Synergizing data-driven and knowledge-based hybrid models for ionic separations

Teslim Olayiwola[1], Luis A. Briceno-Mena[1], Christopher G. Arges[2†], Jose A. Romagnoli[1,*]

[1]*Cain Department of Chemical Engineering, Louisiana State University, Baton Rouge, Louisiana 70803, United States.*
[2]*Department of Chemical Engineering, The Pennsylvania State University, University Park, PA 160802, United States.*
[†]Current address: Argonne National Laboratory, Lemont, IL 60439

[*]Corresponding authors: jose@lsu.edu

Abstract

A hybrid modeling framework has been developed for electrodialysis (ED) and resin-wafer electrodeionization (EDI) in brackish water desalination, integrating compositional modeling with machine learning techniques. Initially, a physics-based compositional model is utilized to characterize the behavior of the unit. Synthetic data is then generated to train a machine learning-based surrogate model capable of handling multiple outputs. This model is further refined using a limited set of experimental data. The effectiveness of this approach is demonstrated by its ability to accurately predict experimental results, indicating a faithful representation of the system's behavior. Through analysis of feature importance facilitated by the machine learning model, a nuanced understanding of the interaction between the chosen ion-exchange resin wafer type and ED/EDI operational parameters is obtained. Notably, it is found that the applied cell voltage has a predominant impact on both separation efficiency and energy consumption. By employing multi-objective optimization techniques, experimental conditions are identified that achieve 99% separation efficiency while keeping energy consumption below 1 kWh/kg.

Keywords: Electrodialysis; Electrodeionization; Hybrid modeling; Brackish water desalination; Optimization

1. Introduction

Electrochemical separation processes are established separation technologies for water desalination and are emerging as a separation platform for many different sectors including nutrient and critical minerals recovery and purification of chemicals generated from processed biomass (e.g., organic acids) [1,2]. The mainstay of electrochemical separation platforms includes electrodialysis (ED), electrodeionization (EDI, and capacitive deionization (CDI) and membrane capacitive deionization (MCDI). These electrochemical units use electric field as the driving force to remove ions from a liquid process stream [3,4]. In emerging applications, electrochemical separations greatest have the ability to perform selective ion separations for metal ion removal from hydrometallurgical streams and recycled black mass [5–8], ammonium and phosphate capture (i.e., nutrient recovery) from waste streams such as agricultural run-off, and organic acids recovery from processed biomass [9,10].

ED, EDI, CDI, and MCDI are mature platforms for water desalination, but are the not the most widespread technologies for desalination because of their higher costs and great energy use when compared to reverse osmosis (RO). However, they do have some advantages over RO because they do not need high pressure piping for desalination. Plus, their modular natural makes them conducive for distributed water treatment and they can be more energy efficient when purifying water with lower concentrations of dissolved salts (for instance, brackish water with less than 5,000 mg/L) [3,11–17]. Generally, electrochemical deionization processes are poised to play an important role in addressing water scarcity for certain scenarios (e.g., brackish water at the lower concentration range) and securing mineral supply chains and managing nutrient cycles.

The electric field is the driving force for removing ions from water in ED and EDI processes. The potential gradient in ED and EDI occurs across a series of alternating ion-exchange membranes (IEMs) partitioning the concentrated and diluate saline streams. The different compartments in the said processes are arranged in a way that allows the cations (i.e., positive ions) in the diluate stream to migrate towards the cation exchange membrane (CEM) so that they permeate across the CEM and collect into the concentrate compartment. Likewise, the anions (i.e., negative ions) selectively pass through the AEM when moving from the diluate compartment to the concentrate compartment. In the concentrate compartment, the anions cannot migrate through the CEM and the cations cannot migrate through the AEM due to Donnan exclusion [1,3]. It is important to point out that ED is similar to EDI except that the diluate stream compartment

contains ion exchange resins (IEX) particles or ion-exchange resin wafers (i.e., immobilized ion-exchange resin particles in a porous wafer configuration) as a way to augment the solution conductivity [18,19]. The IEX particles and resin wafers (RWs) can regenerate the resin particles without the application of acids and bases and provide ions for solution conductivity under very dilute conditions. This occurs via water dissociation at the interface of the oppositely charged particles in the IEX and resin wafer. Water dissociation generates hydroxide ions (OH-) and protons (H+) to regenerate the resins or to augment the process stream ionic conductivity [18–20]. **Fig. 1** shows a simplified pictorial representation of ED and EDI devices.
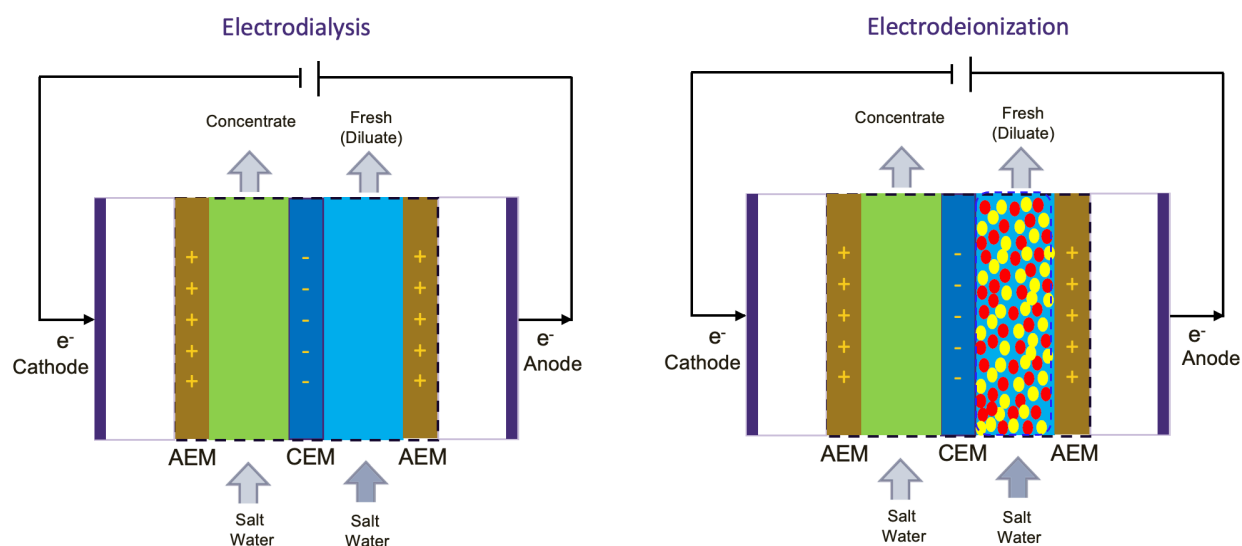
Fig. 1. Simplified process diagrams of electrodialysis (ED) and electrodeionization (EDI). The red and yellow particles depict the cation exchange and anion exchange resins, respectively.

The design of electrochemical deionization units considers many factors: applied cell pair voltage or current density of the stack, concentration of ions in the feed stream, the type of ions in the feed stream, and the properties of the IEMs, IEX particles, and RWs. Furthermore, there is often a tradeoff between the size of the unit, which manifests as a capital expenditure cost, versus energy consumption, an operating expenditure cost. Mathematical models that accurately capture electrochemical deionization system performance with the said considerations is vital to arriving at the highest performing and lowest cost electrochemical separation system.

Regarding ED modeling, several authors [21–23] have proposed mathematical models involving mass balance accounting for ion transport due to convection, migration and diffusion

across IEMs for exchange between the diluate and concentrate compartments. In modeling the ion transport in EDI, most mathematical models describe ion transport along the surface of ion-exchange resins in addition to ion transport via exchange with the resin beads [20,24–26]. The various models are effective for accurately predicting effluent concentrations and voltage/ current profiles that agree with experimental data.

Though mathematical models are a powerful tool for understanding and predicting the behavior of electrochemical systems, they have several limitations that can make them unsuitable for some applications. One limitation of mathematical models is that they are often based on simplifying assumptions that may not be valid in all cases. Also, solving mathematical models is sometimes computationally intensive. As a result of these limitations, there is a need for robust data-driven techniques like machine learning because they do not need to use simplifying assumptions and can be trained on a large set of data that captures the real-world behavior of the system. Machine learning has been applied to a variety of technical fields including various types of chemical processes. More recently, it has been used in the design of polymer membranes for separations [27,28] and fuel cell systems –the latter being an electrochemical process like EDI [29–31]. However, fuel cells and EDI differ from conventional large scale chemical processes that are rich in data because the said electrochemical processes often have smaller amounts of data. Hence, it is difficult to achieve accurate and robust predictions for polymer electrolyte membranes and electrochemical processes using machine learning when there are limited data sets.

Transfer learning (TL) is a useful technique for addressing the challenge of limited data [32–34]. TL involves using a pre-trained task specific data-driven machine learning model (obtained from a large dataset) as a foundation for training a new model (usually a small dataset) [31,34,35]. Given a source domain $\mathcal{D}_s$, and learning task $\mathcal{T}_s$, a target domain $\mathcal{D}_T$ and a learning task $\mathcal{T}_T$, transfer learning aims to help the learning of the target predictive function $f_T(\cdot)$ for the target domain using the knowledge in $\mathcal{D}_s$ and $\mathcal{T}_s$, where $\mathcal{D}_s \neq \mathcal{D}_T$ and $\mathcal{T}_s \neq \mathcal{T}_T$. This technique often enhance a model's accuracy, particularly when the source and target domains are closely related [36,37] and has demonstrated remarkable success in various fields, including materials informatics [35,38,39] and the design of electrochemical systems [31,40].

Herein, we report a hybrid modeling framework that exploits compositional modeling and transfer learning to develop physics-aware models for ED and EDI. First, a mathematical model based on compositional mathematical modeling was used to generate large realization of data to

pretrain a ML i.e., source model. Second, the target ML model with close approximation to experimental data was developed by adjusting the parameters of the pretrained source model with the available experimental data. Thirdly, we identify important factors, such as device operating parameters and material properties, that affect selected performance metrics. Finally, we performed simulations to verify the physical trends obtained from the optimization in relation to experimental intuition. Overall, a hybrid modeling approach was devised from a combination of data-driven and knowledge-based modeling which rendered a powerful tool to accelerate the development of ED and EDI technologies with 99% separation efficiency and <1KWh/kg energy consumed.

2. **Methodology**

In this study, a two-step modeling procedure that exploits compositional modeling and Machine Learning to develop TL-based models for deionization systems as shown in **Fig. 2** are presented. First, a compositional model is obtained using a base model (e.g., mass balances described by coupled differential equations) and the potential drop equation across the compartments. The resulting model is then utilized to generate synthetic training data for a Machine Learning-based surrogate model (*HybridEOS*). Finally, the ML model is refined using Transfer Learning (TL) following the approach previously reported by our team [31]. A significant advantage of using a *HybridEOS* is the ability to perform TL using experimental data to improve the model performance, which can help overcome the limitations of the original compositional model.
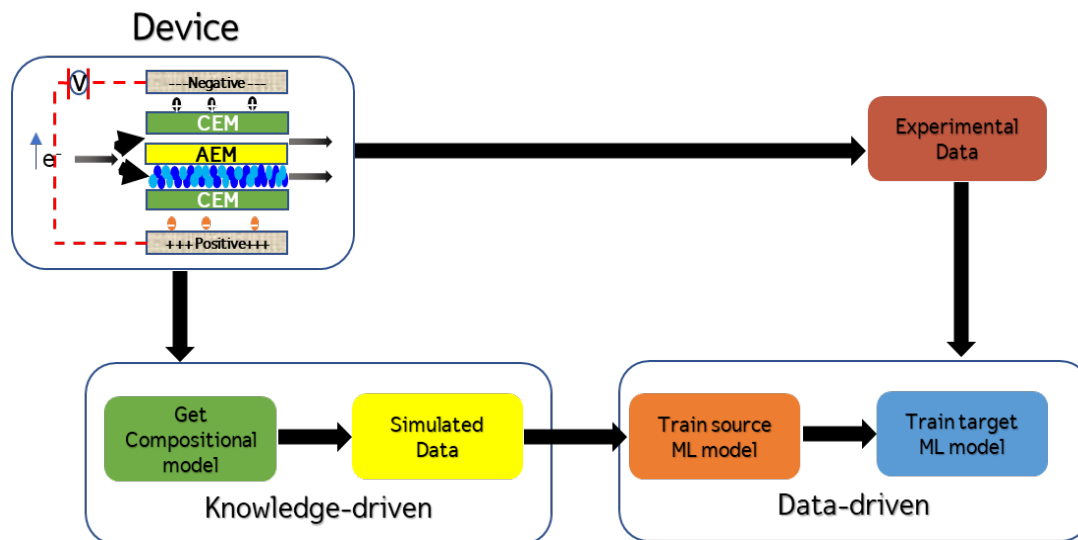
Fig. 2. Two-step modeling approach for the development of physics-aware Machine Learning models.

