Article

Predicting Membrane Fouling of Submerged Membrane Bioreactor Wastewater Treatment Plants Using Machine Learning

Yunyi Zhu, Yuan Wang, Elisabeth Zhu, Zeyu Ma, Hanchen Wang, Chunsheng Chen, Jing Guan, and T. David Waite*

Cite This: https://doi.org/10.1021/acs.est.4c12835
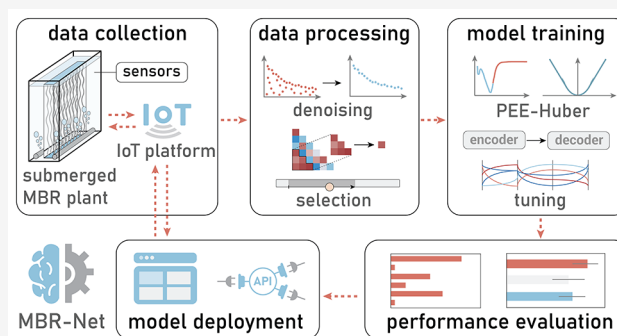
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Membrane fouling remains a significant challenge in the operation of membrane bioreactors (MBRs). Plant operators rely heavily on observations of filtration performance from noisy sensor data to assess membrane fouling conditions and lab-based protocols for plant maintenance, often leading to inaccurate estimations of future performance and delayed membrane cleaning. This challenge is further compounded by the difficulty in integrating existing complex mechanistic models with the Internet of Things (IoT) systems of wastewater treatment plants (WWTPs). By harnessing data obtained from WWTPs, along with innovative data denoising and model training strategies, we developed a machine learning application (MBR-Net) that is capable of forecasting membrane fouling, as indicated by permeability, for a full-scale submerged MBR plant in real time. We show that the trained model can effectively predict one-day-ahead changes in irreversible fouling under different desired fluxes, cleaning conditions and feedwater conditions (with MAPE < 6.45%, MAE < 3.71 LMH bar^{-1}, and R^2 > 0.87 on two independent testing sets). Although data availability presented certain limitations in the model development process, the current results demonstrate the significant value of machine learning in membrane fouling predictions and in providing decision support for fouling mitigation strategies in full-scale WWTPs.

KEYWORDS: membrane bioreactor (MBR), membrane, digital twin, Seq2Seq model, time-series forecasting, LSTM, Internet of Things

INTRODUCTION

Membrane bioreactors (MBRs) are widely used for domestic and industrial wastewater treatment given their ability to generate effluent of superior quality as well as their small footprint and similar capital cost compared with traditional activated sludge treatment systems.[1,2] To date, approximately 100 large and superlarge-scale MBR plants (>50,000 m^3 day^{-1}; SI, Figure S1) and thousands of smaller full-scale commercial MBR plants have been implemented in municipal wastewater treatment worldwide.[3-5] In view of the enactment of increasingly stringent environmental regulations and the increasing scarcity of available water and land, it can be expected that the global MBR market will expand substantially in the coming years. However, the unavoidable decline in membrane filtration performance due to membrane fouling is recognized to be a significant hurdle for MBR operation, which further increases its energy consumption, causes unnecessary loss of productivity, and may well limit the broader use of this technology, especially in less developed regions. Membrane fouling is caused by the deposition of particulate and organic materials either on the membrane surfaces or within the internal structure of the membrane during operation, leading to an increase in the resistance to the passage of water through

the fouled membranes. The increase in membrane resistance (or decrease in membrane permeability) results in a higher transmembrane pressure (TMP) if the MBR is operated under constant flux or a lower flux over time if the MBR is operated under constant TMP. Membrane fouling consists of both recoverable fouling and irrecoverable fouling, depending on whether the fouling can be removed by any mitigation strategies.[6,7] Recoverable fouling can be further separated into four types (Figure 1a):[8-10]

- reversible fouling that can be removed by physical cleaning approaches such as regular cessation of permeation by turning off filtration (a process known as "relaxation") together with continuous air scouring and/or sometimes backwashing (although backwashing has been suggested to be detrimental to fouling control
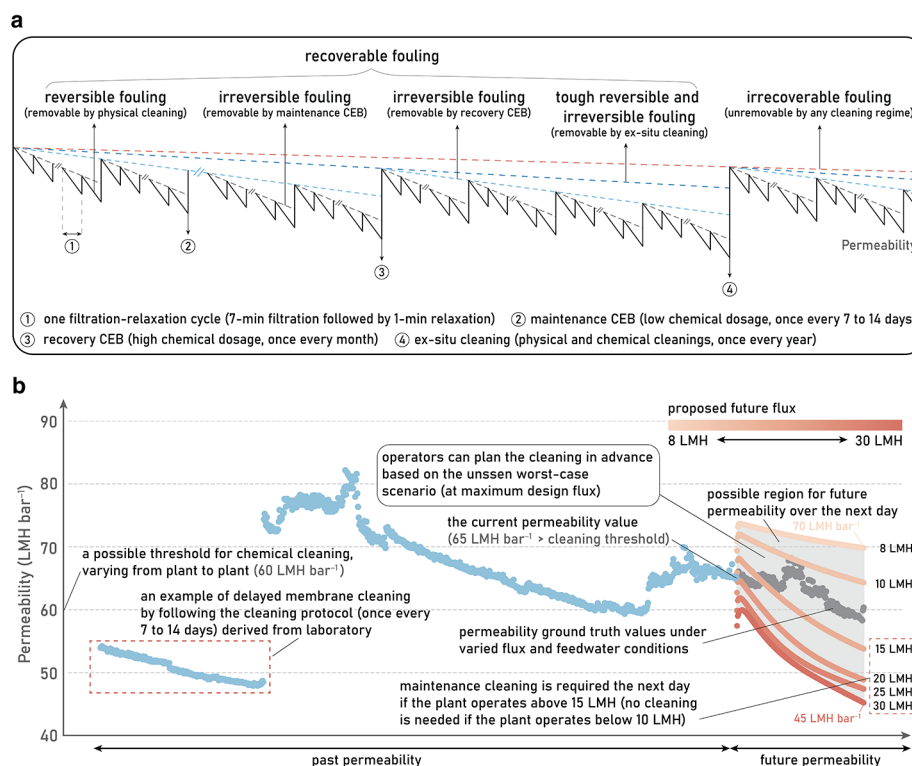
Figure 1. MBR membrane fouling. (a) Accumulation and mitigation of different types of MBR membrane fouling, as indicated by permeability, during long-term MBR operation under ideal constant-flux conditions. Red dashed line: baseline for irrecoverable fouling that cannot be removed by any cleaning approaches; dark-blue dashed line: baseline for tough reversible and irreversible fouling that requires ex situ cleaning to remove; light-blue dashed line: baseline for irreversible fouling that requires recovery by CEB to remove; gray dashed line: baseline for irreversible fouling that requires more frequent maintenance by CEB to remove. (b) Illustration of the possible region for future permeability under different operational conditions with different fixed future flux settings. Blue dots: past permeability; gray dots: permeability ground truth values under varied feedwater conditions and flux settings (flux was adjusted accordingly to meet the required quality and productivity, considering the varied inflow quality and volume to the WWTP); red dots with different color gradients: future permeability (representing membrane fouling) under fixed proposed flux values ranging from 8 to 30 LMH.

in some cases and was not applied in the MBR plants examined in this work[11]);

- irreversible fouling that is removable by in situ "maintenance" chemical cleaning methods such as cleaning in place (CIP: chemical cleaning without draining the membrane tank) and chemically enhanced backwash (CEB: chemical cleaning in conjunction with physical cleaning);
- irreversible fouling that can be removed by in situ "recovery" chemical cleaning methods, which are more intensive with a much higher chemical dosage/loading compared with "maintenance" cleaning; and
- tough reversible and irreversible fouling that requires removing MBR modules from the setup and performing intensive physical and chemical cleaning measures ex situ.

Pressure applied to drive membrane filtration is adjusted in a real wastewater treatment plant (WWTP) operation to meet effluent requirements, considering the frequently varied inflow quality and volume. Previous studies have shown that the compressibility of the fouling layer results in rapid changes in specific filtration resistance, rendering the precise assessment of membrane fouling difficult.[12−14] Consequently, plant operators face challenges in making informed decisions regarding fouling mitigation strategies. Membrane cleaning activities of full-scale plants are often conducted based on

protocols derived from laboratory/pilot studies, or when the permeability (or TMP) reaches a predefined threshold. However, since actual operational and maintenance conditions vary across different plants and altered fluxes (which affect the compressibility of the fouling layer due to corresponding changes in TMP) are also required to manage the uneven volume of municipal wastewater with time within the design range of the WWTP, such protocols often result in delayed membrane cleanings (Figure 1b). With the ongoing advancement of Industry 4.0, WWTPs are increasingly equipped with IoT sensors and cloud-based data storage systems; nevertheless, the captured operational data is typically noisy (SI, Figure S2c,e), which cannot be used by operators to accurately assess and project membrane fouling, and consequently, plant operators face challenges in making informed decisions regarding fouling mitigation strategies.

