Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Experimental and data-driven PVDF hollow fiber direct contact membrane distillation module degradation prediction for water purification using a deep learning method

Amir Hossein Zabihi Tari

azabihit@umd.edu

Department of Mechanical Engineering, University of Maryland, College Park, United States of America
https://orcid.org/0000-0001-8444-6699

Amir H. Keshavarzzadeh

Department of Engineering, University of Cambridge, Cambridge, United Kingdom of Great Britain and Northern Ireland

Pouria Ahmadi

Department of Mechanical and Aerospace Engineering, University of Central Florida, Orlando, United States of America

Research Article

Keywords: Membrane distillation, PVDF degradation, Deep learning, LSTM, Membrane aging, Predictive maintenance, Water purification

Posted Date: September 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10618643/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Experimental and data-driven PVDF hollow fiber direct contact membrane distillation module degradation prediction for water purification using a deep learning method

Amir Hossein Zabihi Tari [a*], Amir H. Keshavarzzadeh [b], Pouria Ahmadi [c]

[a] A James Clark School of Engineering, University of Maryland, College Park, Maryland, United States of America
[b] Department of Engineering, University of Cambridge, Cambridge, United Kingdom of Great Britain and Northern Ireland
[c] Department of Mechanical and Aerospace Engineering, University of Central Florida, Orlando, Florida, United States of America
* Corresponding author: E-mail: azabihit@umd.edu

Abstract

Polyvinylidene Fluoride (PVDF) Direct Contact Membrane Distillation (DCMD) systems offer an efficient solution for clean water production, but are prone to degradation under variable operational conditions such as salinity, temperature, and pressure. This study investigates the aging behavior of PVDF DCMD membranes using experimental cyclic operation data and introduces a predictive model based on long short-term memory (LSTM) neural networks. The results reveal both reversible and irreversible degradation mechanisms contributing to performance decline. A new empirical formula is developed to quantify the reduction in membrane efficiency over time. Experimental observations show a 0.17% decline in normalized flux per cycle during early operation and approximately 4% irreversible loss following each chemical cleaning cycle. The LSTM model achieves high accuracy in predicting degradation trends under various operational scenarios, with mean absolute percentage errors of 0.63% and 0.57% for temperature and pressure variations, respectively. This approach enables proactive control strategies and optimized operating cycles, thereby extending membrane lifespan and ensuring stable performance in real-world applications. By combining empirical observations with deep learning, this work advances predictive maintenance in membrane distillation, offering a practical tool for managing PVDF membrane aging in water purification systems.

Keywords: Membrane distillation, PVDF degradation, Deep learning, LSTM, Membrane aging, Predictive maintenance, Water purification

Acknowledgements

The authors have no acknowledgments to declare.

1 Introduction

Freshwater resources are indispensable for various aspects of human life, including household, agricultural, and industrial purposes. However, the issue of water scarcity has grown increasingly severe, exacerbated by both global warming and rapid population expansion. Water scarcity has significant implications for public health worldwide, often resulting in tragic fatalities due to the lack of access to clean drinking water (de Medeiros Paulino 2019; Shahrim et al. 2023; Shannon et al. 2008). According to data from the World Health Organization (WHO), waterborne diarrheal diseases alone are responsible for more than 2 million deaths each year on a global scale (Ribeiro Pinela et al. 2024). Addressing this challenge necessitates the development of water treatment technologies that can recover a substantial amount of water in an energy-efficient fashion. Among various water treatment methods, the membrane distillation modules stand out as a promising solution. This is particularly pertinent because a substantial portion of the Earth's total water resources consists of saline water, rendering it unsuitable for direct consumption or use (Micale et al. 2009).

Membrane-based purification technologies offer a promising solution for producing clean water and treating wastewater from various industries. Among these, Membrane Distillation (MD) stands out for its ability to meet the growing demand for highly pure water. Using hydrophobic membranes, MD allows water vapor to pass through its porous structure and condense on the permeate side, producing highly pure water. Unlike Reverse Osmosis (RO), which struggles with high salt concentrations, MD can effectively desalinate saline feedwater and further concentrate RO reject solutions, facilitating better disposal (Kalla 2021; Schwantes et al. 2018).

Hollow fiber configuration is the most widely used geometry among membrane separation modules. An advantage of this configuration compared to other geometries is the significantly higher membrane active area per unit volume ratios, hence higher permeate flux (Nunes and Peinemann 2001). A significant limitation of MD is its long-term stability, as it is prone to wetting phenomena, where the saline liquid feed permeates the membrane pores (Zhang and Zhao 2023). Therefore, understanding degradation phenomena and rates can help optimize the operating conditions of the Direct Contact Membrane Distillation (DCMD) module to minimize the risk of early aging.

Polyvinylidene Fluoride (PVDF) is one of the most widely used substances for preparing hollow fiber membranes. PVDF is known for its good thermal stability, chemical resistance, and exceptional mechanical properties. However, the PVDF membrane's main challenge is its susceptibility to collides, organic substances, proteins, and microorganisms, which stems from its hydrophobic properties. This phenomenon is known as membrane fouling. Fouling is always accompanied by a decline in permeate flux and pore blockage. Sodium hypochlorite is typically used as a chemical cleaning agent to remove adsorbed foulants that cannot be effectively removed by hydraulic approaches (Levitsky et al. 2011). The main problem with sodium hypochlorite is that it can alter the chemical and physical properties of the PVDF membrane during continuous operation (Abdullah and Bérubé 2013; Delattre et al. 2012; Hajibabania et al. 2012; Hashim et al. 2011; Puspitasari et al. 2010; Regula et al. 2013), thereby reducing purification performance (Pessoa de Amorim and Afonso Ramos 2006; Kimura et al. 2006; Robinson et al. 2016). As a result, some manufacturers report cumulative lifetime exposure doses in the datasheets for their PVDF DCMD modules. At this threshold, the modules are considered aged, and they must be replaced with new DCMD membranes.

In response to the limitations of traditional theoretical models, the engineering community has increasingly turned to data-driven modeling approaches (de Medeiros Paulino 2019; Nagy et al. 2002; Dürrenmatt and Gujer 2012). Artificial Neural Networks (ANNs) have gained significant popularity as a tool for data-driven modeling. ANNs possess non-linear and non-parametric qualities, enabling them to approximate continuous functions with a high degree of accuracy (Zhang et al. 1998). However, it's worth noting that ANNs have their own set of limitations when it comes to predicting process performance, particularly in scenarios with inherent time-series dependencies. This limitation arises from the absence of a recurrent structure that would enable the utilization of previous outputs as current inputs, essentially overlooking the time-series characteristics present in historical data (Zhang and Man 1998). Consequently, the selection of an appropriate modeling approach that can effectively capture the time-dependent nature of process data becomes paramount.