2.1. Experimental Data

The experimental data used in this study were obtained from the work of Palakkal et al. [4], where they investigated the performance of electrodeionization (EDI) using resin wafers (RWs) formulated with ionomer and non-ionic polymer binders (i.e., polyethylene). The authors produced five types of resin wafers namely Conventional RW with polyethylene binder, RW with cation-exchange ionomer (CEI) binder, RW with anion-exchange ionomer (AEI) binder, Mixed Resin with CEI binder, and Mixed resin with AEI binder. They adopted a single wafer or mixed wafer as the single wafer + oppositely charged ionomer was envisaged to maximize the number of bipolar junction points in the RW to improve the rate of water-splitting because the mixed RWs with and without ionomer binder have a smaller probability that fixed cationic groups meet fixed anionic groups separated by a small gap on the nanoscale. However, they reported EDI data for only three resin wafers namely i) conventional RW with polyethylene binder, ii) RW with cation-exchange ionomer (CEI) binder, and iii) RW with anion-exchange ionomer (AEI) binder. The experimental data and information for the resin wafers include: the chemical structure, average porosity of the resin wafer, and the ion exchange capacity of the RW, and the ionic conductivity of the resin wafer with different concentrations of saline solutions flowing through them. EDI operating parameters and effluent concentration and current density data for brackish water desalination is also available for the three different RWs.[4] It is important to note that the experiments were performed in batch

mode using an initial concentration of 5 g/L NaCl for both the feed and concentrate solutions, a feed flow rate of 25 mL/min, a concentrate flow rate of 50 mL/min, and a cell voltage of approximately 1 V/cell pair. **Fig. 3** shows the chemical structures, their corresponding ionic conductivity in a saline solution and the effluent concentration profile during the EDI process. Further details on the experimental setup and findings can be found in the original reference [4].
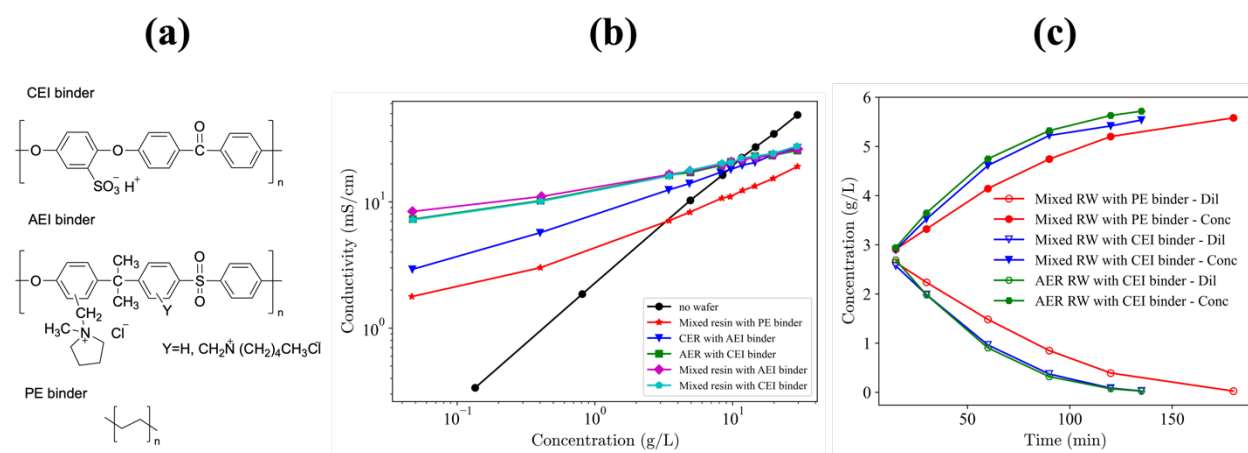
Fig. 3. a) Chemical structures of polyethylene and ionomer binders used in RWs. b) Ionic conductivity (κ) of resin wafers with different binders and ion-exchange resin particles at different sodium chloride concentrations. c) The NaCl concentration profiles of recirculated concentrate and diluate streams over time for EDI demonstrations. These figures were published in[4] and are presented here with permission from Springer Nature.

2.2. Compositional Modeling

We expanded upon the methodologies proposed by Ortiz et al.[21] and Wright et al.[22] for ED to devise a mathematical model tailored for EDI. In our model, we envision the EDI process akin to a reservoir tank, exhibiting characteristics reminiscent of an electric circuit when subjected to an electric field.. The total mass of the salt solution contained in the dilaute and concentrate tank separated by either an AEM or CEM is the sum of bulk flow (i.e., net convective flow of the inlet-outlet stream $J_{bulk,i}$), migration ($J_{mig,i}$) and diffusion ($J_{diff,i}$) of ions i through the AEM and CEM membranes. Herein, accordingly, the mass balance in the tank for an electrodialysis device with two chambers (dilaute, *dil*; concentrate, *con*), shown in **Fig. 4**, is given as follows:

(a)   **(b)**
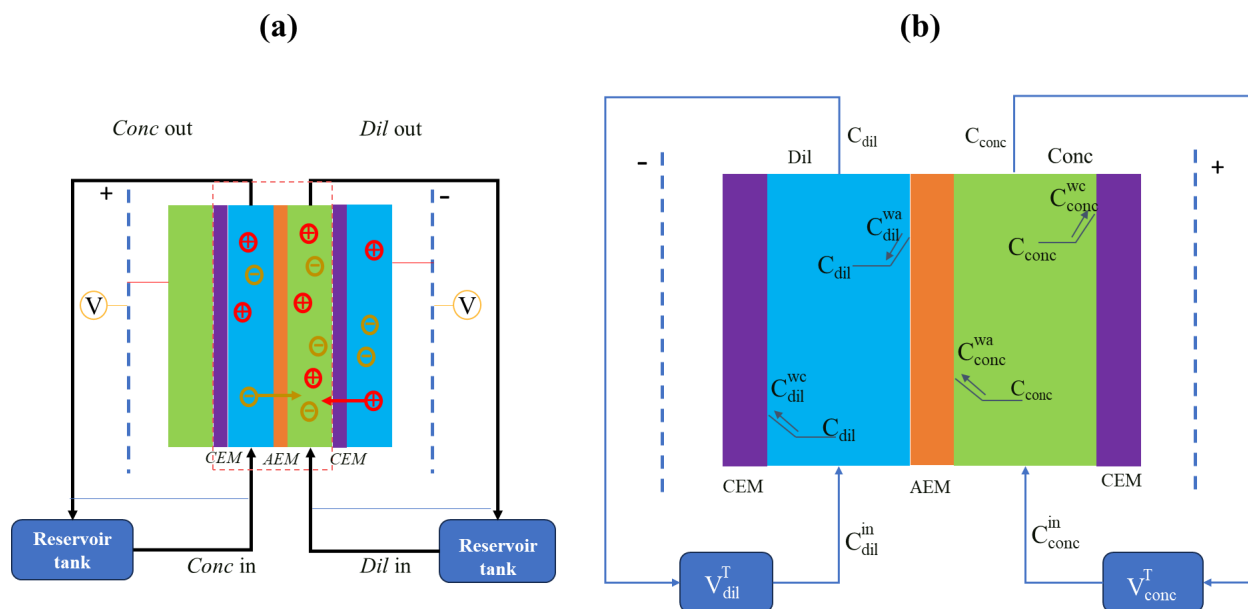
Fig. 4. Schematic of the concentration profiles in bulk and across the boundary layer.

$$V_k \frac{dC_{dil,i}}{dt} = A(J_{bulk,i} + J_{mig,i} + J_{diff,i})$$

$$NV_{k,dil}\frac{dC_{dil}}{dt} = Q_{dil}(C_{dil}^{inlet} - C_{dil}) - \frac{N\phi I}{zF} + \frac{NAD_a}{l_a}(C_{conc}^{wa} - C_{dil}^{wa}) + \frac{NAD_c}{l_c}(C_{con}^{wc} - C_{dil}^{wc}) \quad (1)$$

$$NV_{k,conc}\frac{dC_{con}}{dt} = Q_{con}(C_{con}^{inlet} - C_{con}) + \frac{N\phi I}{zF} - \frac{NAD_a}{l_a}(C_{con}^{wa} - C_{dil}^{wa}) - \frac{NAD_c}{l_c}(C_{con}^{wc} - C_{dil}^{wc}) \quad (2)$$

Where C_{dil}, C_{con}, C_{dil}^{inlet}, C_{con}^{inlet} represents the bulk concentrations (mol/m^3) of the diluate compartment outlet, concentrate compartment outlet, diluate inlet and concentrate inlet streams, respectively; N is the number of cell pairs; ϕ is the current efficiency; z, the charge number; F, Faraday constant (C/mol); D_a, D_c represent the diffusion coefficient of ions in the AEM and CEM, respectively; A, active membrane area (m^2); C_{con}^{wc}, C_{dil}^{wc}, C_{con}^{wa}, C_{dil}^{wa} depicts the concentration at the CEM/concentrate interface, CEM/diluate side, AEM/concentrate side, AEM/ diluate side, respectively; l_a, l_c thickness of the AEM and CEM membrane; Q_{con}, Q_{dil} the volumetric flow rate through the concentrate and diluate stream.