To overcome this aforementioned challenge and improve the fouling assessment accuracy, researchers have extensively investigated advanced techniques for membrane fouling modeling and simulation that, potentially, provide insights into the impact of different operational and maintenance conditions on membrane filtration performance, thereby facilitating the establishment of a more efficient cleaning protocol tailored to the specific fouling conditions of different MBR tanks (also known as MBR trains). However, the complexity of the fouling mechanism and the absence of a
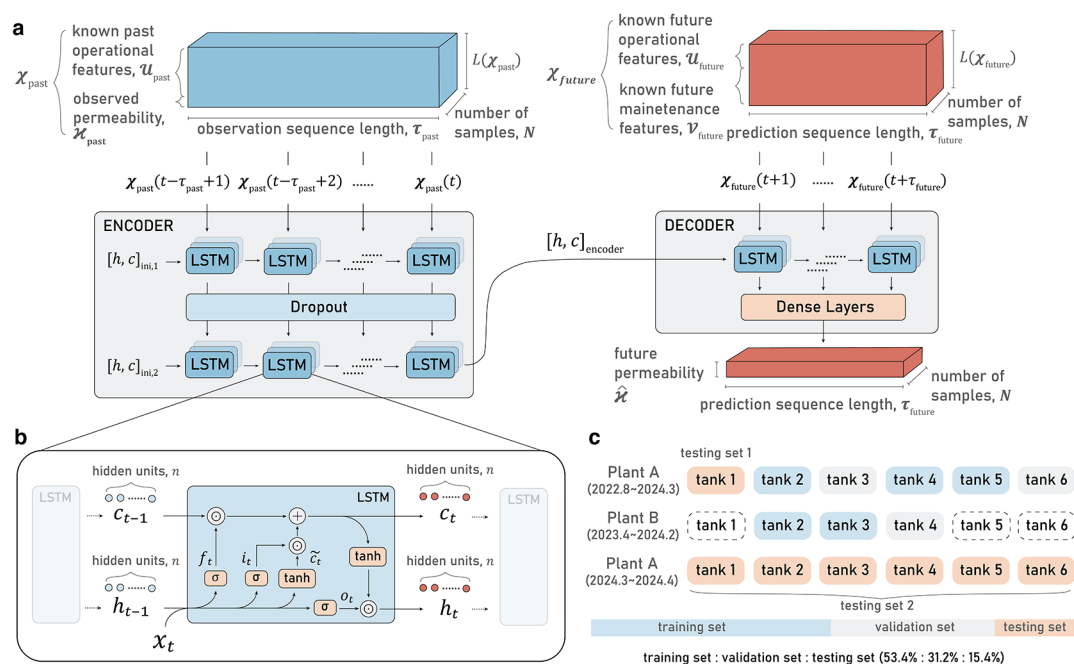
Figure 2. Overview of the model structure and data set details. (a) Structure of the underlying ML model. Blue cuboid: encoder input tensor consisting of known past operational features and observed permeability; red bulky cuboid: decoder input tensor consisting of known future operational features and known future maintenance features; red slender cuboid: decoder output tensor containing future permeability predictions. (b) Structure of LSTM cell. (c) Illustration of data splitting. Blue blocks: MBR tanks included in the training set; gray blocks: MBR tanks included in the validation set; apricot blocks: MBR tanks included in the testing set; white blocks with dashed outline: MBR tanks excluded in the data set due to missing information or messy data patterns.

universally accepted fouling indicator render membrane fouling modeling challenging.[9] The resistance-in-series (RIS) modeling coupled with mass balance analysis has been widely used to model membrane fouling. These empirical (pseudo-mechanistic) models require either the determination of numerous parameters through laboratory experiments or assumptions.[15,16] In view of potential differences between experimental conditions and actual operational conditions, model calibration is necessary before applying these models to support full-scale MBR operation and maintenance. Only a few studies aimed at modeling membrane fouling of full-scale plants have been reported.[17−20] However, these developed models often encounter two main issues: (1) poor performance in making predictions after chemical cleaning is applied as both cake layer resistance and internal fouling resistance would depend on the effectiveness of the cleaning and require redetermination via laboratory pure-water filtration after each cleaning step; and (2) failure to capture the impact of immediate change in operational conditions (e.g., varied TMP values to alter flux in order to manage the uneven volume of municipal wastewater with time) on membrane fouling as the cake layer density is typically assumed to be constant without accounting for the compression or swelling of the cake in most applied RIS models.

A data-driven machine learning (ML) approach has recently been applied as an alternative for modeling membrane fouling in pilot- or full-scale MBR plants.[21−26] The advantage of an ML approach lies in its ability to learn relationships between different variables underlying the fouling mechanism from large amounts of historical data, without requiring any assumption or laboratory-based determination of specific system parameters, as is the case with black-box modeling

approaches. However, most existing studies used data sets collected from a single membrane tank, and the generalizability of the models developed was not well evaluated. In addition, none of these models consider the temporal dependencies of membrane fouling or are able to generate reliable multihorizon (i.e., multistep-ahead) membrane fouling predictions based on past membrane fouling conditions and different future operational and maintenance conditions of MBR plants.

To address these significant deficiencies in MBR membrane fouling modeling and support decision-making regarding membrane maintenance, we developed an ML application (MBR-Net) to achieve real-time membrane fouling predictions for submerged MBR plants. Various water quality, operational, and maintenance data spanning over 2 years were obtained from multiple MBR trains of two full-scale WWTPs and subjected to further processing prior to model development. The underlying long short-term memory (LSTM) encoder−decoder model (Figure 2a,b) was used to generate membrane fouling predictions by forecasting permeability 1 day ahead based on the weekly performance history of the MBR plant. The performance of the optimized model was evaluated across different MBR tanks prior to model deployment. Subsequently, we developed a web interface integrated with the WWTP Internet of Things (IoT) platform to enable operators to interact with the model and provide prompt visualization of model prediction results.

■ MATERIALS AND METHODS

MBR Treatment Plants. The data used in this study were collected from two submerged MBR plants equipped with hollow fiber polyvinylidene difluoride microfiltration membranes with a nominal pore size of 0.4 μm, operated by Beijing

Origin Water in central and south China. These two plants had an anaerobic−anoxic−aerobic−anoxic (AAOA) process as the primary treatment phase followed by six submerged MBR tanks/trains and an ultraviolet disinfection process to produce the treated effluent with a design capacity of larger than 50,000 m^3 day^{-1} for each plant. Each MBR tank of Plant A contained ten membrane modules, whereas each MBR tank of Plant B contained nine membrane modules. Each module consisted of 60 membrane elements, each with a surface area of 35 m^2, providing a total membrane area of 2100 m^2 per module (SI, Figure S1a−e). The feedwater type of both plants was domestic wastewater. The plant operation and maintenance information is provided in the SI, Text S2.

Data Processing. The membrane filtration performance and fouling can be indicated by changes in permeability, $\varkappa$ (LMH bar^{-1}), which is given by Darcy's law:

$$\varkappa = \frac{J}{TMP} = \frac{1}{\eta \cdot R} \tag{1}$$

where the flux J (L m^{-2} h^{-1}, LMH) is the quantity of material passing through a unit area of membrane per unit time, and the TMP (bar) is the driving force of the filtration process (defined as the difference between the gauge pressure during relaxation and the vacuum pressure during filtration). As shown in eq 1, permeability can also be expressed using viscosity η (kg m^{-1} s^{-2}) and resistance R (m^{-1}). A data processing algorithm was applied to denoise the raw data and extract meaningful data points of each filtration−relaxation cycle (SI, Figure S2a−f and Supplementary Note S1). Two offset state parameters, determined through trial and error for each MBR tank, were used for data denoising of plant A to account for flow rate and pressure fluctuations that represent the real operating scenarios rather than leaving the noise caused by sensor error in the data set. Additionally, two adjustable parameters, window size and threshold, were used in the rolling window approach. Values of these two parameters were also determined by trial and error and controlled to be a relatively small number compared with the observation sequence length and data ranges, respectively, ensuring that important peak values instead of noise could be retained. All processed data sets were manually reviewed to identify whether data points unrecognized by the algorithms were noise or valuable observations.

Although the same type of MBR modules were installed in each membrane tank of both Plant A and Plant B, the fouling behavior of each tank varied due to differences in operational and maintenance conditions among the tanks, such as different service times of the modules, different TMP values applied for membrane filtration, different cleaning frequencies, and so on. Therefore, we divided the entire data set after denoising into training, validation, and testing sets based on individual MBR tanks (Figure 2c):

- training set: MBR Tanks 2, 4, and 5 of Plant A from August 2022 to March 2024; MBR Tanks 2 and 3 of Plant B from April 2023 to February 2024;

- validation set: MBR Tanks 3 and 6 of Plant A from August 2022 to March 2024; MBR Tank 4 of Plant B from April 2023 to February 2024;

- testing set: MBR Tank 1 of Plant A from August 2022 to March 2024 (testing set 1); all MBR tanks of Plant A from March to April 2024 (testing set 2).

