen, *A simple model for the vanadium redox battery*.  
883 Electrochimica Acta, 2009. **54**(27): p. 6827-6836.
- 884 50. Zhang, G., et al., *Optimization of porous media flow field for proton exchange membrane*  
885 *fuel cell using a data-driven surrogate model*. Energy Conversion and Management,  
886 2020. **226**: p. 113513.
- 887 51. Kok, M., A. Khalifa, and J. Gostick, *Multiphysics Simulation of the Flow Battery*  
888 *Cathode: Cell Architecture and Electrode Optimization*. Journal of The Electrochemical  
889 Society, 2016. **163**: p. A1408-A1419.
- 890 52. Wan, S., et al., *A coupled machine learning and genetic algorithm approach to the design*  
891 *of porous electrodes for redox flow batteries*. Applied Energy, 2021. **298**: p. 117177.
- 892 53. *Product Page (Fabric)*. [cited 2022 November 10]; Available from:  
893 <https://www.avcarb.com/product-page-fabric/>.
- 894 54. Agar, E., et al., *Reducing capacity fade in vanadium redox flow batteries by altering*  
895 *charging and discharging currents*. Journal of Power Sources, 2014. **246**: p. 767-774.
- 896 55. COMSOL Multiphysics COMSOL AB: Stockholm, Sweeden.
- 897 56. MATLAB. 2022, The MathWorks Inc.: Natick, Massachusetts.
- 898 57. Géron, A., *Hand-On Machine Learning with Scikit-Learn, Keras, and TensorFlow:*  
899 *Concepts, Tools, and Technoques to Build Intelligent Systems*. 2 ed. 2019.
- 900 58. Bishop, C.M., *Pattern Recognition and Machine Learning*. 2006, New York: Springer.
- 901 59. Kamath, C., *Intelligent sampling for surrogate modeling, hyperparameter optimization,*  
902 *and data analysis*. Machine Learning with Applications, 2022. **9**: p. 100373.



- 903 60. Morokoff, W.J. and R.E. Caflisch, *Quasi-Random Sequences and Their Discrepancies*.  
904 SIAM Journal on Scientific Computing, 1994. **15**(6): p. 1251-1279.
- 905 61. Packham, N. and W. Schmidt, *Latin hypercube sampling with dependence and*  
906 *applications in finance*. Journal of Computational Finance, 2010. **13**: p. 81-111.
- 907 62. Mahmoudi, H. and H. Zimmermann, *A new sampling technique for Monte Carlo-based*  
908 *statistical circuit analysis*. 2017. 1277-1280.
- 909 63. Hastie, T.J., R. Tibshirani, and J.H. Friedman. *The Elements of Statistical Learning*.  
910 2001.
- 911 64. Pedregosa, F., et al., *Scikit-learn: Machine Learning in Python*. J. Mach. Learn. Res.,  
912 2011. **12**(null): p. 2825–2830.
- 913 65. Forrester, A., A. Sobester, and A. Keane, *Engineering Design Via Surrogate Modelling:*  
914 *A Practical Guide*. 2008.
- 915 66. Cheng, Z., et al., *Data-driven electrode parameter identification for vanadium redox flow*  
916 *batteries through experimental and numerical methods*. Applied Energy, 2020. **279**: p.  
917 115530.
- 918 67. Settles, B., *Active Learning*. Vol. 6. 2012.
- 919 68. Ren, P., et al., *A Survey of Deep Active Learning*. ACM Comput. Surv., 2021. **54**(9): p.  
920 Article 180.
- 921 69. Settles, B., *From Theories to Queries: Active Learning in Practice*, in *Active Learning*  
922 *and Experimental Design workshop In conjunction with AISTATS 2010*, G. Isabelle, et  
923 al., Editors. 2011, PMLR: Proceedings of Machine Learning Research. p. 1--18.
- 924 70. Regan, T., et al. *Wind Turbine Blade Damage Detection Using Various Machine*  
925 *Learning Algorithms*. in *ASME 2016 International Design Engineering Technical*  
926 *Conferences and Computers and Information in Engineering Conference*. 2016.
- 927 71. Solimine, J. and M. Inalpolat, *An unsupervised data-driven approach for wind turbine*  
928 *blade damage detection under passive acoustics-based excitation*. Wind Engineering,  
929 2022. **46**(4): p. 1311-1330.
- 930 72. Regan, T., C. Beale, and M. Inalpolat, *Wind Turbine Blade Damage Detection Using*  
931 *Supervised Machine Learning Algorithms*. Journal of Vibration and Acoustics, 2017.  
932 **139**(6).
- 933 73. Deb, K., et al. *A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-*  
934 *objective Optimization: NSGA-II*. in *Parallel Problem Solving from Nature PPSN VI*.  
935 2000. Berlin, Heidelberg: Springer Berlin Heidelberg.
- 936 74. Raistrick, I.D., J.R. Macdonald, and D.R. Franceschetti, *Theory*, in *Impedance*  
937 *Spectroscopy*. 2018. p. 21-105.
- 938 75. Potash, R.A., et al., *On the Benefits of a Symmetric Redox Flow Battery*. Journal of The  
939 Electrochemical Society, 2016. **163**(3): p. A338.

941