Because EDI differs from ED, the mass balance equations shown in (1) – (2) are modified to account for the addition of ion-exchange RWs. In an EDI cell, ion removal occurs in two phases: (i) cation/anion diffuse to the cation/anion resin surface and get exchanged with counterions of the

resin; and (ii) ions conduct through the solid phase to the membrane interface then across the membrane and then into the concentrate compartment [20,24,41]. In the diluate compartment, the ionic current primarily flows through the resin because the interstitial solution is depleted of mobile ions and has considerably lower ionic conductivity compared to the charged resin. It is reasonable to assume that the resin phase dictates the ion-transport process. Thus, inside the resin compartment, the mass balance [20,25,26] is described as:

$$V_{bed} \frac{dC_{dil}^R}{dt} = A(J_{i,migration} + J_{i,diffusion}) \tag{3}$$

Assuming diffusion-controlled process, eq. 3 becomes:

$$\frac{dC_{dil}^R}{dt} = a_s J_{i,diffusion} = a_s D_i \frac{(C_{dil} - C_{dil}^R)}{\Delta x} = D_i \cdot \frac{3\epsilon\beta\gamma}{r_p^2(1-\epsilon)} \cdot (C_{dil} - C_{dil}^R) \tag{4}$$

In eq. 4, the area specific ratio of IEX a_s and film thickness Δx are given as $\frac{3\epsilon\beta\gamma}{r_p}$ and $r_p(1-\epsilon)$, respectively. For a packed bed where ϵ represents the packing ratio (the proportion of solid volume to the total volume) and r_p denotes the radius of the particle. It's presumed that β signifies the fraction of the IEX surface available for ion exchange and γ is the ratio of CER to AER within the IEX resin bed. Also, it is assumed that the incorporation of the IEX would change the flow dynamics in the diluate compartment, the migration of ions from solution to resin particle and then to IEM. Thus, the mass balance in the diluate chamber is revised as follows:

$$N\epsilon V_{k,dil} \frac{dC_{dil}}{dt} + N(1-\epsilon)V_{k,dil} \frac{dC_{dil}^R}{dt} = Q_{dil}(C_{dil}^{inlet} - C_{dil}) - \frac{N\phi I}{zF} + \frac{NAD_a}{l_a}(C_{conc}^{wa} - C_{dil}^{wa}) +$$

$$\frac{NAD_c}{l_c}(C_{con}^{wc} - C_{dil}^{wc}) \tag{5}$$

The mass balance for an EDI unit with a recirculation tank is given as:

$$\frac{dC_{dil}^{inlet}}{dt} = \frac{Q_{dil}}{V_{dil}^T}(C_{dil} - C_{dil}^{inlet}) \tag{6}$$

$$\frac{dC_{dil}^{inlet}}{dt} = \frac{Q_{dil}}{V_{dil}^T}(C_{dil} - C_{dil}^{inlet}) \tag{7}$$

For the electric voltage applied across the device, we assumed that the device was analogous to a DC circuit. Herein, the total voltage is given in eq. 8 as the sum of the ohmic drops in a diluate

compartment (E_{ohm}^{dil}), ohmic drop in the concentrate compartment (E_{ohm}^{con}), ohmic drop due to membrane overpotential (E_{mem}), ohmic drop due to diffusion overpotential (E_{diff}), ohmic drop in resin (E_{resins}) and ohmic drop in electrode chambers ($E_{anode} - E_{cathode}$).

$$V = (E_{anode} - E_{cathode}) + N\left(E_{ohm}^{dil} + E_{ohm}^{con} + E_{mem} + E_{diff} + E_{resins}\right) \tag{8}$$

$$V_{cell} = V - (E_{anode} - E_{cathode}) = N\left(E_{ohm}^{dil} + E_{ohm}^{con} + E_{mem} + E_{diff} + E_{resins}\right) \tag{9}$$

$$E_{resins} = E_{resins/mem} + j * A * R_{resins} \tag{10}$$

2.3. Model library

In an EDI unit operating at a constant voltage, it becomes imperative to calculate the electric current flowing through the cell since the mass balance expressions are intrinsically linked to this current. After modifying equation (8-10), we obtain the current density j (A/m^2) as presented in eq. 11. The formula simplifies the membrane potential by combining the Donnan potential with the voltage drop stemming from area resistance of IEMs. Also, the ohmic drop in resin layer is given as the sum of total stack resin resistance and overpotential due to migration from resin to membrane surface.

$$j = \frac{V_{cell} - N \cdot \left(E_{mem} + E_{resin} + E_{diff}\right)}{N \cdot A \cdot \left(r_{ohm}^{dil} + r_{ohm}^{con} + r_a + r_c + r_{resin}\right)} \tag{11}$$

In modeling the ohmic resistance due to resin particles (r_{resin}), we embraced the methodology which considers the dependency of resistance on solution concentration and conductivity and material properties such as ion exchange capacity, packing porosity and density. It's worth mentioning that, in contrast to the EDI device, the resins' ohmic resistance in ED is virtually non-existent. **Table 1** details the model fragments utilized in deriving both current density and effluent concentrations.

**Table 1.** Model library considered in this study.

Model fragment	Functional forms

Ohmic resistances E_{ohm}^{i} where i = con, dil	$$E_{ohm}^{i} = j \cdot A \cdot r_{ohm}^{i}$$ $$r_{ohm}^{i} = \frac{L}{A}\frac{1}{C_i \cdot \Lambda_i}$$ $$\Lambda_i = \Lambda_0 - (B_0\Lambda_0 + B_2) \cdot \frac{\sqrt{C_i}}{1 + B_0 u\sqrt{C_i}}$$
Donnan potential, $E_{mem,i}$	$$E_{mem} = \frac{RT}{F}\left[\alpha_c \ln\left(\frac{\gamma_{con}^{wc} C_{con}^{wc}}{\gamma_{dil}^{wc} C_{dil}^{wc}}\right) + \alpha_a \ln\left(\frac{\gamma_{con}^{wa} C_{con}^{wa}}{\gamma_{dil}^{wa} C_{dil}^{wa}}\right)\right]$$
Resin/membrane potential, $E_{resin/mem}$	$$E_{resin/mem} = \left(2t^{+/-} - 1\right)\frac{RT}{F}\ln\left(\frac{\gamma_{dil}^{R} C_{dil}^{R}}{\gamma_{dil}^{wa} C_{dil}^{wa}}\right)$$
Debye-Huckel Activity coefficient, γ_i	$$-\ln(\gamma_i) = \frac{0.5065\sqrt{C * 10^{-3}}}{1 + 1.298\sqrt{C * 10^{-3}}} - 0.$$
Interface concentrations	$$C_{con}^{wa} = C_{con} + \frac{j\phi}{zFk_m}(t_a^- - t^-)$$ $$C_{dil}^{wa} = C_{dil} - \frac{j\phi}{zFk_m}(t_a^- - t^-)$$ $$C_{con}^{wc} = C_{con} + \frac{j\phi}{zFk_m}(t_c^+ - t^+)$$ $$C_{dil}^{wc} = C_{dil} - \frac{j\phi}{zFk_m}(t_c^+ - t^+)$$ $$t_i^{a/c} = 0.5(\alpha_i + 1)$$

As outlined in **Table 1**, symbols $t_{a/c}^{+/-}$ and $t^{+/-}$ describe the transport number of counterions in IEM and solutions, respectively. Also, $\alpha_{c/a}$, k_m and Λ designates the IEMs perm-selectivity, mass transfer coefficient of ions at the membrane-solution interface, and the molar salt solution conductivity, respectively.

The eq. 1 - 11 previously determine the cell current density (j) versus time under a constant applied voltage as well as the ion concentration profile over time in the diluate and concentrate compartments. The calculations were made in a Python-based program and the simulation results were compared with experimental data to assess the model's accuracy.

2.4. Machine Learning Model
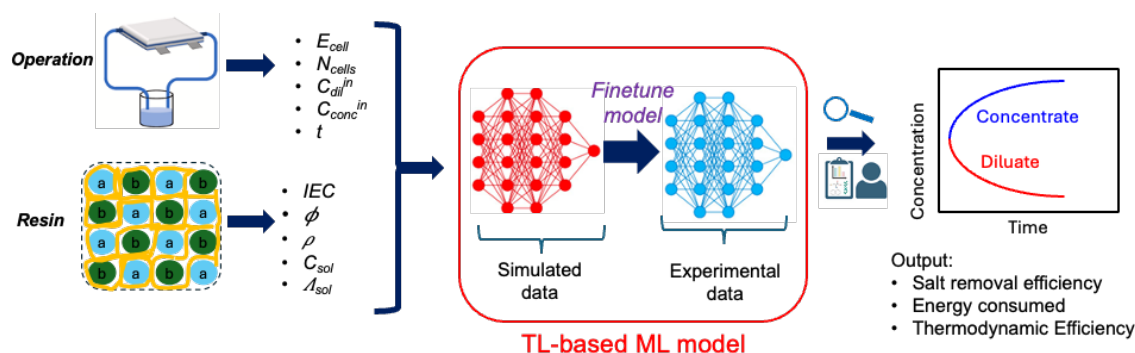
Fig. 5. Featurization and Transfer learning-based strategy implemented in this work. Operational data from the ED/EDI include the cell potential (E_{cell}), number of cell pairs (N_{cells}), inlet concentration in feed/diluate stream (C_{dil}^{in}), inlet concentration in concentrate stream (C_{conc}^{in}), total operation time (t) and material information extracted from the resin include ion exchange capacity (*IEC*), porosity (ϕ), packing density (ρ), solution concentration (C_{sol}) and conductivity (Λ_{sol}).

A transfer learning (TL)-based approach as shown in **Fig. 5** was adopted, and it uses a knowledge-based explicit equation model to generate data based on several combinations of the input variables (using a full 3-level factorial experimental design) and pretrain a source model which is a fully connected neural network (FCN). In creating the experimental design, we considered the following levels: 5 resin wafers (conventional RW with polyethylene binder, RW with CEI binder, RW with AEI binder, mixed resin with CEI binder, and mixed resin with AEI binder) as adopted in the available experimental result [4], feed/diluate stream inlet concentration (1, 2.5, 3.5, 5g/L NaCl), concentrate stream inlet concentration (1, 2.5, 3.5, 5g/L NaCl), number of cells (1, 3, 7, 10), total operation time (30, 60, 90, 120 minutes), and cell potential (0.1, 0.4, 1.0, 2.0). For each RW, their corresponding ion exchange capacity, porosity, and density were extracted. After creating the design space, the corresponding salt removal efficiency (eq. 14) [4], energy consumed per mass (eq. 13) [4], and maximum theoretical thermodynamic efficiency (eq. 14) [42] for each combination were computed.

Salt removal efficiency, $SR = 1 - \frac{C_D}{C_0}$ $\qquad\qquad$ (12)

Energy consumed, EC (KWh/kg): $EC = 2.78 \cdot 10^{-7} \cdot \frac{V \int I \cdot dt}{m}$ $\qquad\qquad$ (13)

Maximum theoretical thermodynamic efficiency, TEE_{max}

$$TEE_{max} = \frac{\overline{\Delta V_R}(WR,SR)}{\frac{|Z^+|+|Z^+|}{|Z^+ \cdot Z^-|} \cdot \frac{RT}{F} \cdot ln \frac{1-WR \cdot (1-SR)}{(1-WR) \cdot (1-SR)}} \quad (14)$$

Where C_B is the brine concentration (i.e. final concentration in the concentrate stream), C_0 is the feed concentration (i.e. inlet concentration into diluate stream), C_D is the diluate stream concentration, m is the mass of ion removed, $\overline{\Delta V_R}(WR,SR)$ is the minimum mean voltage for a reversible separation, v is the total charge $v = v^+ - v^-$, e is the charge on the electron (1.602 * 10^{-19} C), R is the universal gas constant (8.314 J/mol-K), k_B is the Boltzmann constant (1.38 * 10^{-23} J/K) and F is the Faraday constant (96,485 C/mol).

$$m = Q_{dil} \cdot \int [C_0 - C_D(t)] \cdot dt \quad (15)$$

$$WR = \frac{C_B - C_0}{C_B - C_D} \quad (16)$$

$$\overline{\Delta V_R}(WR,SR) = \frac{v}{|v^- \cdot Z^-|} \cdot \frac{k_B T}{e} \cdot \frac{1}{SR} \left\{ \frac{1}{WR} ln \left[\frac{1-WR \cdot (1-SR)}{1-WR} \right] - (1-SR) \cdot ln \left[\frac{1-WR \cdot (1-SR)}{(1-SR) \cdot (1-WR)} \right] \right\} \quad (17)$$

In creating the ML model, the input variables in synthetic and experimental data were scaled to using a min–max approach. To obtain the best model architecture, the scaled synthetic data was fed to the fully connected neural network (FCN) model with parameters (namely number of layers, neurons per layer, and learning rate) tuned with the Bayesian algorithm, facilitated by the KerasTuner [43] optimizer running on TensorFlow [44]—was employed to optimize parameters using mean squared error as the objective function. Herein, the hyperparameters considered include number of layers (1 – 5), neurons per layer (1 – 50), and learning rate (10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}). Using the optimized ML architecture, a source model was trained with the synthetic data and then fine-tuned (i.e. target model) using the 30 available experimental datasets to improve its performance and correct any anomaly arising from the limitations of the explicit equation.