By separating the data sets in this manner, we ensured that the validation set and testing sets remained unseen during parameter fitting in the model training phase as they originated from distinct tanks or time periods. Additionally, considering the time-series nature of our data set, we chose a structured data split based on tanks and time periods, rather than using conventional random splitting or cross-validation approaches, to preserve the temporal integrity of each sub-dataset. If sequences from the same tank with closely related timestamps were randomly split, then there would be a risk of overlapping time-series data, which could result in data from the validation/testing set being included in the training set, thereby artificially improving performance metrics. Therefore, the applied structured split approach avoids the issue of data leakage and ensures a more reliable evaluation of model performance.[27] Permeability, flux, and feed NH_4^+-N of each sub-dataset have relatively similar ranges when compared among different sub-datasets, despite exhibiting different distributions (SI, Supplementary Figure S3c−e). Due to difficulties in recognizing each operational period between two in situ CEBs and missing maintenance information, data obtained from Tanks 1, 5, and 6 of Plant B were excluded in data processing. More details regarding data collection, analysis, and processing can be found in SI Text S2.

ML Model. *LSTM Encoder−Decoder Model Structure for Multihorizon Prediction.* We used an encoder−decoder or sequence-to-sequence (Seq2Seq) structure for the process of heterogeneous past observed and known future sequential data to generate a multihorizon prediction of the target variable (i.e., permeability) (Figure 2a,b). This structure was extracted from the temporal fusion transformer (TFT) that was proven to be effective for multihorizon forecasting across various use cases.[28] The encoder consists of two layers of 900 LSTM cells, and the decoder consists of one layer of 180 LSTM cells, with each cell assigned to process the data at a specific time step in the past or in the future (SI, Text S3). Zero-filled initial state vectors were used to initialize the recurrent process of the LSTM encoder, while the final hidden state and cell state of the LSTM encoder condense the observation sequence's information and serve as the context vector $[h, c]_{encoder}$ to initialize the LSTM decoder.

Encoder features (χ_{past}) include the following:

- the past observed permeability ($\varkappa_{past}$)
- two known past operational features (u_{past}): past flux and past feed NH_4^+-N in an observation window with a length of τ_{past}.

Decoder features (χ_{future}) include the following:

- two known future operational features (u_{future}): future flux and future feed NH_4^+-N.
- four known future maintenance features (v_{future}): future NaOCl (one-hot encoded), future acid (one-hot encoded), future NaOCl dosage, and future acid dosage in a prediction window with a length of τ_{future}.

The encoder tensor $E \in \mathbb{R}^{N \times \tau_{past} \times L(\chi_{past})}$ contains the processed data points/features that are retrieved from the IoT platform, whereas the decoder tensor $D \in \mathbb{R}^{N \times \tau_{future} \times L(\chi_{future})}$ contains the features that require manual input from WWTP operators or domain experts (SI, Supplementary Note S3). We simultaneously output permeability predictions $\hat{\varkappa}$ for τ_{future} steps ahead by an output tensor $O \in \mathbb{R}^{N \times \tau_{future} \times 1}$. Note that we used flux instead of TMP as a
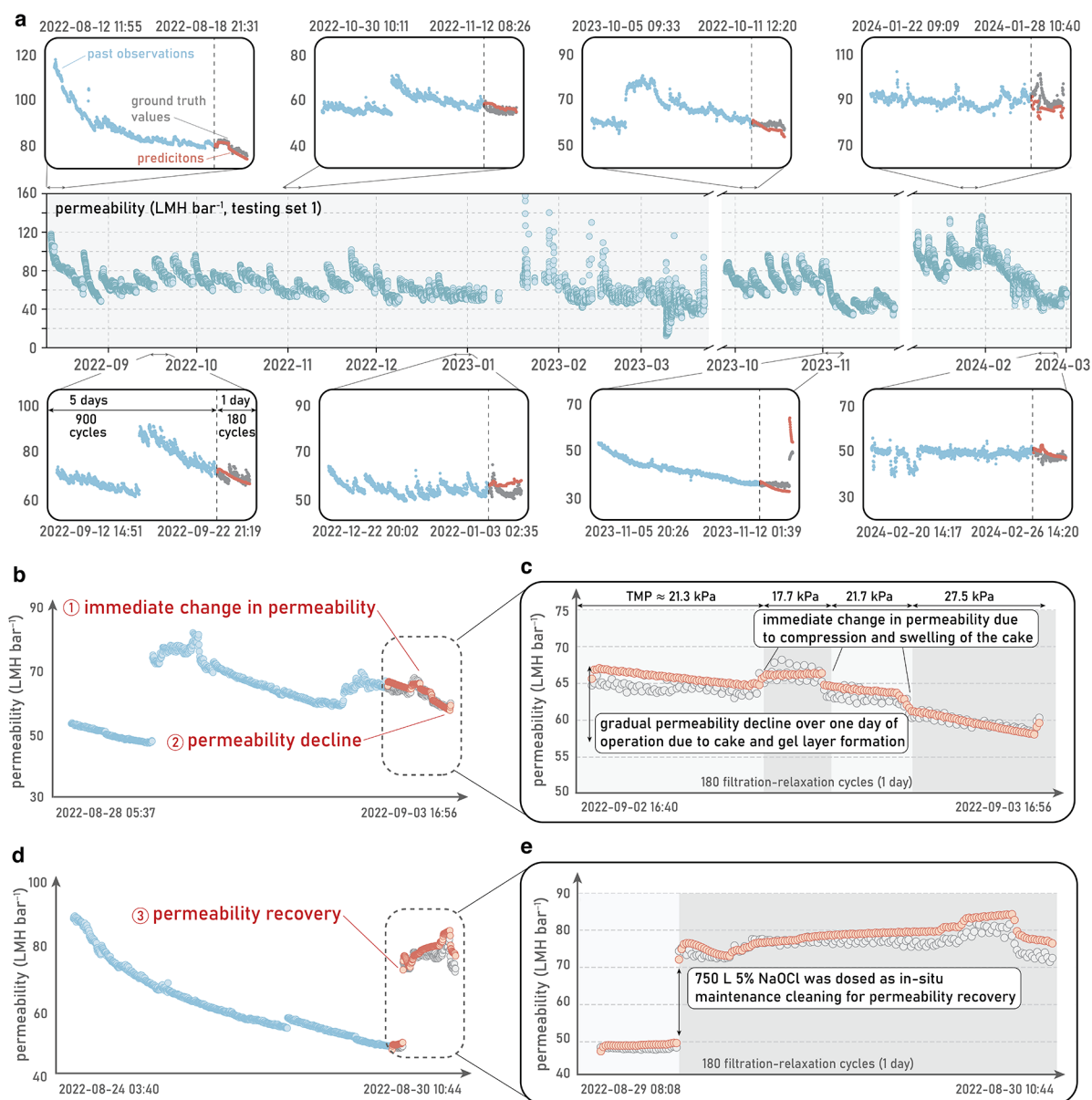
Figure 3. Model predictions and capabilities. (a) Model predictions on testing set 1 at different points of time. Blue dots: past permeability observations; red dots: permeability predictions; gray dots: permeability ground truth values. The middle panel illustrates the decline and recovery of observed permeability (teal dots) over long-term operation. Each CEB can be identified by a sudden substantial increase in the observed permeability after each continuous operation lasting for around 1 week. (b) An example of permeability predictions that demonstrates the model's capability in predicting immediate change in permeability due to cake compression and swelling and predicting the permeability decline due to gel and cake layer accumulation and pore blocking over 1 day operation. (c) A zoomed-in view of the permeability predictions with different average TMP values for each section of operation highlighted by shaded rectangles. (d) An example of permeability predictions that demonstrates the model's capability in predicting permeability recovery due to the implementation of CEB. (e) A zoomed-in view of the permeability predictions with the operational period before and after CEB, highlighted by shaded rectangles.

model feature since flux is the most direct parameter that WWTP operators will monitor when they adjust the pump frequency to change the suction pressure of each MBR tank due to sudden variations in the inflow volume.

Model Training Strategies. Due to the nature of fouling and maintenance of MBR membranes, lesser permeability data were observed in the region when $x > 100$ LMH bar^{-1} and 30 LMH bar^{-1} < x < 50 LMH bar^{-1}; however, these underrepresented data are also critical for the accurate modeling of permeability recovery after cleaning and permeability decline when the membrane was heavily fouled. Additional penalties

were required to ensure that the model does not disproportionately favor the more frequent permeability data points at the expense of the rarer, potentially more important ones. Additionally, noisy data were unavoidable even though they were greatly minimized during the data processing step, and a loss function, such as Huber loss, that is less sensitive to outliers would be expected to be highly beneficial to the improvement of model performance. Therefore, we designed and implemented a custom loss function tailored to the characteristics of the permeability data (SI, Figure S3f).