In recent times, models employing Recurrent Neural Networks (RNNs), a category of deep neural networks, have emerged as valuable tools for time-series prediction tasks. RNNs are characterized by their structure, which incorporates feedback connections that allow them to utilize information from previous time steps to estimate current information. This recurrent structure makes RNNs particularly efficient in predicting intricate hydrologic time series. Among the RNN variants, the Long Short-Term Memory (LSTM) network stands out as an excellent choice for handling lengthy data sequences (Zaytar and El Amrani 2016; Zhang et al. 2018). LSTM networks are equipped with memory cells that differentiate the significance of data and reduce the amount of data that needs to be processed. When aiming for highly accurate long-term predictions in complex plant processes, developing a model using LSTM becomes essential. The specific type of LSTM utilized depends on structural features, such as the incorporation of an attention mechanism. The attention mechanism plays a crucial role in focusing more on essential input data while downplaying the influence of less critical data during weight calculations. This mechanism has found significant applications in domains like image classification (Feng et al. 2020; Ma et al. 2019; Shao et al. 2021) and neural machine translation (Zhang et al. 2020; Choi et al. 2018; Huang et al. 2016). Given its potential to

enhance prediction accuracy, recent studies have advocated for the use of attention-based recurrent networks in time-series prediction tasks. While attention mechanisms have been applied to fields like economics, including stock price changes (Qiu et al. 2020; Chen et al. 2020) and oil price forecasting (Shih et al. 2019; Song et al. 2020), they have seen limited application in the context of water treatment processes. Consequently, the adoption of attention-based LSTM models holds promise for process modeling, particularly in scenarios where long-term time-series data exhibit substantial variability and complexity in input and output variables.

There are very limited investigations on the characteristics of aging and its impact on the purification performance of PVDF DCMDs in the literature. In the current study, a lab-scale experiment is conducted by performing cyclic tests on an intact membrane. We use activated sludge as a model foulant solution. There are five parameters for quantifying the cleaning performance of the PVDF DCMD: i) purification capacity, which indicates the possible maximum filtered volume before cleaning is required (Fane 2007); ii) chemically irreversible loss in permeability, iii) chemically reversible loss in permeability, iv) physically irreversible loss in permeability (Porcelli and Judd 2010), and v) the rate of recovery in the time of chemical cleaning process, which indicates the effect of aging on the performance of the chemical cleaning process (Abdullah and Bérubé 2018).

In the present study, a new measure is introduced for taking into consideration the effect of temperature on aging and calculating the degradation of the cleaning performance over time for a specific number of cycles. Most of the previous studies focused on prognostic approaches that consider the loss in permeability or chemical cleaning as an indicator of the PVDF DCMD performance. Therefore, it is reasonable to see the aging prediction as a time series problem. While previous studies have optimized energy systems using various modeling techniques (Tari et al.

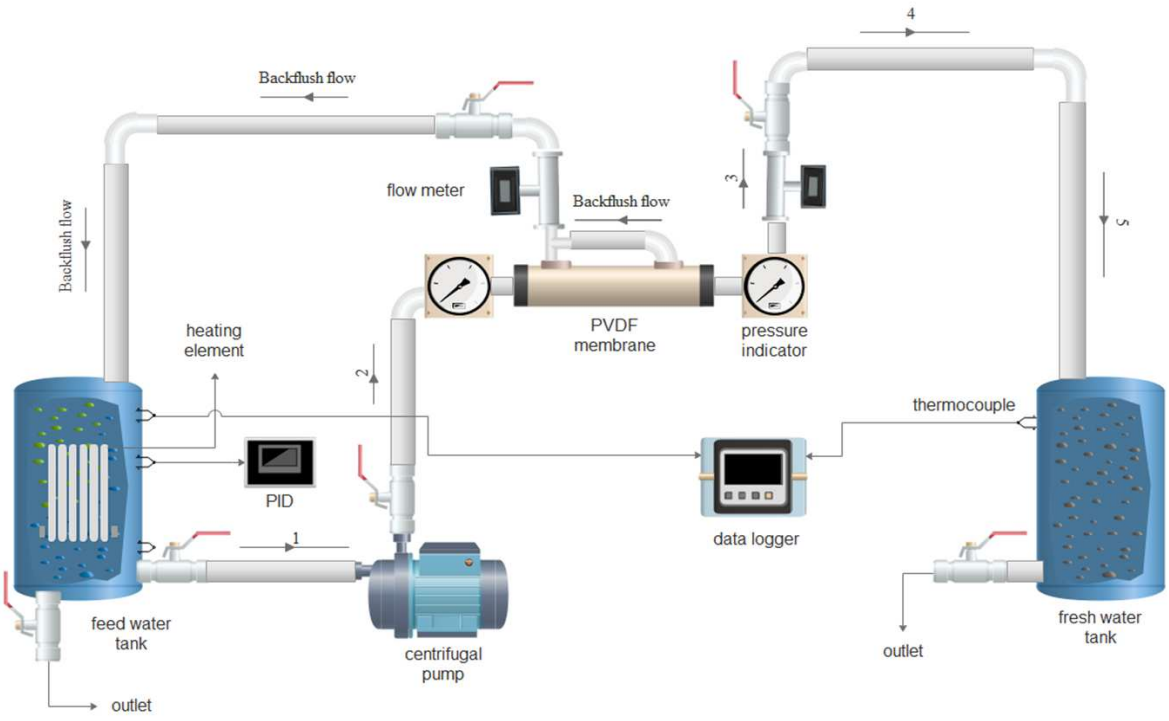
2023), our research introduces a novel data-driven deep-learning model for PVDF DCMD based on the LSTM structure, aiming to predict the aging rate under different operational conditions. Machine learning methods are widely used for forecasting time series, and they have been shown to be a very powerful tool for analyzing complex non-linear problems, such as drinking water treatment processes (Li et al. 2021). The most utilized architecture for time series analysis is RNN. The present study aims to employ this structure to predict the aging rate of PVDF DCMD under different concentrations of cleaning agents and operation conditions.

A novel data-driven deep-learning model for PVDF DCMD is introduced based on the LSTM structure. The developed model is then confirmed with experimental data under different feed concentrations and operation conditions.

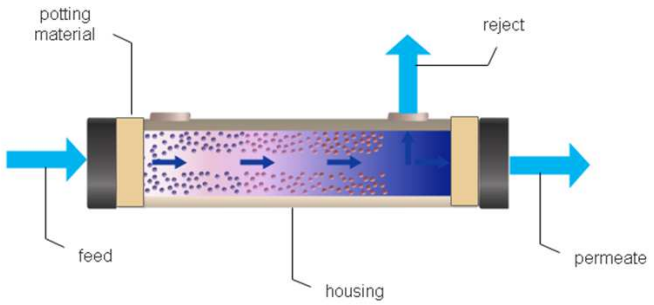
2 Experimental methods

Two reservoirs of inner diameter 64 cm and height 1 m, one with saline, untreated water (stream #1) and the other with fresh, purified water (stream #6) constitute the inlet supply and the output delivery tanks of the PVDF DCMD membrane, respectively. A centrifugal pump supplies saline water to the PVDF DCMD module feed inlet (stream #2), the permeated stream of desalinated water is sent to the freshwater reservoir (stream #6), and the membrane retentate outlet is recycled back into the feed water tank (stream #3). The untreated water tank temperature is controlled via a 400 W sealed water heater connected to a proportional integral derivative (PID) controller and measured using K-type thermocouples. Two thermocouples also measure the outlet stream temperature of the freshwater tank (stream #6) and the permeate outlet stream temperature of the PVDF DCMD module (stream #4). Additionally, two flow meters are utilized to monitor flow rates in the outlet (stream #6) and backpressure pipelines (stream #5). Two barometers are used

for measuring the imposed pressure and to calculate the pressure drop across the PVDF DCMD module. Finally, to measure the salinity of water, a salinity meter is employed. The schematic and the experimental setup can be seen in Fig. 1. The model of each measurement instrument is listed in Table 1, while the technical specifications of the PVDF DCMD module are listed in Table 2.