2.5. Insights from ML model

Extracting insights from developed ML models can be difficult, especially for complex models like neural networks. To deal with this, we adopted the method '*Shapley additive explanation, SHAP*' proposed by Lundberg and Lee [45] to express the importance of variables.

SHAP has been used in wide ranging research namely electrochemical design [30,31], wastewater treatment plants [46], gas separation [47], and bioprocesses [48]. The approach in *SHAP* rests on the idea of obtaining an explanation model *g* for each individual prediction, that is a linear combination of the binary variables z' $\in \{0, 1\}^M$ where *M* is the number of input features.

$$g(z') = \Phi_0 + \sum_{i=1}^{M} \Phi_i z_i' \qquad (18)$$

As shown in eq. 18, the final prediction will then be the additive result of the individual influences $\Phi_i \in IR$. To calculate the SHAP values for each input, an explanation model with all the inputs is evaluated and compared to the evaluations of all possible explanation models that do not include the variable of interest. The final value for a given input is then a weighted average of all the differences. The SHAP values serve as quantifiers of the magnitude of influence exerted by each feature on the model's prediction. A greater positive or negative value in these SHAP scores signifies a correspondingly more substantial impact of the respective feature.

2.6. Optimization

Predictive models help understand the current limitations of electrochemical devices while also identifying pathways for improving performance. Here, we explore the input space of the models to find combinations that lead to maximize the performance metrics. Such a problem can be formulated as a constrained multi-objective optimization task as defined in eq. 19. Here *SR*, *EC*, *TEE* represent the salt removal efficiency, energy consumed per mass, and maximum theoretical thermodynamic efficiency, respectively estimated using surrogate model, *dist* is the kNN distance between a candidate solution and the samples in the training space, and l_i and u_i are the lower and upper bounds of each decision variable (inputs in the ML model) set by their physical realizability. Also, experimentally, SR and TEE range between 0 and 1 while EC is non-negative. By maximizing the SR and TEE_{max} while minimizing the EC and kNN distance, the optimization is guided away from point of extrapolation and towards more trustworthy solutions. The optimization problem defined in eq. (16) was solved using NSGA-II with 100 population size and 100 generations.

$$min/max \qquad f_m(x) = [SR, EC, TEE, dist]$$

$$s.t. \begin{cases} x \in \mathbb{R}^n \\ l_i < x_i < u_i \in \mathbb{R}, \, i = 1, \dots, n \\ 0.1 < SR < 0.99 \end{cases}$$

$$0.2 < TEE < 0.68 \tag{19}$$
$$EC \geq 0$$

3. Results and discussion

3.1. Model Library

In the algebraic expression for the EDI unit defined in eq. (9 - 11), the applied voltage depends on the current density and resistance contributions from the material (membrane and resin) and solution (diluate and concentrate) streams. The diluate stream contains the ion-exchange resin wafer – which consists of cation and/or anion exchange resins immobilized with polymer binder (e.g., polyethylene, cation exchange ionomer, and anion exchange ionomer). For the resin particles, there exists no expression that maps the properties of the resins to the resistance (i.e. inverse of conductivity). To model the conductivity of the resin wafers, a machine learning model utilized ionic exchange capacity (IEC_{resin}), porosity (ϕ), solution concentration (C_{sol}) and solution conductivity (Λ_{sol}) as inputs while resin wafer ionic conductivity was an output (**Fig S1**). The solution conductivity was obtained based on the mathematical expression provided in **Table 1**. The experimental data used in training the model was obtained from the works of Palakal et al[4] and summarized in **Table 1**. Initially, the input matrix was transformed into a standard distribution and afterwards, the testing set containing 20% of the 50 data points was randomly separated from the data set. The remaining 80% of the data set (i.e., the training set) was used for ML model development and optimization. First, the regression models for resin conductivity were generated using the automated ML pipeline provided by AutoML [49]. The pipeline automatically selects XGBoost as the best algorithm outperformed other algorithms including elastic net, extra trees, gradient boosting, AdaBoost, decision tree, K-neighbors, lasso lars, linear SVR, random forest, ridge regression and SGD regression, known for its strong fitting and generalization capabilities. Afterwards, the XGBoost based model was further studied and the results presented here. Using this input matrix, an XGBoost (gradient boosting network) with 4 inputs and 1 output was trained for 1000 epochs with early stopping criterion and the mean squared error as the loss function. After training the model for 1000 epochs as shown in **Fig. 6a**, the model was evaluated, and **Fig. 6b** shows the performance between the models' predictions and experimental training and testing data. The determined parameters for the XGBoost model are as follows: learning_rate = 0.1,

max_depth = 5, n_estimators = 150, while the remaining parameters are set to their default values. Based on the regression analysis, the ML model gives an accurate prediction of the excellent performance having a train accuracy (R^2) of 0.99 and RMSE of 0.35 mS/cm. Furthermore, the model was applied to the remaining 20% test data and resulted in R^2 and RMSE of 0.92 and 1.02 mS/cm, respectively. The accuracy of the ML model is comparable to the experimental data and demonstrated the suitability of the developed ML as a tool to compute the resistance of the diluate stream in an EDI device.

Furthermore, we utilized Shapley values to quantify the influence of the four inputs (IEC_{resin}, ϕ, C_{sol}, Λ_{sol}) on the model output (i.e. resin conductivity, Λ_{resin}), aiming to gain a comprehensive understanding of their respective contributions to this property. Each graph's y-axis represents the feature list, arranged in descending order according to their impact on the overall prediction, with the most influential feature occupying the top position, followed by the less influential ones. Notably, our analysis revealed that the pivotal determinants of resin conductivity are the solution concentration (C_{sol}), solution conductivity (Λ_{sol}), and packing porosity (ϕ), emerging as the most dominant factors impacting this property. This finding concurs with the established knowledge in the realm of IEMs. Specifically, the concentration/conductivity of ions within the solution and the available interstitial spaces for ion transport stand out as paramount determinants of electric field conduction. Additionally, it's worth noting that while the IEC of the resin (IEC_{resin}) does influence Λ_{resin}, its effect is overshadowed by the aforementioned factors. Additionally, it's worth noting that while the IEC of the resin does influence resin conductivity, its effect is overshadowed by the other parameters namely C_{sol}, Λ_{sol}, and ϕ. In an electrochemical system, the concentration of ions within the resin, which is intricately linked to the IEC, ultimately depends on bulk concentration levels. This observation from the trained regression model help rationalized the contribution of solution and material properties on the resulting resin conductivity.
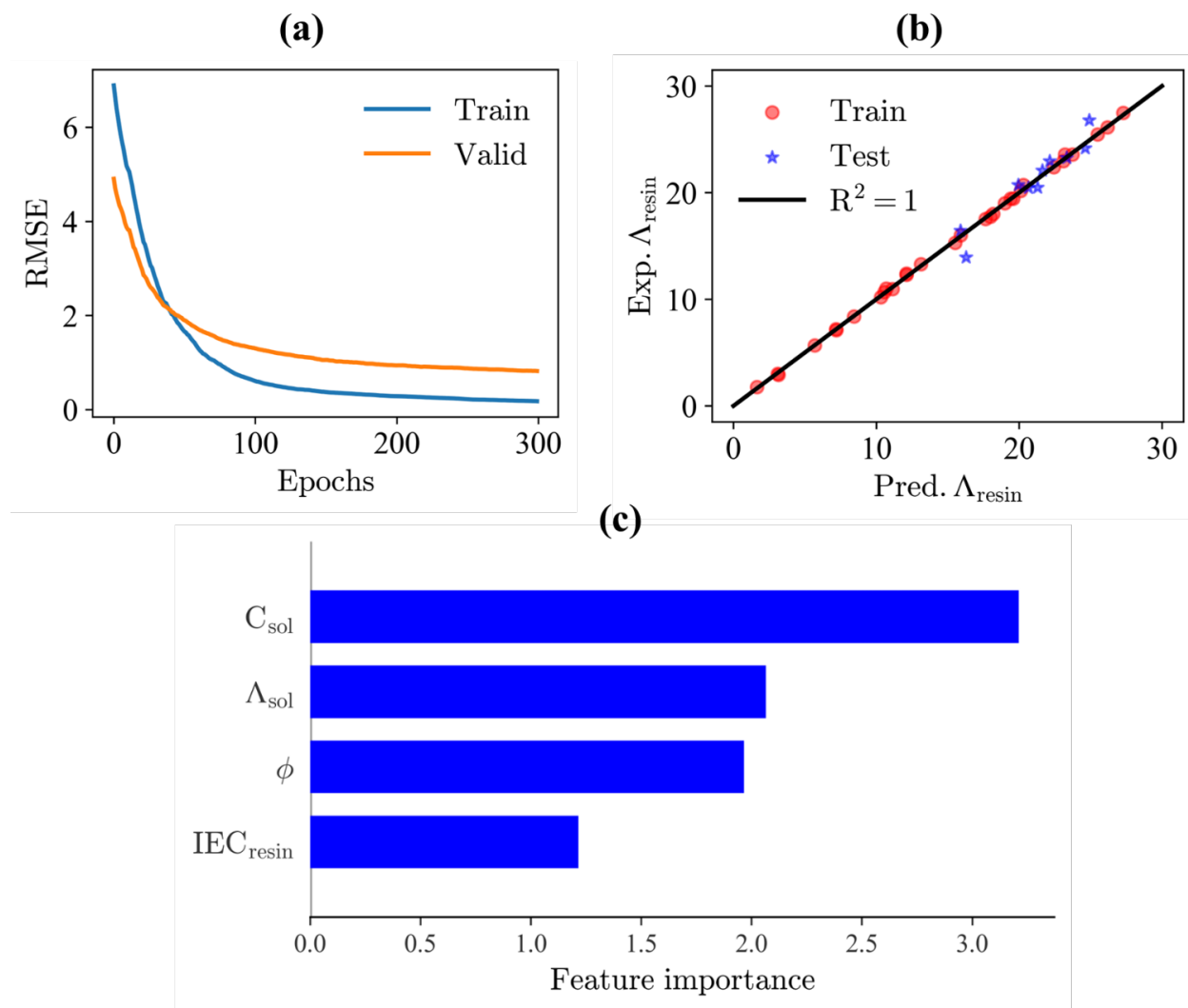
Fig. 6. Performance and feature importance of the ML model to estimate the ionic conductivity of the ion-exchange resin wafers (RWs) (a) error curve showing the RMSE over training epochs (b) trends of experimental resin conductivity versus predicted resin conductivity (c) impact of features on the resin conductivity using SHAP algorithm.

3.2. Compositional Model

To develop a numerical model capable of accurately representing the ion transport behavior in EDI processes, we integrated the resin model discussed in Section 3.1 with the governing equations detailed in Sections 2.2 to 2.3. We obtained the outlet concentration profiles for both the diluate and concentrate streams, as well as the current density profile, by solving the coupled ODEs using a Python ODE solver [50]. For reference, we considered an experimental EDI study [4] with an inlet concentration of 5 g/L NaCl (i.e., brackish water) for both the feed and concentrate

solutions, a feed/diluate flow rate of 25 mL/min, a concentrate flow rate of 50 mL/min, a cell voltage of 1 V/cell pair, and a configuration with 4 cell pairs. The membrane utilized is the Neosepta CMX and AMX while the resin is combination of Purolite PFC100E (cation-exchange resins, CER), Purolite PFA400 (anion-exchange resins, AER) and various binders, leading to the properties listed in **Table S1**.