The weighted polynomial−exponential enhanced Huber loss (PEE-Huber) function, $\mathcal{L}_{\text{PEE−Huber}}$, can be expressed as follows:

$$\mathcal{L}_{\text{PEE-Huber}} = w \cdot \mathcal{L}_{\text{Huber}} \tag{2}$$

$$w = \lambda \cdot (g(x_{\text{norm}}) + \mu) \tag{3}$$

$$g(x_{\text{norm}}) = \frac{f_{\max} - f(x_{\text{norm}})}{f_{\max} - f_{\min}} \tag{4}$$

$$\mathcal{L}_{\text{Huber}} = \begin{cases} \frac{1}{2}(\varkappa - \hat{\varkappa})^2 & \text{if } |\varkappa - \hat{\varkappa}| \leq \delta \\ \delta(|\varkappa - \hat{\varkappa}| - \frac{1}{2}\delta) & \text{if } |\varkappa - \hat{\varkappa}| > \delta \end{cases} \tag{5}$$

where λ and μ are the adjustable multiplier and offset, respectively, to determine the range of weight w. δ is a parameter to determine the transition point between the quadratic region and the linear region (which is less sensitive to large errors and more robust to outliers) of $\mathcal{L}_{\text{Huber}}$[29] $f(x_{\text{norm}})$ is a smooth continuous function that approximates the frequency polygon of normalized permeability data $x_{\text{norm}} \in [0,1]$. $f_{\min}$ and $f_{\max}$ are the global minimum and maximum of $f(x_{\text{norm}})$, respectively. $\varkappa$ and $\hat{\varkappa}$ are ground truth and predicted permeability values, respectively. In this work, λ, μ, and δ were set to 5.0, 1.0, and 5.0, respectively. By deriving w from the permeability distribution, higher weights (heavier penalties) would be assigned to permeability data points that were crucial but in the minority range.[30,31] More details on model training and performance evaluation strategies can be found in the SI Text S3. Codes related to model development and evaluation are available at https://github.com/Jassssz/MBR-Net. Further details regarding the development environment and code availability can be found in SI Text S1.

Application Programming Interfaces (APIs) and Web Interface of MBR-Net. We exposed the ML model to REST APIs to listen for inbound requests and handle event-driven processes. All API end points are located at the base uniform resource locater (URL): https://<APP_NAME>.azurecontainerapps.io/. Hypertext transfer protocol (HTTP) methods such as GET, POST, PUT, and DELETE are supported at each end point to enable the interaction with the resource representation, request messages, and response messages, which are all wrapped up in JavaScript Object Notation (JSON) format (SI, Supplementary Note S2). The MBR-Net web interface serves as the primary point of interaction for users, triggering HTTP requests to APIs through user-friendly button clicks, which, in turn, execute the corresponding backend functions (SI, Supplementary Note S3). We also provide a video demonstrating how real-time membrane fouling predictions can be obtained through the MBR-Net web interface (SI Video S1). More details regarding model deployment can be found in SI Text S3.

■ RESULTS AND DISCUSSION

Impact of Immediate Flux Changes and Cleanings on Membrane Fouling. Since the inflow volume to the WWTPs varies in real operation, the flux of MBR plants must be changed accordingly to meet the wastewater treatment requirements, including effluent quality and productivity. As the driving force of water permeation, the suction pressure

applied by the filtration pump needs to be adjusted and will lead to variations in TMP. Permeability, which is inversely proportional to the membrane resistance, was hence selected as an indicator of membrane fouling in this work, rather than TMP. Continuous air scouring with intermittent relaxation (7 min ON:1 min OFF) was employed as the physical cleaning method in the studied MBR plants to remove reversible fouling and recover membrane permeability attributed to loosely attached fouling materials (Figure 1a). Since the aeration intensity and filtration−relaxation duration were maintained relatively constant for both studied plants, we did not investigate the impact of these two approaches to ameliorate reversible membrane fouling in this work. Instead, more attention was paid to the investigation of irreversible fouling.

Although cake deposition will produce fouling on the membrane surface that is mostly reversible, a small amount of cake deposition that is compressible could also contribute to the irreversible fouling, which is often neglected in membrane fouling studies.[6,32] Previous studies identified that the cake resistance will undergo an immediate increase when the TMP increases as a result of cake compression. Conversely, a lower TMP will lead to the swelling of the cake layer, thus causing a drop in the cake resistance.[12−14,33] Such compressibility of the cake layer might also explain the TMP jump that occurs during long-term constant-flux operation.[12] As illustrated by the operational data (Figure 3c), the membrane permeability increased from 63.64 to 65.52 LMH bar^{-1} when the flux dropped from 13.32 to 11.55 LMH and the applied TMP dropped from 20.94 to 17.63 kPa. The permeability decreased from 65.25 to 63.18 LMH bar^{-1} when the flux increased from 11.56 to 13.69 LMH and the applied TMP increased from 17.72 to 21.67 kPa; similarly, the permeability further decreased from 62.61 to 60.89 LMH bar^{-1} when the flux increased from 14.75 to 16.26 LMH and the applied TMP increased from 23.56 to 26.70 kPa, confirming that the immediate change in permeability or membrane resistance is due to compression or swelling of the cake when increasing or decreasing the suction pressure (given that the water viscosity did not change drastically).

Fouling caused by the formation of a gel layer and associated pore blocking is the main contributor to irreversible fouling and requires chemical cleanings for mitigation.[6,8] A diluted NaOCl solution (1100−2000 ppm for "maintenance" cleaning and 3500−4000 ppm for "recovery" cleaning) was dosed approximately once every 1−2 weeks to oxidize the functional groups of organic and biological foulants with resultant release from the membrane. Either citric or oxalic acid (1100−2000 ppm) was used to remove inorganic foulants, thereby effectively restoring the membrane filtration performance. Permeability data between two consecutive ex situ cleanings were extracted and used in algorithm development (Figure 3a). Membrane permeability was improved to varying degrees because of differences in both the chemical dosage of each CEB and the operating duration between consecutive CEBs.

Model Performance. MBR-Net utilizes the underlying trained LSTM encoder−decoder to predict permeability changes of 180 filtration−relaxation cycles for the upcoming day through a 5 day (900 filtration−relaxation cycles) look-back horizon (Figure 2a and SI Figure S2g). The developed model is able to (1) predict an immediate change in permeability caused by cake compression and swelling (attributed to variations in suction pressure resulting from the required changes in flux). In one testing sample (Figure
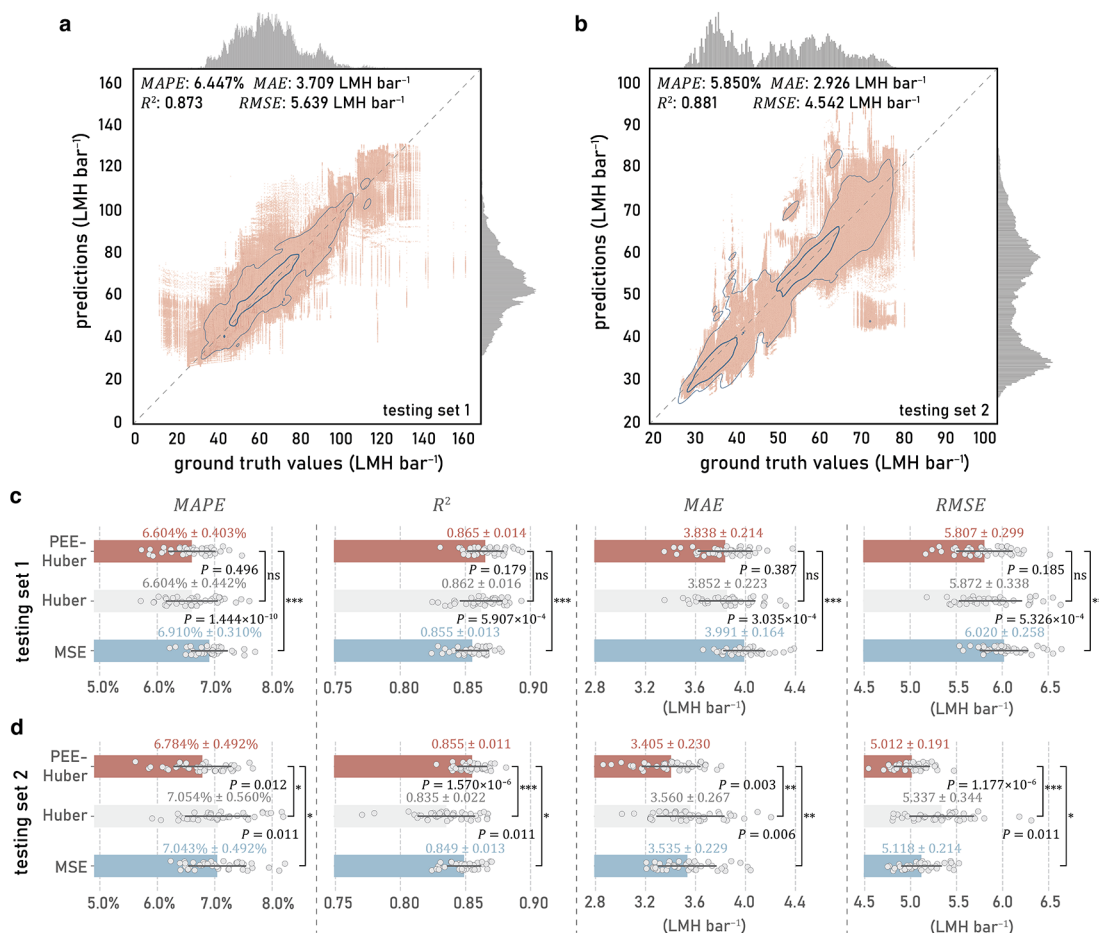
Figure 4. Model performance results (a, b). Model performance on testing set 1 and testing set 2. The horizontal coordinate and the vertical coordinate of apricot dots represent the ground truth permeability values and predictions. Histograms of observed and predicted permeability values are shown at the top and right side of the plot. Thick blue contour lines of the kernel density estimate plot enclose regions where the probability density is smaller than 0.33; thin blue contour lines enclose regions where the probability density is smaller than 0.66. The dashed diagonal line indicates the perfect prediction. (c, d) Average MAPE, R^2 MAE, and RMSE values of models trained using different loss functions (sample size equals 40 for each group of models with regard to different loss functions) on testing set 1 and testing set 2 with error bars indicating the standard deviation. Red bars: the group of models trained using the polynomial−exponential enhanced Huber loss function; gray bars: the group of models trained using the Huber loss function; blue bars: the group of models trained using the MSE loss function. Anderson−Darling, Shapiro−Wilk, and Kilmogorov−Smirnov tests were performed to confirm the data normality of each group, with the test results reported in SI Tables S6 and S7. One-tailed two-sample t-tests ($\alpha = 0.05$) were employed to test the statistical significance with t-values reported in SI Tables S8 and S9. $P \leq 0.001$ (***); $0.001 < P \leq 0.01$ (**); $0.01 < P \leq 0.05$ (*); $P > 0.05$ (ns).