(a)



(b)



Fig. 1. (a) 3D schematic of the PVDF DCMD system setup. (b) Schematic of the PVDF DCMD module. (c) Experimental setup

Table 1. Equipment used for testing (pump, flowmeter, barometer)

Part	Model
Data logger	Prova 800
Heater element	400 W
PID controller	REX C900FK02-V*AN
Solid state relay	SSR-40 DA
Salinity meter	SA287
Thermocouple	K
Pump	Centrifugal Type: CM-100
Flowmeter	BNYZWOT inline flow meter 0.5-5GPM

Part	Model
Pressure gauge	DuraChoice 2 - Type: EN 837-1

Table 2. PVDF DCMD module characteristics

Parameter	Specifications
Model	DCMD module
Membrane material	PVDF (Polyvinylidene Fluoride)
Pore size	0.45 μm
Porosity	0.7
Membrane thickness	200 μm
Channel depth	5 mm
Width	0.5 m
Length	1.3 m

We performed 620 accelerated stress cycles (ASC) to estimate the degradation rate of the PVDF module by independently controlling the feed water flow rate (stream #2) according to a fluctuating profile between 0 and 18 lit/min. Each cycle starts with the activated sludge tank (stream #1) at different temperatures between 25°C and 40°C, set by a PID controller, while the permeate outlet stream (stream #4) is always maintained between 20-40°C. The activated sludge is purified at different permeate stream flow rates (stream #4), from zero up to a maximum value of 128 liters/hour based on the testing cycle. The ASC test is terminated when the transmembrane pressure reaches 70 kPa, at which point the PVDF module needs chemical cleaning. Following the cleaning procedure, the irreversible efficiency reduction is calculated

using clean water permeation tests. The cleaning procedure and the clean water permeation tests can be found in Abdullah and Bérubé (2018).

The cycles are divided into batches to better understand the degradation behavior over time. Each batch consists of a set number of cycles, typically 60, and represents a distinct phase of operation from the start of one peak in pressure to the next. This division into batches allows for a more detailed analysis of the degradation process, making it easier to identify trends and patterns in the performance of the DCMD module. Additionally, it helps in isolating the impact of various operational parameters on the module's performance, thereby providing more granular insights into the degradation mechanisms at play. Each batch is designed to capture the progressive aging and recovery effects, enabling us to track the reversible and irreversible changes in membrane performance more accurately.

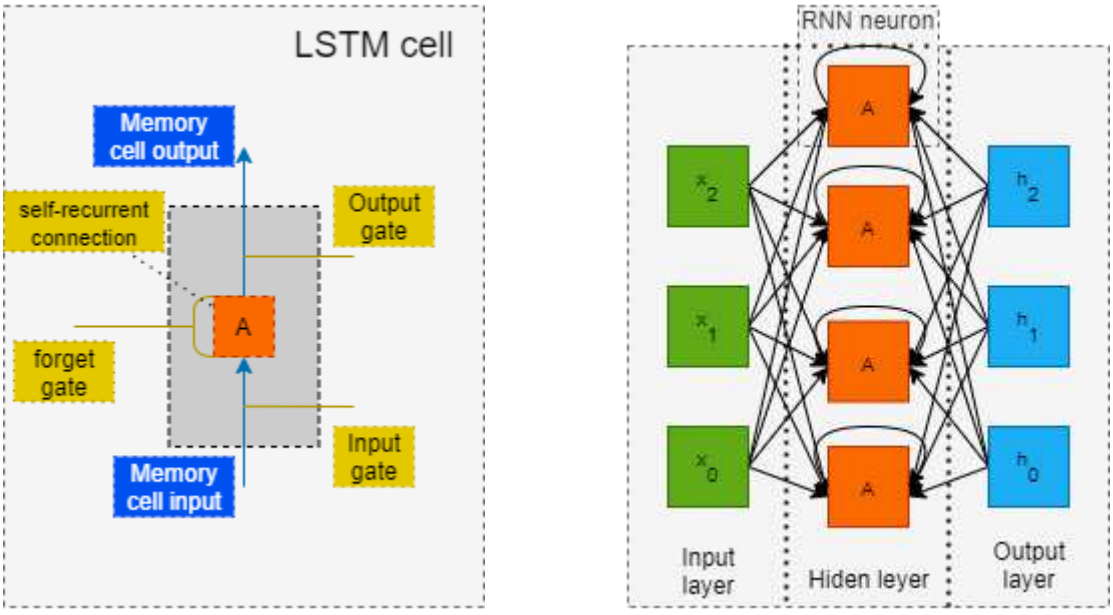
Following the cleaning steps, the membrane only partially recovers its original performance. The loss in performance that cannot be restored is attributed to irreversible degradation processes, such as chemical interactions and physical damage that occur during the unit's operation. This degradation affects the membrane's aging rate, which can be evaluated by measuring the difference in performance before and after cleaning. This approach has been previously used to determine the relationship between membrane aging and cleaning efficiency (Abdullah and Bérubé 2018). In these experiments, sodium hypochlorite (NaOCl) is used as the chemical cleaning agent.

3 Numerical methods

Feedforward networks are only able to remember the formative moments of training (Gers et al. 2002). There are three layers like the common feedforward network in a traditional RNN (Fig. 2a): hidden, input, and output. Nevertheless, RNN and feedforward networks are different in terms of the self-recurrent connections of the neurons.

3.1 Machine learning method

For the sake of elaboration, the inputs of RNN neurons are used to indicate the present time step, while the effects of previous time steps are noted.



(a) (b)

Fig. 2. An example of RNN and LSTM memory cell. (a) High-level description of each LSTM neuron consisting of input, forget, and output gate. The main idea of a self-recurrent connection is to connect the output signal to the input layer. Hence, with this architecture, output response is considered as an additional input variable. (b) The general structure of deep learning algorithms consists of input and output layers with several hidden layers in between

RNN includes "memory" to enhance the performance, however, problems of vanishing or discovering gradients have appeared as the main obstacles for using it in forecasting long time series (Gers et al. 2002). Within the RNN gradient back-propagation stage, the gradient signal ends up multiplying by the weight matrix many times related to the connections between the recurrent hidden layer's neurons. In long-term

time dependencies, the eigenvalue of the weight matrix is not able to learn the data because the gradient signal is minor.

This is called gradient vanishing. However, exploding gradients will occur when the eigenvalue is greater than one. This is because the gradient signal is very high. Using backpropagation through time (BPTT), the formula of the error backpropagation can be specifically stated as follows (Werbos 1990):

$$\delta_k{}^T = \delta_t{}^T \prod_{i=k}^{t-1} diag\,[f'(\boldsymbol{net_i})]\,\boldsymbol{W} \quad (1)$$

in which δ represents the error for each neuron, $\boldsymbol{W}$ shows the self-recurrent connection's weight matrix, and $\boldsymbol{net_i}$ is the neuron's weighted input. Assessing every term numerical value in $\delta_t{}^T$, the upper bound norm can be obtained as:

$$\left\|\delta_k{}^T\right\| \le \left\|\delta_t{}^T\right\| \prod_{i=k}^{t-1} \|diag\,[f'(\boldsymbol{net_i})]\|\,\|\boldsymbol{W}\| \le \left\|\delta_t{}^T\right\| \beta_f \beta_W{}^{t-k} \quad (2)$$

Where β_W and β_f denote the norms of the upper bound for the weight matrix $\boldsymbol{W}$ and diagonal matrix *diag* $[f'(\boldsymbol{net_i})]$, respectively.

As per condition (2), the exponential function of $\beta_f\ \beta_W$ passes the figure of error from time t to time k. This value is considered as upper bound. Thus, for the exceptionally large $t - k$ (blunder been passed for a more extended period), the value of (2) turns out to be either very large ($\beta_f\ \beta_W > 1$) or incredibly small ($\beta_f\ \beta_W < 1$) compared to the pre-expressed issue of inclination vanishing/detonating, individually.