Fig. 8 and **Fig. S2** depict the concentration profiles of bulk concentrations in the diluate and concentrate channels. The model predictions for EDI concentration profiles for the EDI unit using resin wafers with anion and cation exchange resin particles (i.e., mixed resin) bound with a cation exchange ionomer (i.e., sulfonated poly(arylene ether ether ketone) (SPEEK)) binder and anion exchange resin (ARE) with SPEEK binder showed close agreement with experimental values. Additionally, we conducted a statistical analysis of the fitting procedure, presenting the results in **Table 2**. We assessed the prediction accuracy of the compositional model using mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2). For the mixed resin with SPEEK (CEI) binder, the MAE values for model predictions were found to be 0.22, 0.19, and 0.02 for concentration in concentrate streams, concentration in diluate streams, and current density, respectively. Correspondingly, the RMSE values were 0.26, 0.24, and 0.02. These levels of prediction accuracy closely align with experimental data, illustrating that the numerical model does not falsely inflate accuracy due to data leakage. **Table 2** demonstrates that the MAEs of the numerical model predictions are as low as ~0.01 A and ~0.16 g/L, while the RMSEs are ~0.02 A and ~0.20 g/L for current and concentrations, respectively. These results underscore the capability of the proposed numerical model to accurately represent ion transport behavior in electrochemical setups like electrodialysis and electrodeionization.

(a)

(b)
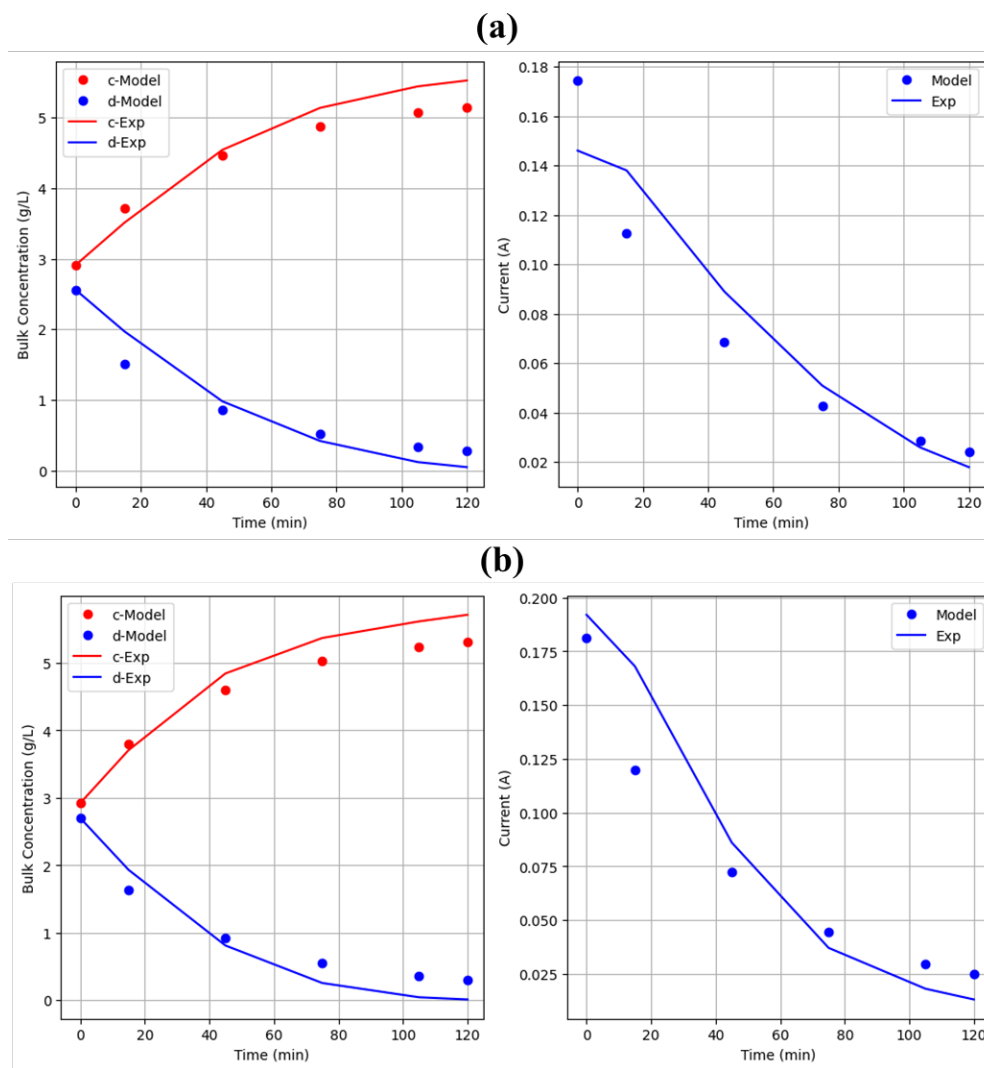
Fig. 8. Training results for the continuum EDI model. **(left)** predicted (label 'Model') and experimental ('Exp') effluent concentration on the concentrate (label 'c') and diluate (label 'd') compartment. **(right)** predicted (Model) and experimental (Exp) current from Pallakal et al.[4] **(a)** Mixed Resin with SPEEK(CEI)-binder (b) AER with SPEEK(CEI) binder.

Table 2. Performance of the Compositional Model for the available experimental data from ref.[4]

| Data	Concentrate (c)	Diluate (d)	Current
Mixed Resin-SPEEK (CEI)-binder	RMSE = 0.26 R^2 = 0.93 MAE = 0.22	RMSE = 0.24 R^2 = 0.94 MAE = 0.19	RMSE = 0.02 R^2 = 0.87 MAE = 0.02
AER with SPEEK (CEI) binder	RMSE = 0.29 R^2= 0.92 MAE = 0.24	RMSE = 0.25 R^2= 0.94 MAE = 0.22	RMSE = 0.02 R^2 = 0.90 MAE = 0.02
Mixed Resin with PE binder	RMSE = 0.20 R^2 = 0.96 MAE = 0.16	RMSE = 0.28 R^2= 0.92 MAE = 0.22	RMSE = 0.02 R^2 = 0.70 MAE = 0.01

After validating the model for brackish water deionization, we conducted sensitivity analyses for parameters including the number of cell pairs, applied cell voltage (V/cell), feed + concentrate concentration, and active membrane area. **Fig. 9** illustrates the effects of varying these parameters on the concentration profiles in the diluate and concentrate channels, as well as the resulting current. Additionally, **Fig. S3** presents changes in flowrates, volume of the recirculation stream, and concentration of the concentrate and diluate. Our findings indicate that increasing the number of cell pairs leads to lower current flow and greater ion removal. Likewise, increasing the applied voltage results in higher current flow and greater ion removal within the deionization chamber.
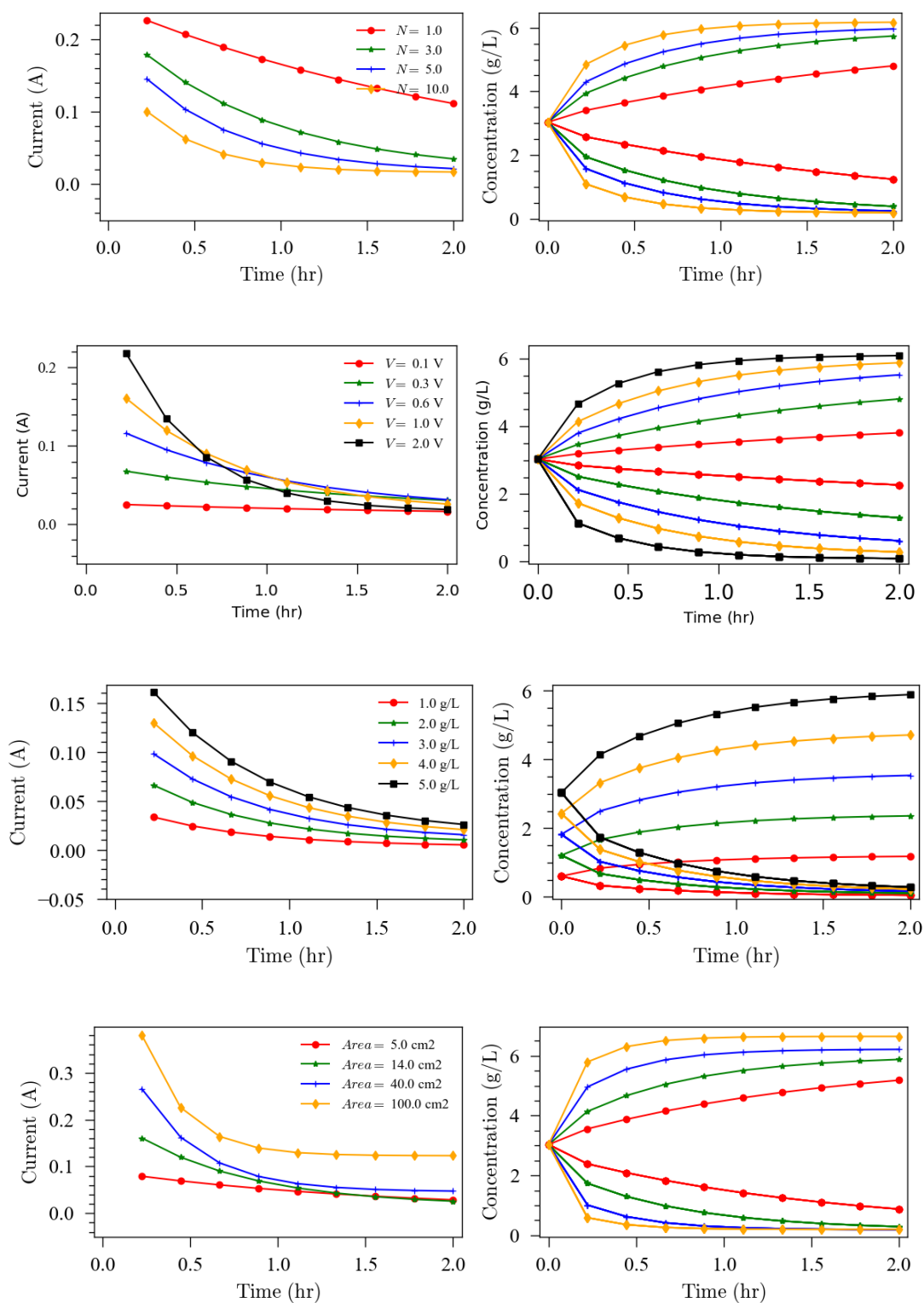
Fig. 9. Effect of number of cells (N), applied cell voltage (V), feed concentration, and active membrane area (area) on the device performance. **(left)** predicted current **(right)** predicted effluent concentration on the concentrate and dilute sides.