3b,c), when TMP undergoes a sudden decrease from 20.94 to 17.63 kPa, the predicted permeability increases from 64.92 to 66.22 LMH bar^{-1} compared to the actual increase in the observed permeability from 63.64 to 65.52 LMH bar^{-1}. Conversely, when the TMP increases suddenly from 17.72 to 21.67 kPa and from 23.56 to 26.70 kPa, the predicted permeability decreases from 66.39 to 64.71 LMH bar^{-1} and from 62.90 to 61.68 LMH bar^{-1}, respectively, compared to the observed permeability decrease from 65.25 to 63.18 LMH bar^{-1} and from 62.61 to 60.89 LMH bar^{-1}. The model can also (2) predict the gradual permeability decline over a 1 day operation resulting from foulant accumulation, as evidenced by the decrease in the predicted permeability from 65.63 LMH bar^{-1} at the first filtration−relaxation cycle (observed permeability: 64.73 LMH bar^{-1}; flux: 13.71 LMH; TMP: 21.17 kPa) to 64.52 LMH bar^{-1} after 100 cycles (observed permeability: 63.34 LMH bar^{-1}; flux: 13.67 LMH; TMP: 21.59 kPa) and, further, to 59.64 LMH bar^{-1} after 180 cycles

(observed permeability: 60.35 LMH bar^{-1}; flux: 13.78 LMH; TMP: 22.82 kPa) (Figure 3b,c). Additionally, the trained model is capable of (3) predicting permeability recovery after the irreversible fouling layer was removed by CEB, as illustrated by the increase in the predicted permeability from 49.69 to 74.93 LMH bar^{-1}, compared to the increase in the observed permeability from 48.51 to 75.24 LMH bar^{-1} (Figure 3d,e). To generate real-time permeability predictions, encoder features are automatically fetched from the WWTP's IoT platform, while decoder features with respect to expected future operational and maintenance conditions require manual input from MBR operators (SI, Figure S8b and Supplementary Video S1). The necessity of an encoder−decoder-based model and its significantly better performance compared with other simple regression models are demonstrated in the SI Table S14.

To demonstrate the intertank robustness of the model, we evaluated the performance of the best-performing model using
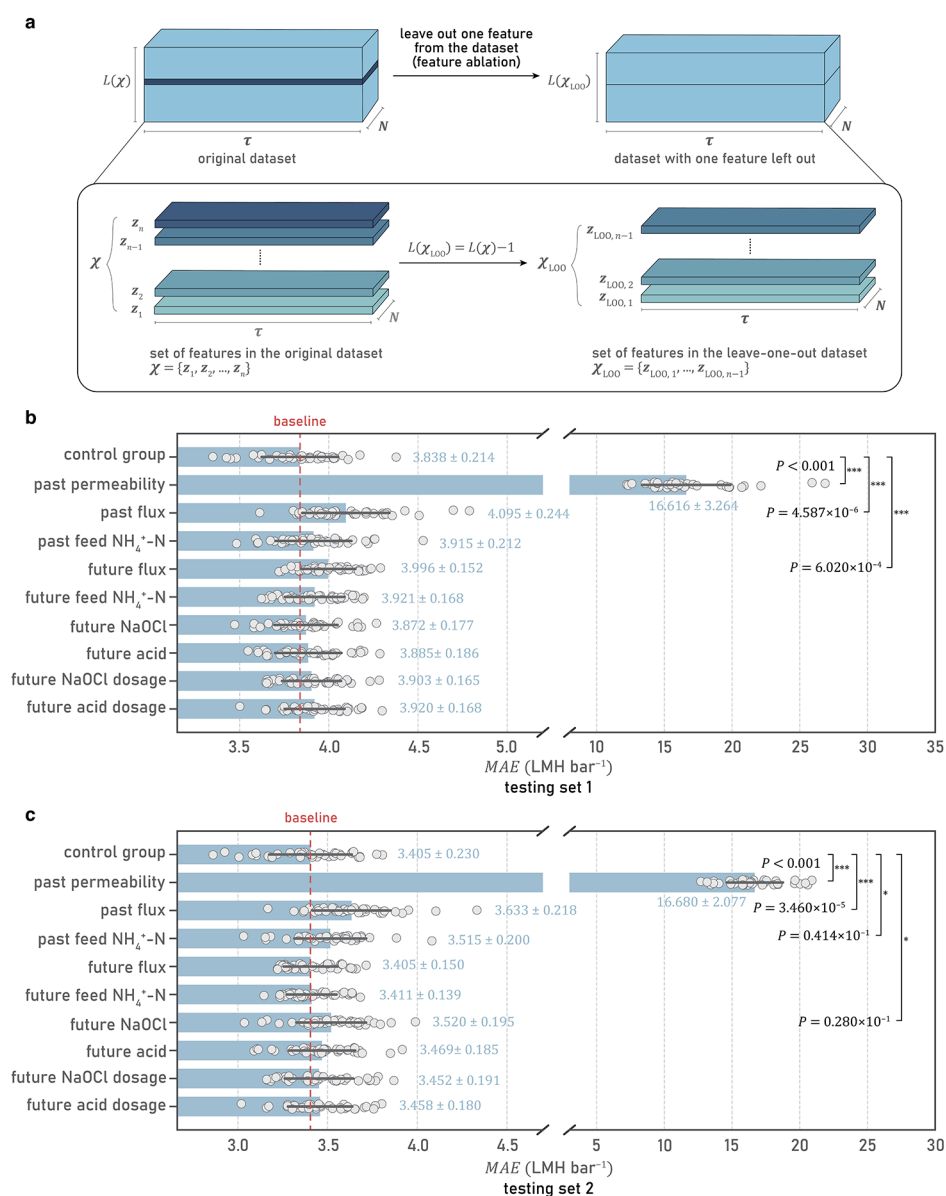
Figure 5. Leave-one-out ablation results. (a) An illustration of data set manipulation for the leave-one-out feature ablation experiment. (b, c) Average MAE values of models trained using different leave-one-out data sets on testing set 1 and testing set 2 (blue bars) with error bars indicating the standard deviation (sample size equals 40 for each group). Anderson−Darling, Shapiro−Wilk, and Kilmogorov−Smirnov tests were used to test the data normality of each group with test results reported in the SI Tables S10 and S11. Two-tailed Mann−Whitney tests ($\alpha = 0.05$) were employed to determine the statistical significance with U-values reported in SI Tables S12 and S13. $P \leq 0.001$ (***); $0.001 < P \leq 0.01$ (**); $0.01 < P \leq 0.05$ (*); $P > 0.05$ (ns). The average MAE value of the control group (that is the PEE-Huber group in performance evaluation) was set as the baseline.

a testing set (i.e., testing set 1 as specified in Materials and Methods, Figure 2c) consisting of data obtained from tank 1 of Plant A, which is operated and serviced under conditions that differ from those used on the other MBR tanks. We also evaluated the temporal robustness of MBR-Net on a testing set (i.e., testing set 2 as specified in Figure 2c) comprising data from all tanks of Plant A during a different time period compared with the training and validation sets. Overall, the developed model exhibits a relatively consistent performance on both testing sets, achieving a MAPE of 6.45%, an R^2 of 0.87, an MAE of 3.71 LMH bar^{-1}, and an RMSE of 5.64 LMH bar^{-1} on testing set 1 (Figures 4a and S5a,b) and a *MAPE* of 5.85%, an R^2 of 0.88, an MAE of 2.93 LMH bar^{-1}, and an RMSE of

4.54 LMH bar^{-1} on testing set 2 (Figures 4b and S5c,d). This two-pronged performance evaluation approach demonstrates both the exceptional adaptability of the model to various operational and maintenance conditions and its consistent performance across various stages of membrane fouling.