Such issues are the principal inspiration for the LSTM model presenting a clever construction known as an LSTM cell (Fig. 2b). Despite the self-intermittent connection neuron, a forget gate, output gate, and input gate are included in comparison with the customary RNN network. The LSTM-based RNN can be developed by replacing RNN neurons in the hidden layer with the memory cell of LSTM.

By designing the self-repetitive connections in LSTM cells, the connections between the actual cells and their current circumstance can be understood. Changing the self-intermittent neurons' weight to 1.0, an LSTM cell's state can stay steady. In view of the preparation prerequisites, the repetitive connection can be impacted by the forget gate through the ability of the LSTM cell which can neglect or recollect the past advance state. On the other hand, the current situation can be varied by feeding the signal data by using the input gate. Consequently, the output gate can settle on the choice of whether the impact to different neurons is needed, by separating the output data.

Fig. 3 clarifies the LSTM cell in a numerical mathematical way. The necessary training data within a long period can be saved by the level stream pathway at the base by employing the organization cell state. At each time step, the cell state ought to be refreshed, and X_t addresses the current time step's input, h_t and h_{t-1} are the result of the past and present time steps, separately, C_t and C_{t-1} illustrate the present and previous time step's memory.

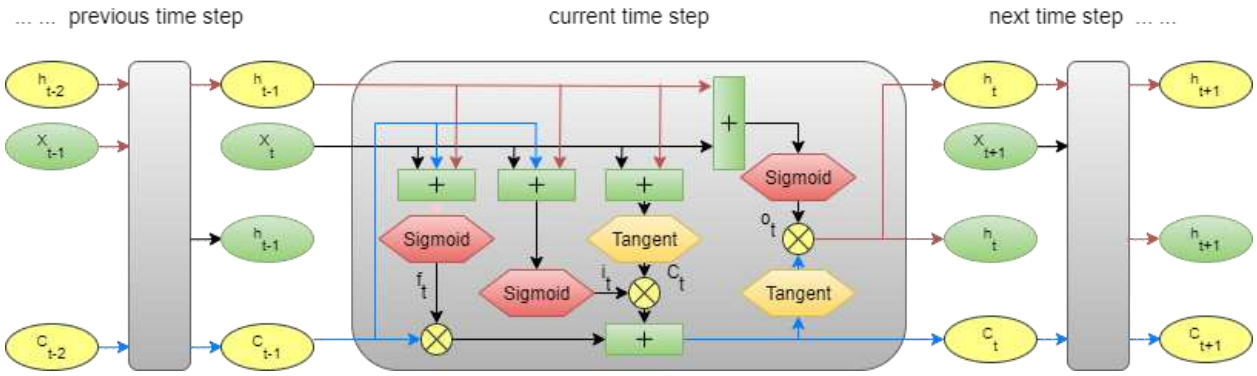


Fig. 3. The LSTM memory cell's architecture

The details of mathematical calculations related to LSTM are given in the supplementary section. First, the gate can be mathematically explained as:

$$g(x) = \sigma(\boldsymbol{W}x + b) \tag{3}$$

Then, for the present time step, the input gate layer numerical value can be achieved as:

$$i_t = \sigma[\boldsymbol{W}_i.(X_t, h_{t-1}, C_{t-1}) + b_i] \quad (4)$$

In the meantime, the candidate value $\widetilde{C_t}$ is made and added as:

$$\widetilde{C_t} = \tanh[\boldsymbol{W}_C.(X_t, h_{t-1}) + b_C] \quad (5)$$

An update is created by the above-mentioned i_t and $\widetilde{C_t}$ to the cell state. The forget gate layer is calculated as:

$$f_t = \sigma[\boldsymbol{W}_f.(X_t, h_{t-1}, C_{t-1}) + b_f] \quad (6)$$

The updated cell state is stated as:

$$C_t = f_t^{\circ} C_{t-1} + i_t^{\circ} \widetilde{C_t} \quad (7)$$

Ultimately, the output gate layer chooses to obtain the part of the state in terms of cell state as:

$$\begin{cases} o_t = \sigma[\boldsymbol{W}_o.(X_t, h_{t-1}) + b_o] \\ h_t = o_t^{\circ} \tanh(C_t) \end{cases} \quad (8)$$

where $tanh\ (z)$ shows the output function, $\sigma(z)$ represents the activation function, $\boldsymbol{W}_f$, $\boldsymbol{W}_i$, $\boldsymbol{W}_o$ and $\boldsymbol{W}_C$ are the weight matrices, b_f , b_i, b_o and b_C are bias vectors, ° indicates the multiplication by the elements (for more details, please refer to Appendix A).

The prediction model can be evaluated using criteria such as Mean Absolute Error Percentage (MAPE) and coefficient of determination (R2). MAPE is a measure of the prediction accuracy of a forecasting method in statistics. It usually expresses the accuracy as a ratio defined by equation 9. The lower the MAPE value, the more accurate the prediction. On the other hand, the higher the value of R2, the better the prediction, and if R2 = 1, the prediction will be completely consistent with the experimental data. The equations of these criteria are as follows:

$$MAPE = \frac{1}{N}\sum_1^N \frac{|X(t) - \hat{X}(t)|}{|X(t)|} \quad (9)$$

$$R^2 = 1 - \frac{\sum_1^N (X(t) - \hat{X}(t))^2}{\sum_1^N (X(t) - \bar{X}(t))^2} \quad (10)$$

In the above equations, X(t) represents the amount of actual water salinity measured, (t) represents the amount of water salinity predicted by the degradation model, (t) represents the average amount of water salinity data, and N also represents the number of values of salinity measured (Ma et al. 2018).

A cyclic operating condition is required because the detection of performance degradation under steady-state operation would not be possible via deep learning algorithms. Therefore, the prediction of aging would be inaccurate for different working conditions. Designing such cycles assists in augmenting the input data. In addition, when following such a method, the effect of each operation condition can be captured separately from other operation conditions. Therefore, these regimes of operation can lead to a broad landscape of data by providing the opportunity for the deep learning structure to fully capture the characteristics of each operation condition and map it to a respective aging rate. To elaborate, after training the LSTM, end users can specify their operation condition (inlet temperature, inlet pressure, and inlet flow rate) and obtain an accurate estimate for the aging rate. However, if the experiments were completed under constant operation conditions, the obtained estimated aging rate would only be valid for that specific operation condition and could not provide acceptable results for other operation conditions.

4 Results and Discussion

We adjust the cyclic pressure, the inlet flow rate, and temperature to investigate the performance of the DCMD module under different working conditions and aging phenomena in the PVDF DCMD unit. As illustrated in Fig. 4, the inlet pressure for each cycle varies from 0 to 2.3 bar (gauge pressure). The inlet flow rate reaches a peak of 18 lpm in each cycle. The data were sampled at intervals of 2 minutes. The inlet flow rate increases in the DCMD unit, after the inlet pressure

setpoint is imposed. The cyclic inlet pressure and flow rate are produced artificially and in each cycle, the value of this step function is increased until the upper limit of pressure, which is a sign of clogging, is reached.