3.3. ML-based Surrogate Model

Once the compositional model is obtained, synthetic data is generated by simulating system responses to changes in the input variables. In generating the data, we consider the following materials for the EDI process: 5 RWs, 1 type of an AEM, and 1 type of a CEMs. These materials are described in **Table S1** and **S2,** respectively. These three components of the EDI process resulted in 9 variables which includes operational variables such as cell voltage, feed/diluate concentration, concentrate concentration, number of cells and total operational time, and resin properties such as ion exchange capacity, porosity, and diameter. To attain a well-balanced dataset, we meticulously structured the values for the input variables by employing a comprehensive 3-level factorial experimental design encompassing 9 variables. This rigorous approach is chosen with the intention of achieving both a uniform distribution of data and an ample number of instances for training, thereby mitigating any potential limitations stemming from a shortage of pretraining data in subsequent stages of the analysis. Using the generated data, an ML-based surrogate model "*HybridEOS*" can be trained to capture the information contained in the compositional model. The architecture and training scheme for a fully connected ANN was optimized to minimize the prediction using the number of layers, number of nodes, learning rate and batch size as decision variables and training, testing and validation error as objective functions. The *HybridEOS* was trained to predict three outputs, namely the salt removal efficiency (SR), energy consumed per mass (EC), and maximum theoretical thermodynamic efficiency (TEE_{max}). The training and validation sets consists of 5683 data points and the test set consists of 299 data points that were kept from the model during training and used for model evaluation only. Afterwards, these parameters were added to the data to create a dataset with dimensionality in the order of the experimental data used in the transfer learning phase. Afterwards, the model was trained and the selected architecture after optimization with Bayesian Optimizer was an ANN with 4 layers with architecture 9:50:42:29:50 and a learning rate of 0.001. **Fig. 9(a)** and **Fig. S4** present the training results of the ML model, showing a remarkable overlap between the *HybridEOS* predictions and the actual values. These outcomes affirm the successful training of the *HybridEOS* models, allowing for accurate estimation of SR, EC, and TEE_{max} in other material outlet streams. In **Fig. 9(b)** and **Table 2**, the performance metrics (RMSE, MAE, R^2) of the ML model estimates are depicted, demonstrating their strong performance in terms of estimation accuracy.
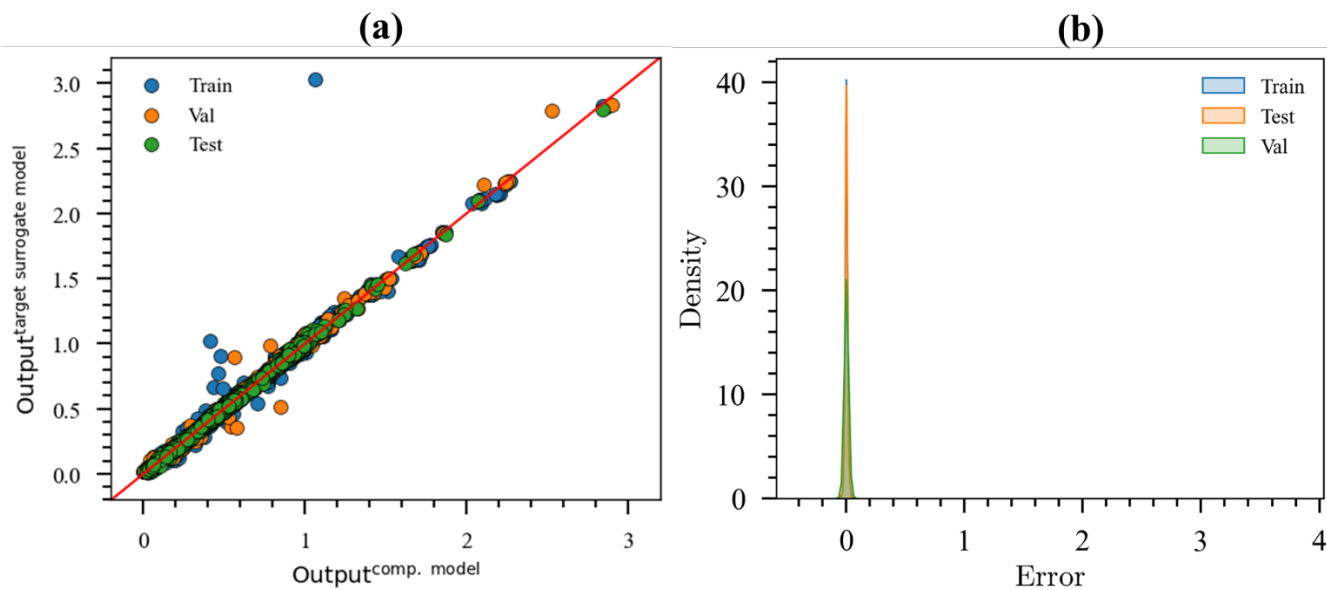
Fig. 9. Training results for the source model using simulated data from a knowledge-based model and optimized ANN architectures. The nested plot displays the convergence of the MSE error curve.

Once the source model was built and trained, TL was used to refine the model using the available experimental data consisting of 30 samples were obtained of which 27 were used during training/ validation and the rest as test set. A proper TL scheme was obtained by conducting 3 steps namely (i) freezing the layers in the source model *(ii)* retrain the new model with learning rate of 0.001 for 2000 epochs *(iii)* unfreeze the source model layers in *(ii)* at low learning rate (0.00001) and train for a short time (2 epochs). **Fig. 10** shows the performance of the TL scheme. The cross-validation shows a small improvement in the validation error after the TL step and an increase in the variance of the training and testing errors. Our TL model was trained on the same descriptors as the source model detailed above, including the hyperparameter procedure. Using these features, we were able to develop an TL model that reproduces experimental values with high accuracy including the validation data. This result supports the consensus on the advantages associated with the TL scheme.
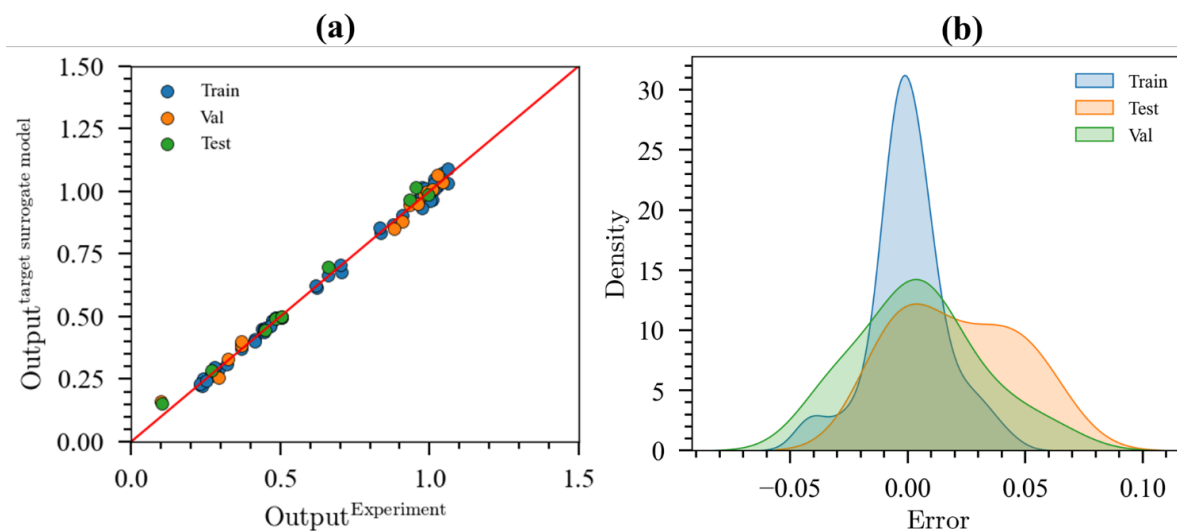
Fig. 10. Training results for the target model using experimental data. The nested plot displays the convergence of the MSE error curve.

Table 3. Performance metrics of the source and target model.

Dataset	Source Model			Target Model		
	RMSE	MAE	R^2	RMSE	MAE	R^2
Train	0.02	0.01	0.99	0.02	0.01	1.00
Val	0.08	0.01	0.94	0.03	0.02	0.99
Test	0.01	0.01	1.00	0.03	0.03	0.99

3.4. Model Interpretation

To understand the learning mechanism of the ML model for the selected EDI performance metrics (salt removal efficiency, SR; energy consumed EC; maximum theoretical thermodynamic efficiency, TEE_{max}), the TL-based NN model was employed for predictive analysis. The investigation delved into the intricate relationship between input parameters and model outputs. Subsequently, SHapley Additive exPlanations (SHAP) values were computed from a carefully selected dataset within the simulated parameter space. The derived importance scores were visualized through SHAP summary plots (**Fig. 11**, **S5** and **S6**), shedding light on the influence of individual inputs on model predictions. In **Fig. 11** (left), the importance of each input element is ranked from most to least while in **Fig. 11** (right), the diagram displays the value of each input element through a range of colors, while the x-axis SHAP value denotes the effect of each factor.
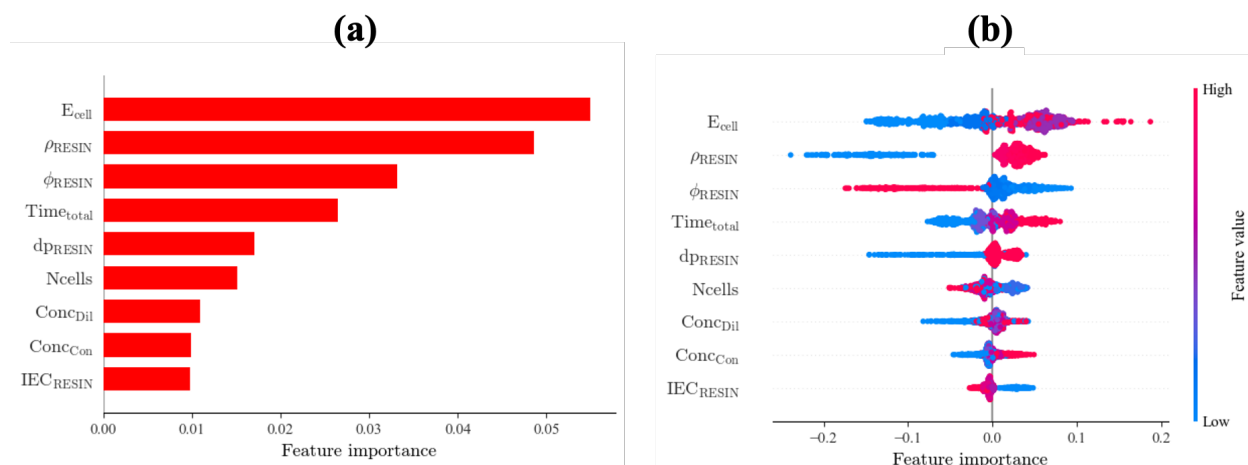
Fig. 11. Investigation of the acquired knowledge using the surrogate model, comprising (a) a SHAP summary plot and (b) the SHAP importance regarding the model's predictive capacity for salt removal efficiency (SR). The color scale signifies variable values, with red indicating high values and blue representing low values.

As observed in **Fig. 11**, the applied cell voltage (E_{cell}) was shown to have a significant factor with both favorable and detrimental impacts on the removal efficiency of ED/EDI device. The second crucial factor was found to be operational time ($Time_{total}$) with both positive and negative effects. Resin properties rank as the third, fourth and fifth most important factors. For example, the influence of resin particle diameter (dp_{resin}) was determined to be more negative; however, a positive correlation was also noted. These findings suggest that there exists an optimum limit for E_{cell}, $Time_{total}$, and dp_{resin}, at which the ED/EDI device had the peak salt removal efficiency. The effects of the three properties are in tandem with experimental reports showing that increasing the applied voltage, operational time and resin properties leads to changes in the removal efficiency. For the energy consumption data presented in **Fig. S5**, the cell voltage ranks as the most important factor contribution to energy consumption in ED/EDI device while the number of cells in the ED/EDI stack contributes most to the thermodynamic efficiency. Generally, the different feature studied showed wide ranging impact of the three different device metrics and requiring a well guided approach to select the optimal values thus necessitating a need for optimization algorithm as posited in Section 3.5. Also, though we have small dataset for the fine-tuning stage, it is advisable to revisit the conclusion with an experimental database with more varied features especially the device parameters, which is suggested for future research in this area.