We also assessed the robustness of the model with regard to the choice of the loss function (eqs 2, 5, and S10). Hyperparameters were tuned individually for three groups of models trained using different loss functions (SI Table S1). On testing set 1 (Figure 4c and SI Table S8), models trained using a PEE-Huber loss function exhibit substantial improvements ($P < 0.001$) in all performance indicators compared with models trained using a mean squared error (MSE) loss
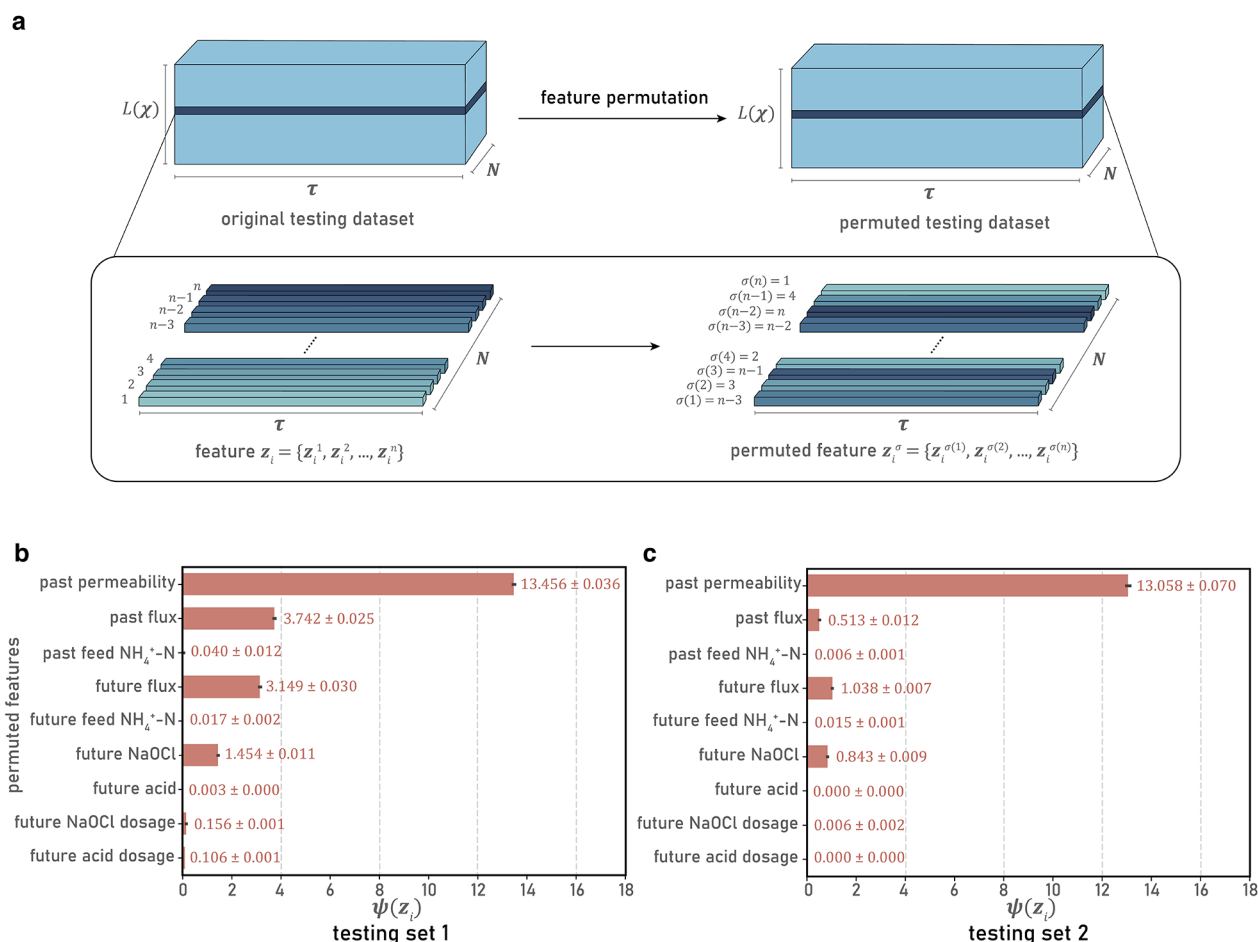
Figure 6. Feature permutation results. (a) An illustration of data set manipulation for the feature permutation experiment. (b, c) Permutation feature importance computed on testing set 1 and testing set 2. Red bars represent the average importance score of each feature with error bars indicating the standard deviation (sample size equals 5 for each model feature).

function. However, the PEE-Huber group shows a similar performance on testing set 1 compared to the conventional Huber group with no significant improvement ($P > 0.05$) in all performance indicators. We surmise that this difference in comparison results could be attributed to the fact that most of the permeability data points in testing set 1 are normally distributed within the range of 40−100 LMH bar^{-1} as are the majority of the permeability data points in the training and validation sets (SI Figure S3c). On the other hand, testing set 2 contains more permeability data points within the range of 30−40 LMH bar^{-1}—a range where fewer permeability data points were present in the training and validation sets. To improve the prediction accuracy within this range, higher penalties need to be applied to these rarer data points. As a result, the PEE-Huber group significantly outperforms the Huber group with no additional weights applied to the loss function on testing set 2. Marked improvements ($P < 0.05$) in all performance indicators are identified on testing set 2 when we compare the PEE-Huber groups with either the Huber group or the MSE group, indicating that the implementation of the weighted Huber loss function is effective in dealing with imbalanced data sets with outliers and leads to improvements in the overall prediction accuracy (Figure 4d and SI Table S9).

Feature Importance Analysis. We performed leave-one-out feature importance (also known as feature ablation) and

feature permutation experiments to quantify the contribution of each feature to the model performance (Figures 5a, 6a, and SI Supplementary Text S3). Based on the results of feature ablation (Figure 5b,c), past permeability is considered the most important feature, as indicated by the substantial increase in MAE values (performance drops) to 16.62 ± 3.26 and 16.68 ± 2.08 LMH bar^{-1} compared with the control group, which has MAE values of 3.84 ± 0.21 and 3.41 ± 0.23 LMH bar^{-1} on testing set 1 and testing set 2, respectively. The second most important feature suggested by the feature ablation experiment is the past flux. Leaving out past flux results in a significant increase in MAE values to 4.10 ± 0.24 and 3.63 ± 0.22 LMH bar^{-1}, with $P < 0.001$, on testing set 1 and testing set 2, respectively. Leaving out future flux, past feed NH$_4^+$-N and future NaOCl also lead to different levels of significant performance drops compared with the control group on either testing set 1 or testing set 2, with $P < 0.05$.

Similarly, the feature permutation results show that past permeability provides the most distinct performance difference, achieving importance scores $\psi(z_i)$ of 14.46 ± 0.04 and 13.06 ± 0.07 LMH bar^{-1} on testing set 1 and testing set 2, respectively (Figure 6b,c). Because the past permeability sequence reflects the current membrane aging and irreversible fouling conditions, incorrect membrane fouling information will be encoded and passed to the LSTM decoder with the resultant

I

generation of permeability predictions with a huge bias if the past permeability is permuted (SI Figure S6a). The permutation results also demonstrate that flux is crucial, with past flux achieving $\psi(z_i)$ values of 3.74 ± 0.03 and 0.51 ± 0.01 LMH bar^{-1} on testing sets 1 and 2 and future flux achieving $\psi(z_i)$ values of 3.15 ± 0.03 and 1.04 ± 0.01 LMH bar^{-1} on testing sets 1 and 2, respectively (Figure 6b,c). As discussed previously, change in flux (or productivity) requires the adjustment of suction pressure, eventually leading to a change in cake compressibility. A noticeable bias can be found on permeability predictions if past flux is permuted (SI Figure S6b), as incorrect relationships between past flux and past permeability can be learned by the model. Similarly, if future flux is permuted, the LSTM decoder will generate the permeability predictions based on the provided variance in flux, and immediate changes in permeability will be incorrectly forecast (SI Figure S6c). Future NaOCl is another important feature for achieving accurate permeability predictions, as indicated by feature permutation results with $\psi(z_i)$ values of 14.46 ± 0.04 and 13.06 ± 0.07 LMH bar^{-1} on testing sets 1 and 2, respectively (Figure 6b,c). Permeability recovery cannot be accurately predicted if future NaOCl is permuted (SI Figure S7a,b).