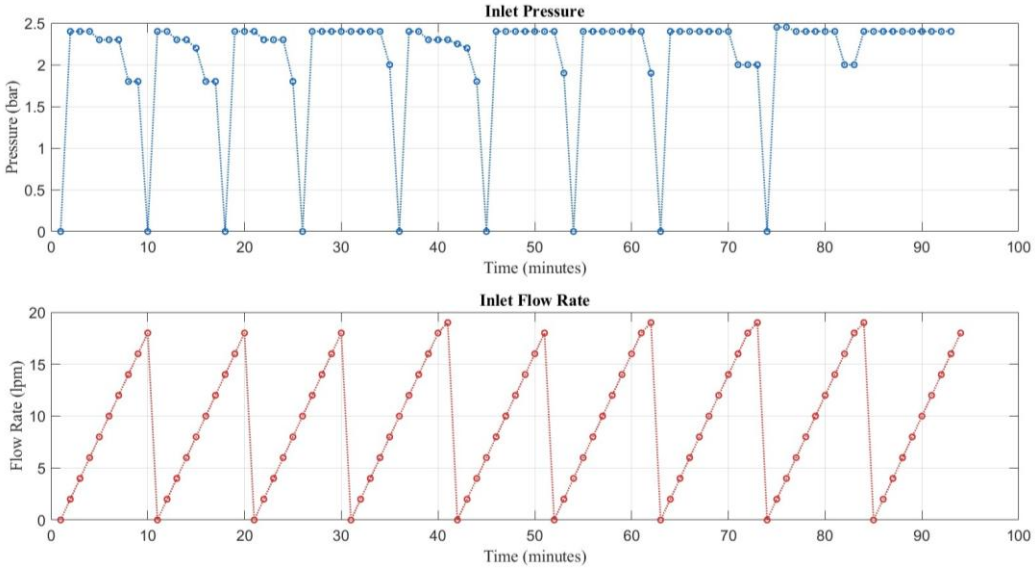
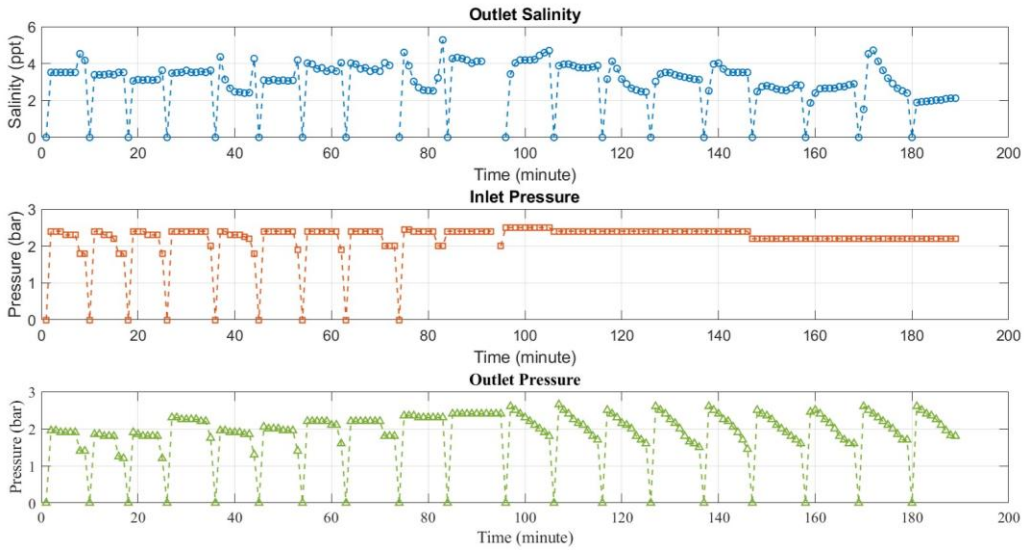


Fig. 4. Experimental cycles for inlet flow rate and pressure. This figure represents part of the designed experimental cycles that are imposed on the PVDF DCMD unit. These cycles, with different inlet pressures and flow rates, are produced to investigate the landscape of the PVDF DCMD module aging rate. By imposing such experimental regimes, the deep learning architecture could capture the effect of all influential parameters on aging.

Fig. 5 illustrates a sample of the obtained data for inlet and outlet parameters in a specific batch of cycles. Fig. 5 illustrates the last 96 minutes of the fifth batch and the first 91 minutes of the sixth batch of cycles. Before conducting the tests, the tests were performed in such a way that every 60 cycles were one batch. Each batch is such that it is from one peak to the next peak. Also, each batch is such that the rate of aging and dispersion of data increased from the first batch to the ninth.

The outlet pressure increases moderately over time, mainly due to the increase in salt concentration during the operation. After these cycles, the DCMD module reached a state that required chemical cleaning, which will be discussed in the following sections. After cleaning, the sixth batch started.

297 At the beginning of the sixth batch, the outlet pressure reached a peak of 2.7 bar. However, this
298 pressure continued increasing until the end of the batch, which led to the clogging of the module
299 again. As with the fifth batch, at the end of the sixth batch, backflush signals were induced to keep
300 the module working and gather as much data as possible. At this point, chemical cleaning of the
301 module was required again. The experiment continued with the same pattern for other batches.
302 The rationale for designing the cycles of experimentation following this method is described next.



303

304 Fig 5. Relation between inlet pressure and outlet salinity (quality of the purified water) and pressure in the fifth and
305 sixth batches. This figure represents the general shape of experiments performed on the module. It shows the last
306 cycles in the fifth batch before clogging occurred, and the first cycles in the sixth batch.

307 Fig. 6 shows samples of output data that depict the effect of operation conditions on permeate
308 flow, consisting of two main sections: the last 96 minutes of the fifth batch before chemical
309 cleaning and the first 91 minutes of the sixth batch of cycles. Therefore, Fig. 6 represents a few
310 cycles before and after cleaning and not the entire experimental cycle of each batch.

311 The operational parameters are induced as either different pulses or step functions. Fig 6-a depicts
312 the relation between backflush flow rate and the quantity of purified water, showing the effect of
313 backflush pulses while inlet pressure is kept constant.

314 As shown in Fig. 6-a, in the clogging period, the permeate flow experiences a sudden decrease in
315 value. After this sudden decrease, the permeate flow recovered to a lower value in comparison to
316 the permeate flow before clogging due to imposing backflush flow rate pulses; thereafter, the
317 permeate flow rate starts to gradually decrease until it reaches approximately 6 lpm. In this state,
318 the DCMD module is very sensitive to backflush pulses. At this point, clogging occurs and
319 chemical cleaning is required at the end of the fifth batch of cycles. The cyclic experiment is
320 continued after cleaning. After chemical cleaning, the cyclic operation of the module continues
321 without backflush. The backflush pulses are only induced when clogging is indicated, that is when
322 the pressure of the module experiences a sudden increase. After this point, the pulses in the
323 permeate flow are due to other cyclic operational parameters such as inlet pressure and flow rate.

324 Fig. 6-b shows the impact of the inlet pressure on the module's cleaning efficiency. The
325 performance of the DCMD module is much more sensitive to inlet pressure than the other
326 parameters. Prior to clogging, the pressure was kept constant; after chemical cleaning, the
327 backflush valve remained closed and pressure was induced by way of longer step functions. Due
328 to clogging before chemical cleaning, the pressure of the module would increase dramatically to
329 its critical upper limit (3 bar) if the backflush valve remained closed. By inducing backflush pulses,
330 the pressure of the unit was kept within the operating limit to gather as much data as possible in
331 the last minutes prior to clogging. After cleaning, pressure was induced using longer step functions
332 but with lower values. The value of the inlet pressure was increased gradually until it reached the
333 clogging state again.