3.5. Device optimization

Monitoring whether the model is extrapolating when predicting a new configuration in device optimization is important for the modeling framework to be useful in the development of new technologies. Herein, to get a better understanding of the synthetic data used to train the source model, we explored the use of 2D spatial distribution of the data and determination of the pareto optimal and nearest neighbor distance. **Fig. 12(a)** shows the PaCMAP projection for the data used to train the source surrogate model. PaCMAP was selected because it is known to preserve both the local and global structure of the data in the original space. As shown in **Fig. 12(a)**, the pretraining data is evenly distributed across the space defined by the factorial design used to generate it. However, some empty space can be found within the range of the reduced dimensions. In terms of extrapolation, points located in these empty regions, as well as points outside the range, could be designated as points of extrapolation applying the kNN distance threshold criteria. Furthermore, from the pretraining data, a kNN distance threshold was obtained using the knee method [51] and shown in **Fig. 12(b)**. The threshold is then defined to be the inflection point in the kNN distance versus sample number and approximates the kNN distance for the outermost point that belongs to a cluster in the training set and all points beyond this distance are considered POEs. As it will be shown in the following section, the extrapolation information is passed to the optimization algorithm to encourage it to find good solutions while minimizing the degree of extrapolation.
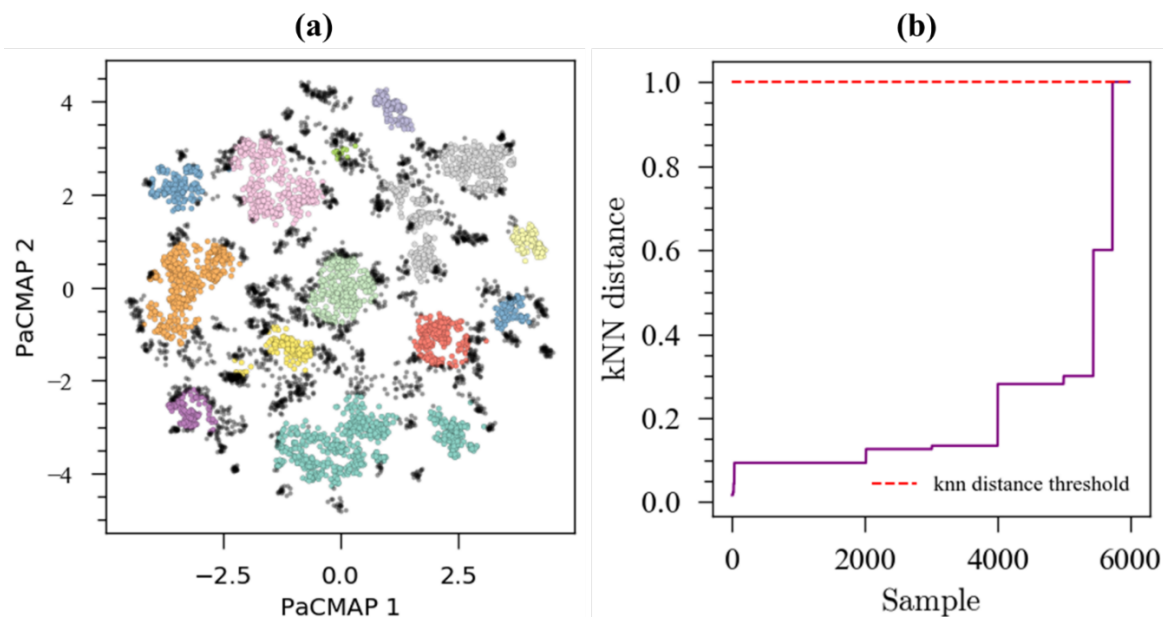
Fig. 12. Extrapolation analysis of the dataset and surrogate model in optimization task (a): PaCMAP 2D projection for the pretraining data with different colors representing different HDBSCAN-based clusters. (b) Computed distances between samples and threshold distance based on k-Nearest Neighbor algorithm.

During device optimization, a given candidate solution would be considered a point of extrapolation if its k-nearest neighbor distance with respect to the source training data is greater than a predetermined threshold shown in **Fig. 12(b)**. To help in the development of better electrochemical desalination processes, this goal takes the form of exploring the input space of the models to find combinations that lead to good performance metrics. To facilitate this analysis, we once again consider the surrogate model and a Pareto front analysis (as shown in Fig. 13) targeting high separation efficiency, low energy consumption and high thermodynamic efficiency. The surrogate model shows that high removal efficiency (> 90%) is generally associated with low thermal energy efficiency and high energy consumption. This observation between the removal efficiency and thermodynamic energy efficiency agrees with reported values (**Fig. S10**) in ref [42] where the author reported low characteristic thermodynamic efficiency with increasing separation efficiency at any water recovery rate. To conclude, we found a maximum of separation efficiency at 99% at energy consumption level lower than 1 kWh/kg using the proposed methodology.

(a)

(b)
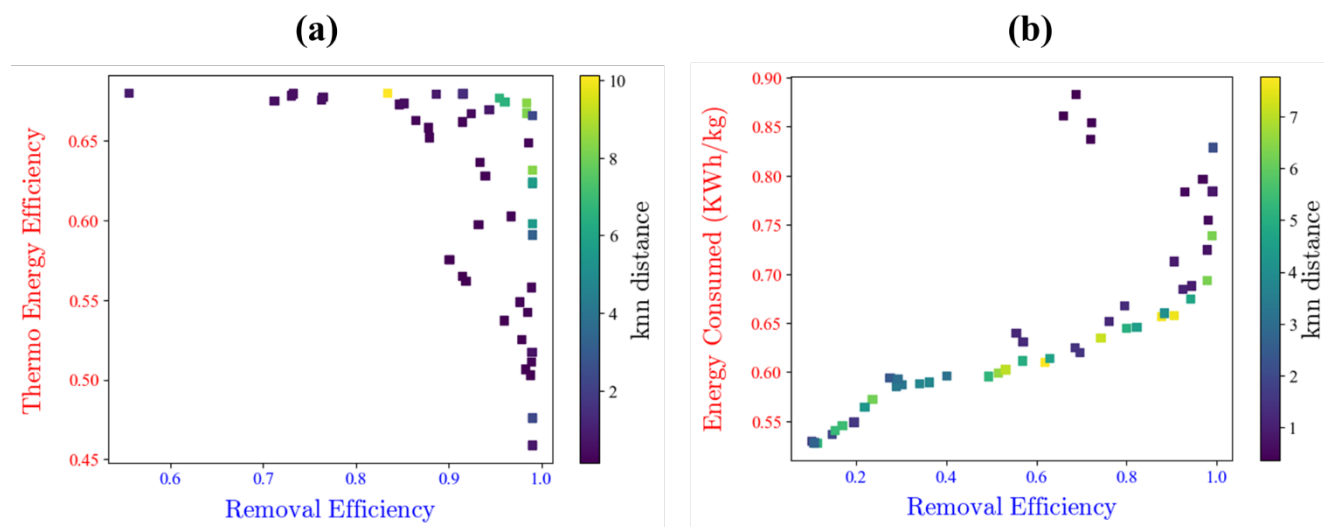
Fig. 13. Pareto front showing optimal trade-offs between different combinations of high removal efficiency, low energy consumption and high thermodynamic efficiency. The Pareto points are colored by their KNN distance.

4. Conclusions

A hybrid machine learning approach has been devised to model electrodialysis (ED) and electrodeionization (EDI) employing different resin wafer (RW) materials for brackish water desalination. This approach not only delves into the diverse EDI operating parameters affecting desalination performance but also lays emphasis on the initial data generation using a knowledge-based model. The process initiates with the creation of a source model through pretraining, culminating in a final target model capable of accurately predicting experimental results. The efficacy of this methodology is underpinned by thorough analyses of model accuracy, interpretation, as well as optimization and extrapolation monitoring exercises. This automated model synthesis strategy holds promise for conducting preliminary analyses of selected material and operational parameters for electrochemical separation devices.

Abbreviations

Parameter	Interpretation
ML	Machine Learning
TL	Transfer Learning
$J_{bulk,i}$	net convective flow of component i in and out of the compartment (mol/s.m^3)
$J_{mig,i}$	migration of ions i through the AEM and CEM membranes (mol/s.m^3)
$J_{diff,i}$	diffusion of ions i from bulk to AEM/CEM interface (mol/.m^3)
C_{dil}	outlet bulk concentration (mol/m^3) of the diluate compartment
C_{con}	outlet bulk concentration (mol/m^3) of the concentrate compartment
C_{dil}^{inlet}	inlet bulk concentrations (mol/m^3) of the diluate compartment
C_{conc}^{inlet}	inlet bulk concentrations (mol/m^3) of the concentrate compartment
C_{con}^{wc}	concentration at the CEM/concentrate interface (mol/m^3)
C_{dil}^{wc}	concentration at the CEM/diluate interface (mol/m^3)
C_{con}^{wa}	concentration at AEM/concentrate interface (mol/m^3)
C_{dil}^{wa}	concentration at AEM/diluate interface (mol/m^3)
l_a	AEM thickness (m)
l_c	CEM thickness (m)
N	Number of cell pairs
V	Applied voltage
j	Current density (A/m^2)
ϕ	current efficiency
α_i	Perm-selectivity of ion (i) in membrane
z	charge number
D_a	diffusion coefficient of ions in AEM (m^2/s)
D_c	diffusion coefficient of ions in CEM (m^2/s)
A	active membrane area (m^2)

C_0	Inlet concentration into diluate stream (mol/m^3)
C_B	Final concentration in the concentrate stream (mol/m^3)
C_D	Final outlet concentration in the diluate stream (mol/m^3)
m	mass of ion removed (kg)
WR	Water recovery
SR	Salt removal efficiency
EC	Energy consumed (KWh/kg)
TEE_{max}	Maximum theoretical thermodynamic efficiency
$\overline{\Delta V_R}(WR, SR)$	minimum mean voltage for a reversible separation
v	the v is the total charge
e	charge on the electron
k_B	Boltzmann constant (J/K)
F	Faraday constant (C/mol)
R	Universal Gas constant (J/mol-K)
POE	Point of extrapolation
ED/EDI	Electrodialysis / Electrodeionization
ML	Machine Learning
TL	Transfer Learning

Data Availability

The datasets and code generated during this study are available at GitHub:

Acknowledgments

References

[1] A. Campione, L. Gurreri, M. Ciofalo, G. Micale, A. Tamburini, A. Cipollina, Electrodialysis for water desalination: A critical assessment of recent developments on process fundamentals, models and applications, Desalination 434 (2018) 121–160. https://doi.org/10.1016/j.desal.2017.12.044.

[2] R.P. Lively, The refinery of today, tomorrow, and the future: A separations perspective, AIChE J. 67 (2021) e17286. https://doi.org/10.1002/aic.17286.

[3] M.A. Alkhadra, X. Su, M.E. Suss, H. Tian, E.N. Guyes, A.N. Shocron, K.M. Conforti, J.P. de Souza, N. Kim, M. Tedesco, K. Khoiruddin, I.G. Wenten, J.G. Santiago, T.A. Hatton, M.Z. Bazant, Electrochemical Methods for Water Purification, Ion Separations, and Energy

Conversion, Chem. Rev. 122 (2022) 13547–13635. https://doi.org/10.1021/acs.chemrev.1c00396.

[4] V.M. Palakkal, L. Valentino, Q. Lei, S. Kole, Y.J. Lin, C.G. Arges, Advancing electrodeionization with conductive ionomer binders that immobilize ion-exchange resin particles into porous wafer substrates, Npj Clean Water 3 (2020) 1–10. https://doi.org/10.1038/s41545-020-0052-z.

[5] W. Li, H. Yusha, L. Yifei, F. Jiangtao, H. Ning, M. Li, Desalination mechanism of modified activated carbon/carbon nanotubes composite electrode, Water Supply 19 (2019) 2054–2060. https://doi.org/10.2166/ws.2019.082.

[6] T.A. Saleh, M. Mustaqeem, M. Khaled, Water treatment technologies in removing heavy metal ions from wastewater: A review, Environ. Nanotechnol. Monit. Manag. 17 (2022) 100617. https://doi.org/10.1016/j.enmm.2021.100617.

[7] H.I. Uzun, E. Debik, Economical approach to nitrate removal via membrane capacitive deionization, Sep. Purif. Technol. 209 (2019) 776–781. https://doi.org/10.1016/j.seppur.2018.09.037.