Model Limitations. The feature importance results also clearly indicate that the developed model is able to predict two expected permeability changes of real-world MBR plants: (i) immediate change in permeability (or short-term fouling behaviors of irreversible fouling) due to cake compression and swelling and (ii) permeability recovery (or amelioration of irreversible fouling) due to the implementation of CEB. However, model limitations are evident with regard to the low sensitivity of certain features and the absence of some potentially important features. It can be inferred that the current options of wastewater quality parameters may be inadequate for representing the impact of wastewater characteristics on membrane fouling. This is indicated by the similar MAE values achieved with respect to the control group on both testing sets when leaving out the future feed NH_4^+-N, and by the achieved $\psi(z_i) < 0.05$ LMH bar^{-1} for both past and future feed NH_4^+-N in feature permutation experiments. Biomass and sludge characteristics such as mixed liquor suspended solids (MLSS) concentration could be a suitable model feature as reported in previous studies on MBR membrane fouling.[34−36] However, since the WWTPs examined here exhibit a relatively constant MBR tank MLSS concentrations and lack accurate MLSS sensors, we have not yet incorporated any biomass characteristics into the model. Additionally, it can be anticipated that future acid dosage will exhibit low contributions to the model since only a few acid cleaning events were performed within the time span of testing set 1, and no acid cleaning was conducted on any MBR tanks within the time span of testing set 2. As a result of similar data limitations and/or missing data problems, several other potentially important parameters related to operational conditions, such as aeration intensity and temperature, were also omitted in the data processing step. These identified limitations present opportunities for future model enhancements, and in the future, we will further expand the MBR data set to include additional MBR process parameters. With ongoing refinement, we expect that the model will be capable of capturing more nonlinear multivariate relationships underlying the membrane fouling behavior.

Implications and Future Prospects. The main challenge in membrane fouling modeling is capturing the nonlinear relationships between various factors contributing to the observed complex fouling behavior. These nonlinear relationships are the primary reason why mechanistic models, developed using predetermined parameters from laboratory experiments, have limited value in optimizing decision-making and mitigating fouling in full-scale MBR plants. Due to the lack of reliable simulations tailored to real-world MBR fouling processes, considerable engineering skills and empirical knowledge are required by plant operators to successfully operate and service WWTPs, relying on subjective evaluations of the past and current MBR performance. Nevertheless, projections about future membrane fouling conditions under different operational conditions are often inaccurate. The ML model presented here provides a possible solution by gaining insights into past performance and relationships between operational conditions and permeability from large historical data sets of MBR plants, thereby enabling the forecasting of membrane fouling and, through this, supporting decision-making in the operation of WWTPs. Specifically, the developed model is focused on simulating the immediate change in irreversible fouling caused by variations in operational conditions, the expected accumulation of irreversible fouling over 1 day of operation, and the removal of irreversible fouling after the proposed implementation of CEB, achieving MAPE < 6.45%, MAE < 3.71 LMH bar^{-1}, $R^2 > 0.87$, and RMSE < 5.64 LMH bar^{-1} on two independent testing sets. The MBR-Net package described here may be integrated with the monitoring and control systems currently used at existing plants, thereby enabling the development of intelligent WWTPs.

"Less is more" encapsulates the esteemed idea that the harmonious synergy resulting from simplicity can often yield more robust impacts than excessive complexity. Synaptic pruning that occurs in the human brain is a typical example of this principle, enhancing the brain's neural transmission efficiency. In the development of frontend ML applications, this principle similarly motivates and has been widely applied in both feature reduction[37,38] and model architecture design.[39,40] In our work, we enhance the model efficiency by extracting critical MBR features prior to model training and applying dropouts during training while making the model structure more compact by adapting the LSTM encoder– decoder structure using insights from the TFT architecture. These approaches facilitate the ability of the ML model to resolve real-world membrane fouling tasks associated with the operation of MBR WWTPs. However, membrane fouling and fouling amelioration are recognized to be highly complicated processes. As a result, modeling of MBR membrane fouling and recovery may conform more to the dictum that "less is not always more" and require the inclusion of a range of diverse parameters. Due to the relatively similar feedwater characteristics and certain operational conditions of the MBR plants for which extensive data sets were available, we did not include static or other missing variables such as membrane tank MLSS, characteristics of physical cleaning (such as relaxation frequency and aeration intensity), or characteristics of the biomass in the present work. After sufficient data were obtained from more MBR plants, these MBR parameters could be incorporated into the workflow of model development and performance evaluation. By retraining the model with an expanded data set and adding more model complexity

accordingly, we believe that the effect of "synaptic blooming" could be achieved and a more generic membrane fouling model that is applicable to multiple MBR plants could be established. Finally, we envision MBR-Net as a primer for future ML studies on membrane fouling of full-scale MBR plants and, more generally, ML studies on performance deterioration and recovery in the long-term operation of WWTPs in general.

■ ASSOCIATED CONTENT

ⓈⒾ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.4c12835.

> Supplementary texts, including (i) nomenclature, (ii) development of environment description and code availability, (iii) MBR data collection, processing, splitting, and distillation, and (iv) model training, performance evaluation, and deployment; supplementary notes on (i) data processing algorithm, (ii) description of web APIs, and (iii) logic of web APIs; supplementary figures including MBR module and full-scale MBR plant, data denoising and time-series data sets, feature reduction and custom loss function, hyperparameter tuning, performance evaluation, model predictions, and web application architectural layers and user interface; supplementary tables including hyperparameter tuning results, normality test, pair-wise t-test, two-sample t-test, Mann−Whitney t-test, and performance of other simple regression models; supplementary video: Frontend web app and real-time membrane fouling prediction (PDF)

■ AUTHOR INFORMATION

Corresponding Author

T. David Waite − *UNSW Centre for Transformational Environmental Technologies (CTET), Yixing 214200, China; Water Research Centre, UNSW Sydney, Sydney, NSW 2052, Australia;* ⓞ orcid.org/0000-0002-5411-3233; Phone: +61-2-9385 5060; Email: d.waite@unsw.edu.au

Authors

Yunyi Zhu − *UNSW Centre for Transformational Environmental Technologies (CTET), Yixing 214200, China; Water Research Centre, UNSW Sydney, Sydney, NSW 2052, Australia;* ⓞ orcid.org/0000-0002-8634-4431

Yuan Wang − *UNSW Centre for Transformational Environmental Technologies (CTET), Yixing 214200, China; Water Research Centre, UNSW Sydney, Sydney, NSW 2052, Australia;* ⓞ orcid.org/0000-0003-1728-5762

Elisabeth Zhu − *UNSW Centre for Transformational Environmental Technologies (CTET), Yixing 214200, China; School of Computer Science and Engineering, UNSW Sydney, Sydney, NSW 2052, Australia*

Zeyu Ma − *Beijing Origin Water, Beijing 100000, China*

Hanchen Wang − *School of Computer Science and Engineering, UNSW Sydney, Sydney, NSW 2052, Australia*

Chunsheng Chen − *Beijing Origin Water, Beijing 100000, China*

Jing Guan − *Beijing Origin Water, Beijing 100000, China*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.est.4c12835

■ REFERENCES

(1) Judd, S. *The MBR Book: Principles and Applications of Membrane Bioreactors for Water and Wastewater Treatment*; Elsevier, 2010.

(2) Xiao, K.; Xu, Y.; Liang, S.; Lei, T.; Sun, J.; Wen, X.; Zhang, H.; Chen, C.; Huang, X. Engineering Application of Membrane Bioreactor for Wastewater Treatment in China: Current State and Future Prospect. *Front. Environ. Sci. Eng.* **2014**, *8* (6), 805−819.

(3) Hai, F. I.; Yamamoto, K.; Lee, C.-H. *Membrane Biological Reactors: Theory, Modeling, Design, Management and Applications to Wastewater Reuse*; IWA Publishing, 2018.

(4) Xiao, K.; Liang, S.; Wang, X.; Chen, C.; Huang, X. Current State and Challenges of Full-Scale Membrane Bioreactor Applications: A Critical Review. *Bioresour. Technol.* **2019**, *271*, 473−481.

(5) Vaccari, M.; Abbà, A.; Bertanza, G.; Collivignarelli, C. An Evidence-Based Survey on Full-Scale Membrane Biological Reactors: Main Technical Features and Operational Aspects. *Appl. Sci.* **2022**, *12* (13), 6559.

(6) Di Bella, G.; Di Trapani, D. A Brief Review on the Resistance-in-Series Model in Membrane Bioreactors (MBRs). *Membranes* **2019**, *9* (2), 24.