Although the inlet temperature altered during the cycles, in general, the change in the permeate flow was not significant as the pattern of the inlet pressure remained the same. However, the inlet temperature changes showed a considerable effect on the aging rate. After cleaning, when the inlet pressure was induced through longer step functions and inlet flow pulses were also induced, the module showed a steep response without delay. It was shown that the higher the inlet flow rate, the higher the stability of the system, which in turn prevented the PVDF DCMD system from sudden cut-off. As expected, the results demonstrated that the cleaning efficiency decreased moderately due to degradation. The degradation rate of the PVDF DCMD module is discussed in the next subsection.

Fig. 6-c represents the impact of inlet temperature on the performance of the PVDF DCMD module. The experiments concerning the effect of inlet temperature on degradation consisted of two main patterns: cyclic step functions followed by pulses. As the operation continued, the outlet temperature remained higher on average than the inlet temperature; in other words, the temperature of the module increased during operation, which can lead to faster aging. The effect of this increase in temperature can be captured by the LSTM structure to estimate the effect of temperature on aging. Hence, temperature increases can be considered an additional parameter related to the aging and cleaning performance.

Through the increase in the number of cycles, the average difference between the inlet and outlet temperature increased. The highest observed temperature difference was 10 °C. Experiments were conducted at different inlet temperatures, ranging from room temperature to 45°C. Fig. 6-d represents the value of the outlet backflush temperature. As shown, the DCMD module experienced a delay related to the backflush temperature, however, the responses of the backflush temperature were smoother than those of the inlet temperature.

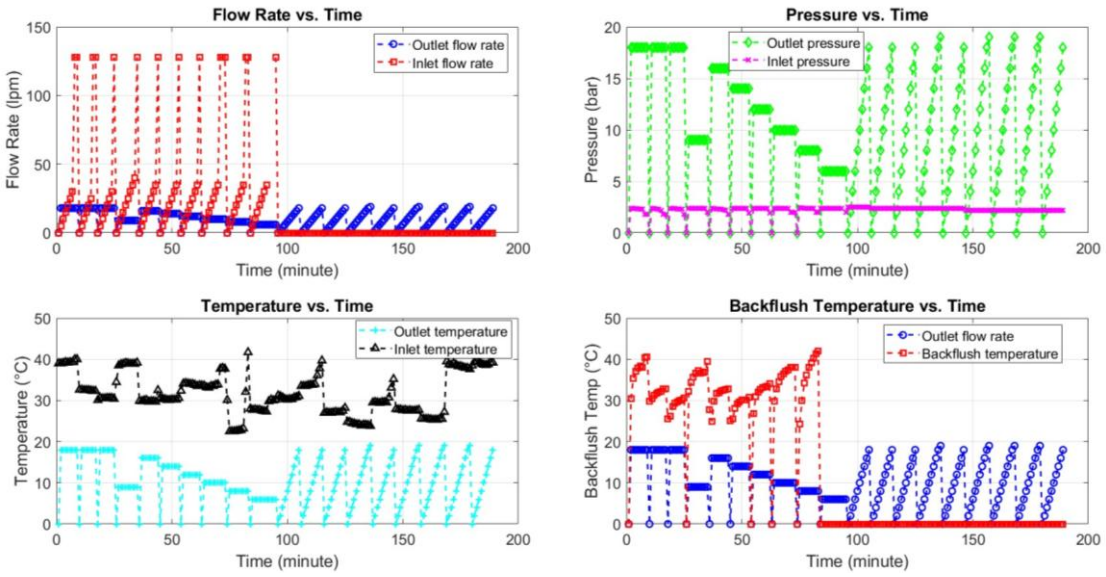
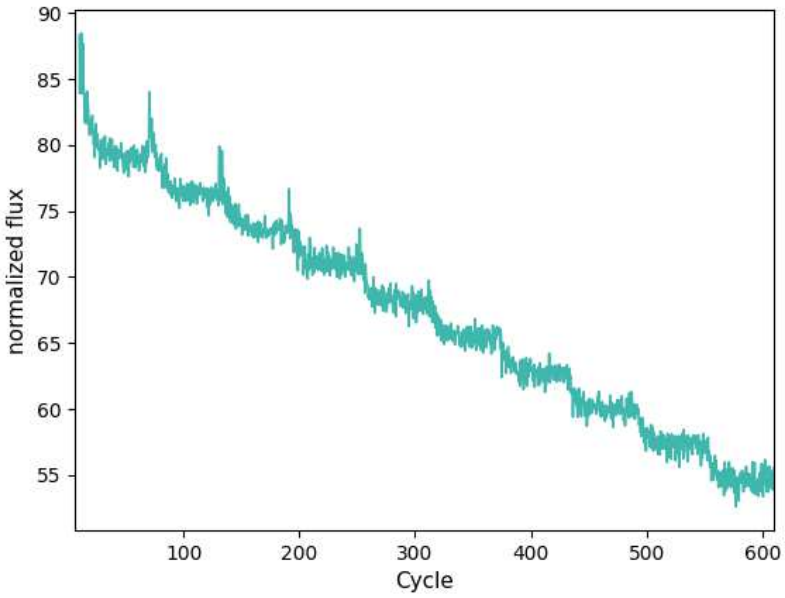


Fig. 6. Relation between operational parameters and permeate flow. This figure compares different input parameters to permeate flow for a few cycles before chemical cleaning and a few cycles after chemical cleaning. These are sample data that were fed to LSTM. With the aid of cyclic operational parameters, LSTM can capture the relation between operating parameters and the aging rate. The experimental cycles at the beginning of batches after chemical cleaning were different from the cycles close to the clogging of the unit due to the use of backflush pulses. This was mainly because, with the imposed novel cycles at the time of clogging, more data could be gathered.

Fig. 7 demonstrates the degradation rate of the PVDF DCMD module. The permeate fluxes were normalized based on the maximum permeate flux of the module in its virgin state under the suggested operation condition in the PVDF DCMD module datasheet. The starting normalized flux was 88%; this value decreased to approximately 78% at the end of the first batch. This decrease became faster in the first 40 cycles in the first batch compared to the subsequent 20 cycles. This decrease in permeate flux was due mainly to the increase in concentration in subsequent cycles. However, the normalized flux increased at the start of the second batch because the module had enough time between these two-cycle batches to reach the ambient temperature after chemical cleaning. Although the normalized flux increased in the second batch, it did not reach the same value as in the first batch. As demonstrated in Fig. 8, the normalized flux at the start of the second