[8] L. Chen, C. Wang, S. Liu, L. Zhu, Investigation of adsorption/desorption behavior of Cr(VI) at the presence of inorganic and organic substance in membrane capacitive deionization (MCDI), J. Environ. Sci. 78 (2019) 303–314. https://doi.org/10.1016/j.jes.2018.11.005.

[9] S.-J. Kim, J.-H. Choi, J.-H. Kim, Removal of acetic acid and sulfuric acid from biomass hydrolyzate using a lime addition–capacitive deionization (CDI) hybrid process, Process Biochem. 47 (2012) 2051–2057. https://doi.org/10.1016/j.procbio.2012.07.020.

[10] T. Kulkarni, A.M.I. Al Dhamen, D. Bhattacharya, C.G. Arges, Bipolar Membrane Capacitive Deionization for pH-Assisted Ionic Separations, ACS EST Eng. 3 (2023) 2171–2182. https://doi.org/10.1021/acsestengg.3c00041.

[11] B. Sauvet-Goichon, Ashkelon desalination plant — A successful challenge, Desalination 203 (2007) 75–81. https://doi.org/10.1016/j.desal.2006.03.525.

[12] S.-Y. Pan, S.W. Snyder, H.-W. Ma, Y.J. Lin, P.-C. Chiang, Energy-efficient resin wafer electrodeionization for impaired water reclamation, J. Clean. Prod. 174 (2018) 1464–1474. https://doi.org/10.1016/j.jclepro.2017.11.068.

[13] R. Zhao, S. Porada, P.M. Biesheuvel, A. van der Wal, Energy consumption in membrane capacitive deionization for different water recoveries and flow rates, and comparison with reverse osmosis, Desalination 330 (2013) 35–41. https://doi.org/10.1016/j.desal.2013.08.017.

[14] Y.-J. Kim, J.-H. Kim, J.-H. Choi, Selective removal of nitrate ions by controlling the applied current in membrane capacitive deionization (MCDI), J. Membr. Sci. 429 (2013) 52–57. https://doi.org/10.1016/j.memsci.2012.11.064.

[15] D.-H. Lee, T. Ryu, J. Shin, J.C. Ryu, K.-S. Chung, Y.H. Kim, Selective lithium recovery from aqueous solution using a modified membrane capacitive deionization system, Hydrometallurgy 173 (2017) 283–288. https://doi.org/10.1016/j.hydromet.2017.09.005.

[16] W. Tang, P. Kovalsky, B. Cao, D. He, T.D. Waite, Fluoride Removal from Brackish Groundwaters by Constant Current Capacitive Deionization (CDI), Environ. Sci. Technol. 50 (2016) 10570–10579. https://doi.org/10.1021/acs.est.6b03307.

[17] J.J. Lado, R.E. Pérez-Roa, J.J. Wouters, M.I. Tejedor-Tejedor, C. Federspill, J.M. Ortiz, M.A. Anderson, Removal of nitrate by asymmetric capacitive deionization, Sep. Purif. Technol. 183 (2017) 145–152. https://doi.org/10.1016/j.seppur.2017.03.071.

[18] P.-C. Tseng, Z.-Z. Lin, T.-L. Chen, Y. Lin, P.-C. Chiang, Performance evaluation of resin wafer electrodeionization for cooling tower blowdown water reclamation, Sustain. Environ. Res. 32 (2022) 36. https://doi.org/10.1186/s42834-022-00145-8.

[19] T. Ho, A. Kurup, T. Davis, J. Hestekin, Wafer Chemistry and Properties for Ion Removal by Wafer Enhanced Electrodeionization, Sep. Sci. Technol. 45 (2010) 433–446. https://doi.org/10.1080/01496390903526709.

[20] A.S. Kurup, T. Ho, J.A. Hestekin, Simulation and Optimal Design of Electrodeionization Process: Separation of Multicomponent Electrolyte Solution, Ind. Eng. Chem. Res. 48 (2009) 9268–9277. https://doi.org/10.1021/ie801906d.

[21] J.M. Ortiz, J.A. Sotoca, E. Expósito, F. Gallud, V. García-García, V. Montiel, A. Aldaz, Brackish water desalination by electrodialysis: batch recirculation operation modeling, J. Membr. Sci. 252 (2005) 65–75. https://doi.org/10.1016/j.memsci.2004.11.021.

[22] N.C. Wright, S.R. Shah, S.E. Amrose, A.G. Winter, A robust model of brackish water electrodialysis desalination with experimental comparison at different size scales, Desalination 443 (2018) 27–43. https://doi.org/10.1016/j.desal.2018.04.018.

[23] M.M. Generous, N.A.A. Qasem, S.M. Zubair, The significance of modeling electrodialysis desalination using multi-component saline water, Desalination 496 (2020) 114347. https://doi.org/10.1016/j.desal.2020.114347.

[24] E. Glueckauf, Electro-deionization through a packed bed, Br. Chem. Eng. 4 (1959) 646–651.

[25] I. Monzie, L. Muhr, F. Lapicque, G. Grévillot, Mass transfer investigations in electrodeionization processes using the microcolumn technique, Chem. Eng. Sci. 60 (2005) 1389–1399. https://doi.org/10.1016/j.ces.2004.10.010.

[26] H.M. Verbeek, L. Fürst, H. Neumeister, Digital simulation of an electrodeionization process, Comput. Chem. Eng. 22 (1998) S913–S916. https://doi.org/10.1016/S0098-1354(98)00179-3.

[27] L. Yao, Y. Li, Q. Cheng, Z. Chen, J. Song, Modeling and optimization of metal-organic frameworks membranes for reverse osmosis with artificial neural networks, Desalination 532 (2022) 115729. https://doi.org/10.1016/j.desal.2022.115729.

[28] J. Wang, K. Tian, D. Li, M. Chen, X. Feng, Y. Zhang, Y. Wang, B. Van der Bruggen, Machine learning in gas separation membrane developing: Ready for prime time, Sep. Purif. Technol. 313 (2023) 123493. https://doi.org/10.1016/j.seppur.2023.123493.

[29] A. Mistry, A.A. Franco, S.J. Cooper, S.A. Roberts, V. Viswanathan, How Machine Learning Will Revolutionize Electrochemical Sciences, ACS Energy Lett. 6 (2021) 1422–1431. https://doi.org/10.1021/acsenergylett.1c00194.

[30] L.A. Briceno-Mena, G. Venugopalan, J.A. Romagnoli, C.G. Arges, Machine learning for guiding high-temperature PEM fuel cells with greater power density, Patterns (2021) 100187. https://doi.org/10.1016/j.patter.2020.100187.

[31] L.A. Briceno-Mena, J.A. Romagnoli, C.G. Arges, PemNet: A Transfer Learning-Based Modeling Approach of High-Temperature Polymer Electrolyte Membrane Electrochemical Systems, Ind. Eng. Chem. Res. 61 (2022) 3350–3357. https://doi.org/10.1021/acs.iecr.1c04237.

[32] Q. Yang, An Introduction to Transfer Learning, in: C. Tang, C.X. Ling, X. Zhou, N.J. Cercone, X. Li (Eds.), Adv. Data Min. Appl., Springer, Berlin, Heidelberg, 2008: pp. 1–1. https://doi.org/10.1007/978-3-540-88192-6_1.

[33] Q. Yang, Y. Zhang, W. Dai, S.J. Pan, Transfer Learning, Cambridge University Press, Cambridge, 2020. https://doi.org/10.1017/9781139061773.

[34] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How transferable are features in deep neural networks?, in: Adv. Neural Inf. Process. Syst., Curran Associates, Inc., 2014. https://proceedings.neurips.cc/paper_files/paper/2014/hash/375c71349b295fbe2dcdca9206f20a06-Abstract.html (accessed May 15, 2023).

[35] J. Shi, F. Albreiki, Yamil J. Colón, S. Srivastava, J.K. Whitmer, Transfer Learning Facilitates the Prediction of Polymer–Surface Adhesion Strength, J. Chem. Theory Comput. (2023). https://doi.org/10.1021/acs.jctc.2c01314.

[36] L.A. Briceno-Mena, C.G. Arges, J.A. Romagnoli, Machine learning-based surrogate models and transfer learning for derivative free optimization of HT-PEM fuel cells, Comput. Chem. Eng. 171 (2023) 108159. https://doi.org/10.1016/j.compchemeng.2023.108159.

[37] L.A. Briceno-Mena, Hybrid Modeling for Electrochemical Systems, LSU Dr. Diss. (2023). https://doi.org/10.31390/gradschool_dissertations.6053.

[38] R. Ma, Y.J. Colón, T. Luo, Transfer Learning Study of Gas Adsorption in Metal–Organic Frameworks, ACS Appl. Mater. Interfaces 12 (2020) 34041–34048. https://doi.org/10.1021/acsami.0c06858.

[39] H. Yamada, C. Liu, S. Wu, Y. Koyama, S. Ju, J. Shiomi, J. Morikawa, R. Yoshida, Predicting Materials Properties with Little Data Using Shotgun Transfer Learning, ACS Cent. Sci. 5 (2019) 1717–1730. https://doi.org/10.1021/acscentsci.9b00804.

[40] Z. Deng, X. Lin, J. Cai, X. Hu, Battery health estimation with degradation pattern recognition and transfer learning, J. Power Sources 525 (2022) 231027. https://doi.org/10.1016/j.jpowsour.2022.231027.

[41] L. Alvarado, A. Chen, Electrodeionization: Principles, Strategies and Applications, Electrochimica Acta 132 (2014) 583–597. https://doi.org/10.1016/j.electacta.2014.03.165.

[42] S. Lin, Energy Efficiency of Desalination: Fundamental Insights from Intuitive Interpretation, Environ. Sci. Technol. 54 (2020) 76–84. https://doi.org/10.1021/acs.est.9b04788.

[43] T. O'Malley, E. Bursztein, J. Long, F. Chollet, H. Jin, L. Invernizzi, others, KerasTuner, (2019). https://github.com/keras-team/keras-tuner.

[44] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, (2016). https://doi.org/10.48550/arXiv.1603.04467.

[45] S.M. Lundberg, S.-I. Lee, A Unified Approach to Interpreting Model Predictions, in: Adv. Neural Inf. Process. Syst., Curran Associates, Inc., 2017. https://papers.nips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html (accessed May 15, 2023).

[46] D. Wang, S. Thunéll, U. Lindberg, L. Jiang, J. Trygg, M. Tysklind, Towards better process management in wastewater treatment plants: Process analytics based on SHAP values for tree-based machine learning methods, J. Environ. Manage. 301 (2022) 113941. https://doi.org/10.1016/j.jenvman.2021.113941.

[47] J.W. Barnett, C.R. Bilchak, Y. Wang, B.C. Benicewicz, L.A. Murdock, T. Bereau, S.K. Kumar, Designing exceptional gas-separation polymer membranes using machine learning, Sci. Adv. 6 (2020) eaaz4301. https://doi.org/10.1126/sciadv.aaz4301.

[48] S.K. Khanal, A. Tarafdar, S. You, Artificial intelligence and machine learning for smart bioprocesses, Bioresour. Technol. 375 (2023) 128826. https://doi.org/10.1016/j.biortech.2023.128826.

[49] A. Moez, PyCaret: An open source, low-code machine learning library in Python, (2020). https://www.pycaret.org.

[50] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, İ. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods 17 (2020) 261–272. https://doi.org/10.1038/s41592-019-0686-2.

[51] V. Satopaa, J. Albrecht, D. Irwin, B. Raghavan, Finding a "Kneedle" in a Haystack: Detecting Knee Points in System Behavior, in: 2011 31st Int. Conf. Distrib. Comput. Syst. Workshop, 2011: pp. 166–171. https://doi.org/10.1109/ICDCSW.2011.20.