(7) Cai, C.; Sun, W.; He, S.; Zhang, Y.; Wang, X. Ceramic Membrane Fouling Mechanisms and Control for Water Treatment. *Front. Environ. Sci. Eng.* **2023**, *17* (10), 126.

(8) Wang, Z.; Ma, J.; Tang, C. Y.; Kimura, K.; Wang, Q.; Han, X. Membrane Cleaning in Membrane Bioreactors: A Review. *J. Membr. Sci.* **2014**, *468*, 276−307.

(9) Drews, A. Membrane Fouling in Membrane Bioreactors—Characterisation, Contradictions. *Cause and Cures. J. Memb. Sci.* **2010**, *363* (1), 1−28.

(10) Kraume, M.; Wedi, D.; Schaller, J.; Iversen, V.; Drews, A. Fouling in MBR: What Use Are Lab Investigations for Full Scale Operation? *Desalination* **2009**, *236* (1), 94−103.

(11) Wu, J.; Le-Clech, P.; Stuetz, R. M.; Fane, A. G.; Chen, V. Effects of Relaxation and Backwashing Conditions on Fouling in Membrane Bioreactor. *J. Membr. Sci.* **2008**, *324* (1), 26−32.

(12) Poorasgari, E.; Bugge, T. V.; Christensen, M. L.; Jørgensen, M. K. Compressibility of Fouling Layers in Membrane Bioreactors. *J. Membr. Sci.* **2015**, *475*, 65−70.

(13) Bugge, T. V.; Jørgensen, M. K.; Christensen, M. L.; Keiding, K. Modeling Cake Buildup under TMP-Step Filtration in a Membrane Bioreactor: Cake Compressibility Is Significant. *Water Res.* **2012**, *46* (14), 4330−4338.

(14) Jørgensen, M. K.; Keiding, K.; Christensen, M. L. On the Reversibility of Cake Buildup and Compression in a Membrane Bioreactor. *J. Membr. Sci.* **2014**, *455*, 152−161.

(15) Li, X.; Wang, X. Modelling of Membrane Fouling in a Submerged Membrane Bioreactor. *J. Membr. Sci.* **2006**, *278* (1), 151−161.

(16) Viet, N. D.; Jang, D.; Yoon, Y.; Jang, A. Enhancement of Membrane System Performance Using Artificial Intelligence Technologies for Sustainable Water and Wastewater Treatment: A Critical Review. *Crit. Rev. Environ. Sci. Technol.* **2022**, *52* (20), 3689−3719.

(17) Philippe, N.; Stricker, A.-E.; Racault, Y.; Husson, A.; Sperandio, M.; Vanrolleghem, P. Modelling the Long-Term Evolution of Permeability in a Full-Scale MBR: Statistical Approaches. *Desalination* **2013**, *325*, 7−15.

(18) Ludwig, T.; Gaida, D.; Keysers, C.; Pinnekamp, J.; Bongards, M.; Kern, P.; Wolf, C.; Sousa Brito, A. L. An Advanced Simulation

Model for Membrane Bioreactors: Development, Calibration and Validation. *Water Sci. Technol.* **2012**, *66* (7), 1384−1391.

(19) Dalmau, M.; Atanasova, N.; Gabarrón, S.; Rodriguez-Roda, I.; Comas, J. Comparison of a Deterministic and a Data Driven Model to Describe MBR Fouling. *Chem. Eng. J.* **2015**, *260*, 300−308.

(20) Wintgens, T.; Rosen, J.; Melin, T.; Brepols, C.; Drensla, K.; Engelhardt, N. Modelling of a Membrane Bioreactor System for Municipal Wastewater Treatment. *J. Membr. Sci.* **2003**, *216* (1), 55−65.

(21) Kovacs, D. J.; Li, Z.; Baetz, B. W.; Hong, Y.; Donnaz, S.; Zhao, X.; Zhou, P.; Ding, H.; Dong, Q. Membrane Fouling Prediction and Uncertainty Analysis Using Machine Learning: A Wastewater Treatment Plant Case Study. *J. Membr. Sci.* **2022**, *660*, No. 120817.

(22) Mirbagheri, S. A.; Bagheri, M.; Bagheri, Z.; Kamarkhani, A. M. Evaluation and Prediction of Membrane Fouling in a Submerged Membrane Bioreactor with Simultaneous Upward and Downward Aeration Using Artificial Neural Network-Genetic Algorithm. *Process Saf. Environ. Prot.* **2015**, *96*, 111−124.

(23) Jiang, T.; Zhang, H.; Gao, D.; Dong, F.; Gao, J.; Yang, F. Fouling Characteristics of a Novel Rotating Tubular Membrane Bioreactor. *Chem. Eng. Process. Process Intensif.* **2012**, *62*, 39−46.

(24) Niu, C.; Li, X.; Dai, R.; Wang, Z. Artificial Intelligence-Incorporated Membrane Fouling Prediction for Membrane-Based Processes in the Past 20 Years: A Critical Review. *Water Res.* **2022**, *216*, No. 118299.

(25) Nam, K.; Heo, S.; Rhee, G.; Kim, M.; Yoo, C. Dual-Objective Optimization for Energy-Saving and Fouling Mitigation in MBR Plants Using AI-Based Influent Prediction and an Integrated Biological-Physical Model. *J. Membr. Sci.* **2021**, *626*, No. 119208.

(26) Lai, Y.; Xiao, K.; He, Y.; Liu, X.; Tan, J.; Xue, W.; Zhang, A.; Huang, X. Machine Learning for Membrane Bioreactor Research: Principles, Methods, Applications, and a Tutorial. *Front. Environ. Sci. Eng.* **2025**, *19* (3), 34.

(27) Yang, M.; Zhu, J.-J.; McGaughey, A.; Zheng, S.; Priestley, R. D.; Ren, Z. J. Predicting Extraction Selectivity of Acetic Acid in Pervaporation by Machine Learning Models with Data Leakage Management. *Environ. Sci. Technol.* **2023**, *57* (14), 5934−5946.

(28) Lim, B.; Arık, S. Ö.; Loeff, N.; Pfister, T. Temporal Fusion Transformers for Interpretable Multi-Horizon Time Series Forecasting. *Int. J. Forecast.* **2021**, *37* (4), 1748−1764.

(29) Huber, P. J. *Robust Estimation of a Location Parameter*; Springer New York: New York, NY, 1992.

(30) Géron, A. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*; O'Reilly Media, Inc., 2022.

(31) Ebert-Uphoff, I.; Lagerquist, R.; Hilburn, K.; Lee, Y.; Haynes, K.; Stock, J.; Kumler, C.; Stewart, J. Q. CIRA Guide to Custom Loss Functions for Neural Networks in Environmental Sciences--Version 1. **2021**, *arXiv (Machine Learning)*. 10.48550/arXiv.2106.09757 (accessed February 2025).

(32) Diez, V.; Ezquerra, D.; Cabezas, J. L.; García, A.; Ramos, C. A Modified Method for Evaluation of Critical Flux, Fouling Rate and in Situ Determination of Resistance and Compressibility in MBR under Different Fouling Conditions. *J. Membr. Sci.* **2014**, *453*, 1−11.

(33) Robles, A.; Ruano, M. V.; Ribes, J.; Seco, A.; Ferrer, J. Mathematical Modelling of Filtration in Submerged Anaerobic MBRs (SAnMBRs): Long-Term Validation. *J. Membr. Sci.* **2013**, *446*, 303−309.

(34) Bagheri, M.; Mirbagheri, S. A. Critical Review of Fouling Mitigation Strategies in Membrane Bioreactors Treating Water and Wastewater. *Bioresour. Technol.* **2018**, *258*, 318−334.

(35) Bagheri, M.; Akbari, A.; Mirbagheri, S. A. Advanced Control of Membrane Fouling in Filtration Systems Using Artificial Intelligence and Machine Learning Techniques: A Critical Review. *Process Saf. Environ. Prot.* **2019**, *123*, 229−252.

(36) Le-Clech, P.; Jefferson, B.; Judd, S. J. Impact of Aeration, Solids Concentration and Membrane Characteristics on the Hydraulic Performance of a Membrane Bioreactor. *J. Membr. Sci.* **2003**, *218* (1), 117−129.

(37) Pudjihartono, N.; Fadason, T.; Kempa-Liehr, A. W.; O'Sullivan, J. M. A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. *Front. Bioinforma.* **2022**, *2*, No. 927312.

(38) Blum, A. L.; Langley, P. Selection of Relevant Features and Examples in Machine Learning. *Artif. Intell.* **1997**, *97* (1), 245−271.

(39) Hinton, G.; Vinyals, O.; Dean, J. Distilling the Knowledge in a Neural Network. **2015**, *arXiv (Machine Learning)*. 10.48550/arXiv.1503.02531 (accessed February 2025).

(40) Asher, M. J.; Croke, B. F. W.; Jakeman, A. J.; Peeters, L. J. M. A Review of Surrogate Models and Their Application to Groundwater Modeling. *Water Resour. Res.* **2015**, *51* (8), 5957−5973.