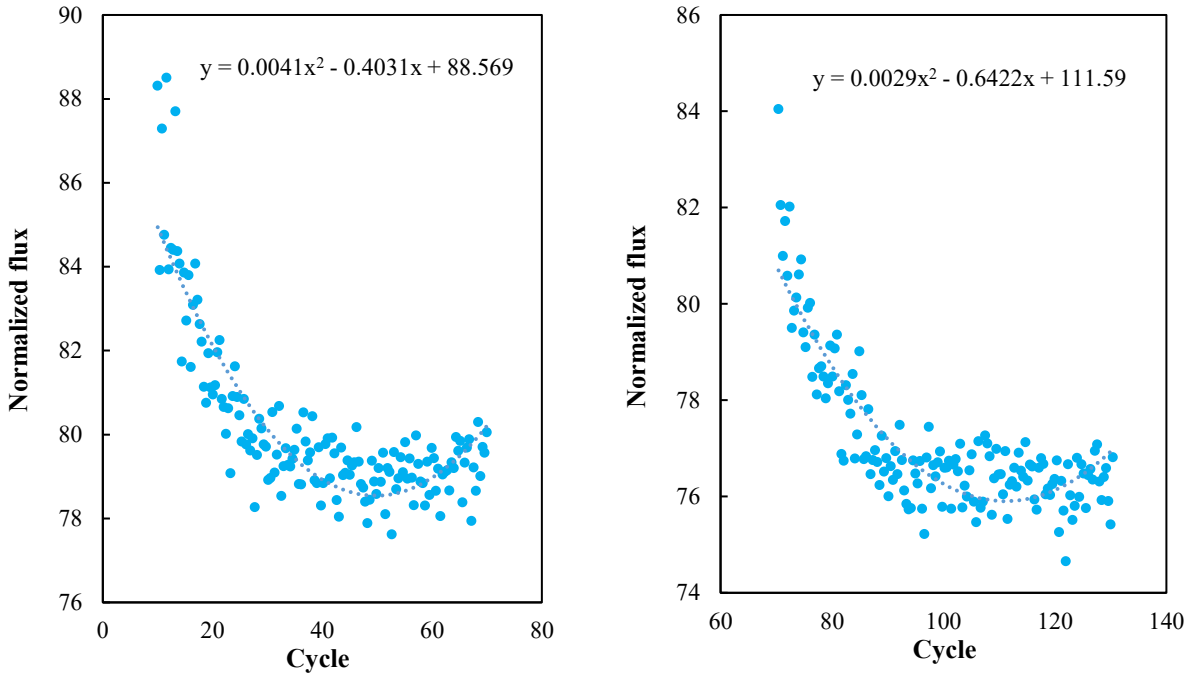
batch was approximately 84%. This 4% difference represents the permanent damage induced by the purification and cleaning process. The decrease of 10% (88% to 78%) in the first 60 cycles represented a reversible normalized flux decrease. In the second 60 cycles, the module experienced a 9% reversible normalized flux decrease and irreversible degradation of roughly 4%. In the sixth batch, the reversible decrease in normalized flux was approximately 5%, while the reduction in efficiency was 4%, which shows that the difference between the reversible and irreversible reduction in normalized flux decreased considerably from the sixth batch onwards. On the other hand, the sharp decrease in the first few batches settled from the sixth batch onwards. The normalized flux of the PVDF module continued to decrease up to roughly 55%. From this point, approximately half of the inlet water could not be purified. Therefore, using modules from this point is not beneficial. Furthermore, in the last batches, the main reduction in permeate flow was due to irreversible aging rather than a reversible reduction in permeate flow, which shows that the rate of irreversible damage accelerated from the sixth batch. The reversible and irreversible normalized flux reduction rates in each batch of the experiment are illustrated in Fig. 9.



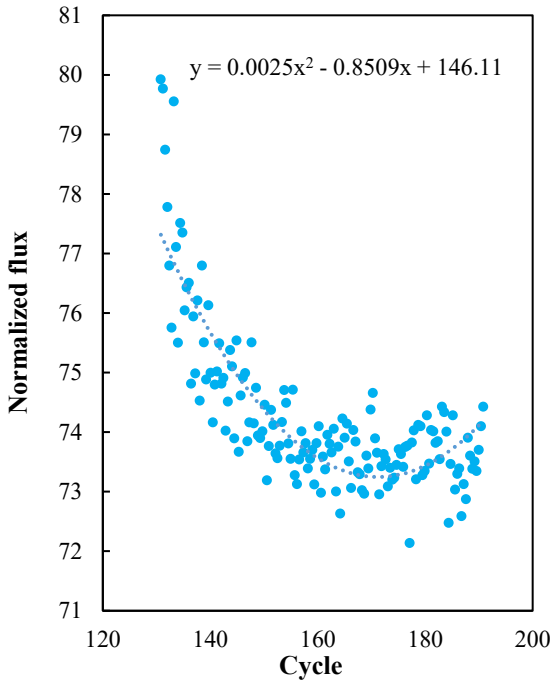
389 Fig. 7. Decrease rate of normalized flux over cycles.

390

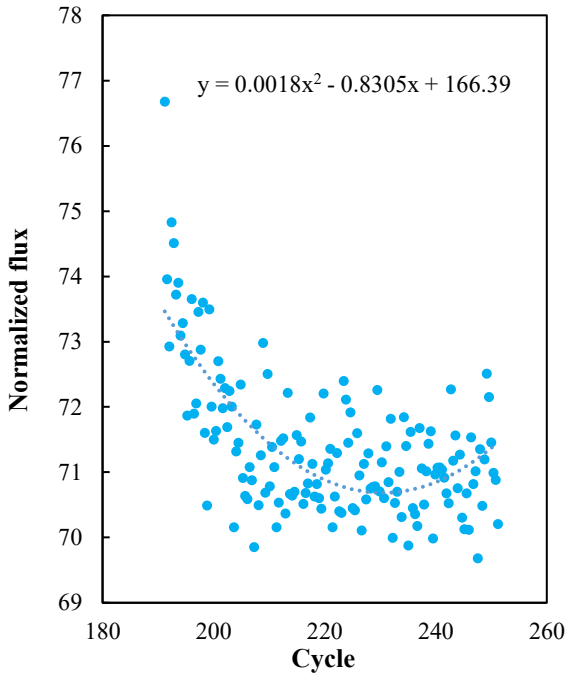
391



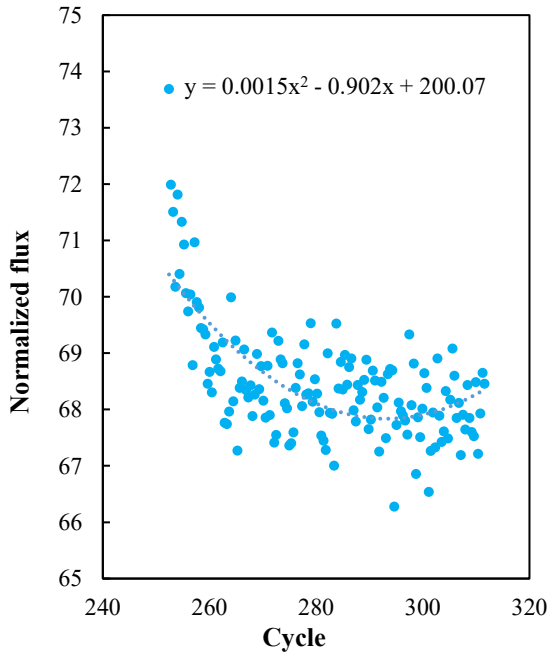
(a) (b)


y = 0.0025x² - 0.8509x + 146.11
Normalized flux
Cycle
81
80
79
78
77
76
75
74
73
72
71
120
140
160
180
200

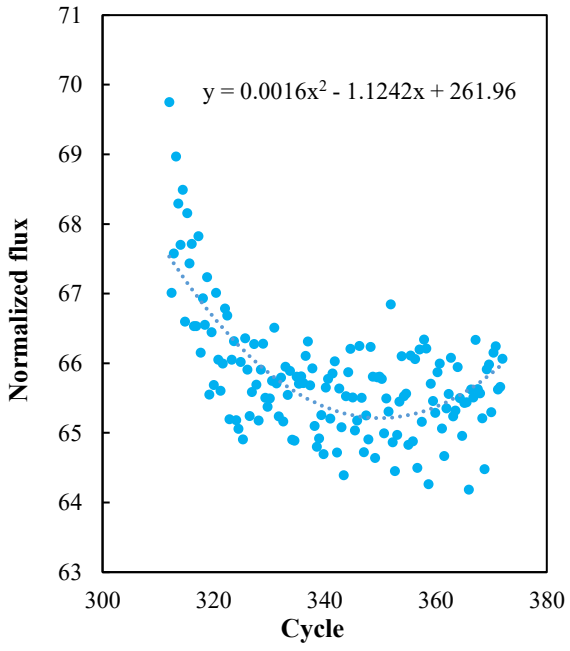
(c)


y = 0.0018x² - 0.8305x + 166.39
Normalized flux
Cycle
78
77
76
75
74
73
72
71
70
69
180
200
220
240
260

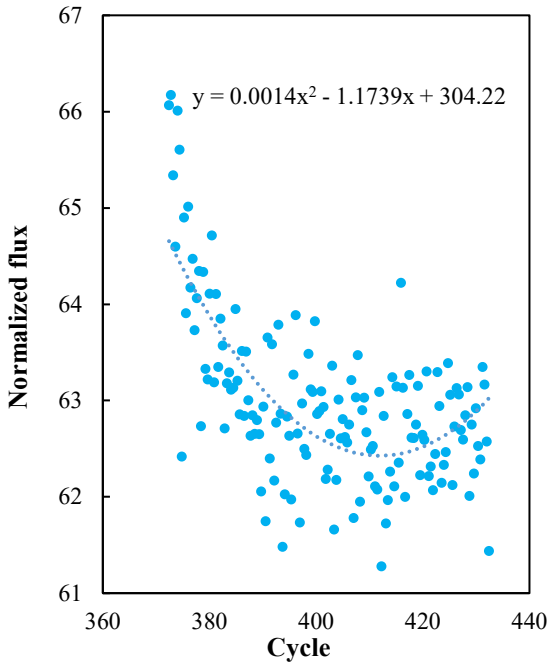
(d)


y = 0.0015x² - 0.902x + 200.07
Normalized flux
Cycle
75
74
73
72
71
70
69
68
67
66
65
240
260
280
300
320

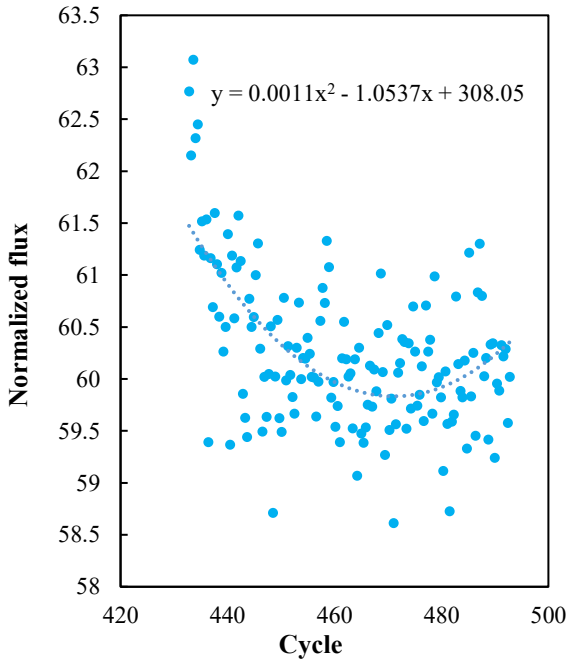
(e)


y = 0.0016x² - 1.1242x + 261.96
Normalized flux
Cycle
71
70
69
68
67
66
65
64
63
300
320
340
360
380

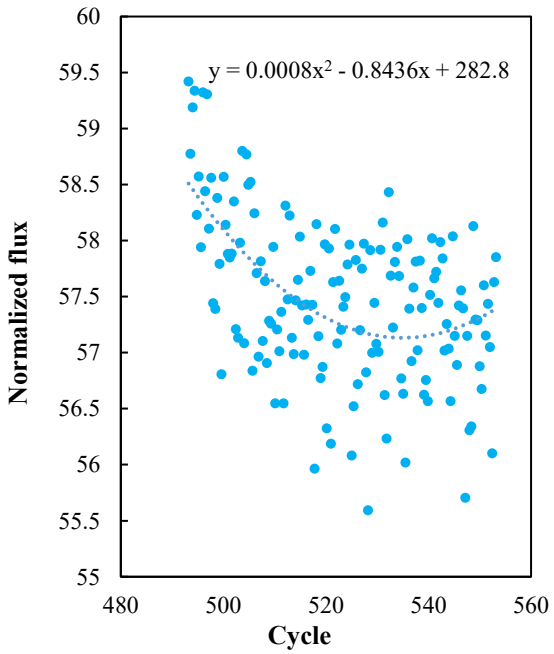
(f)



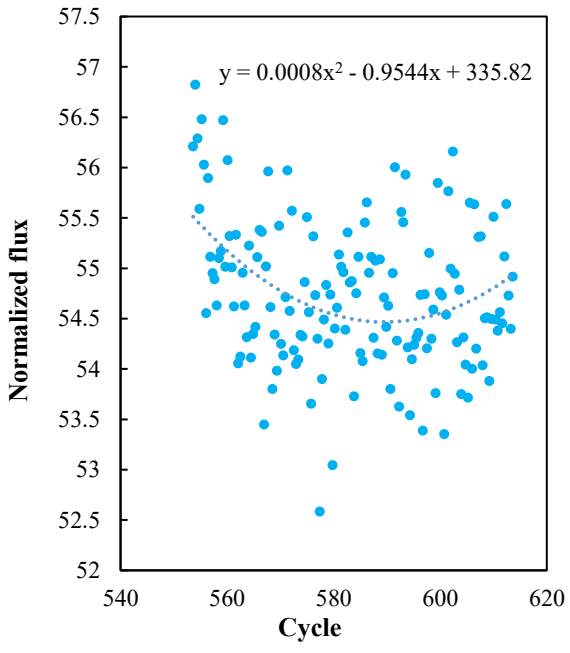
(g)



(h)



(i)



(j)

392 Fig. 8. Aging rate of the DCMD module for each batch.

As shown, the sharp decrease in the first few cycles alleviated from the sixth batch. Additionally, the difference between the reversible and irreversible reduction of normalized flux decreased across batches.

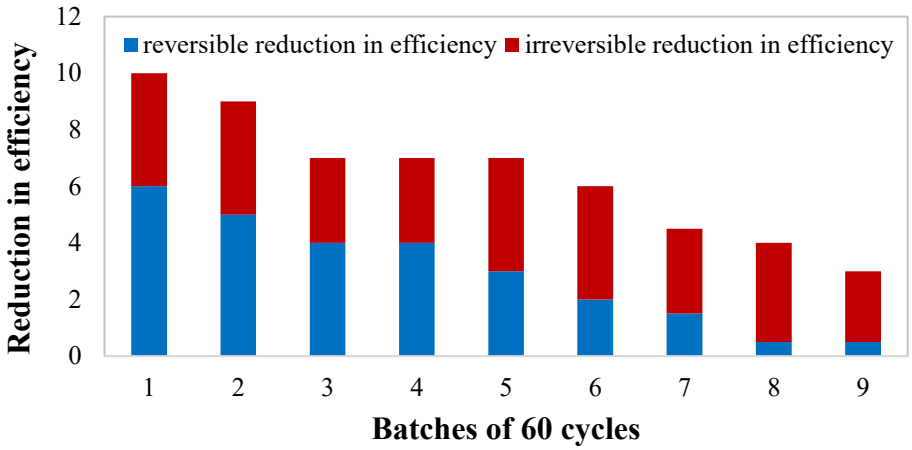


Fig. 9. Reversible and irreversible reduction in normalized flux in each batch over 60 cycles.

The time series analysis technique was implemented as the first step for predicting the aging of the DCMD module. This method employs the vector autoregressions (VAR) and integrated autoregressive moving average (ARIMA) models to predict time series data, and was implemented in this study to predict the aging of the DCMD module under different operational conditions. This aging prediction is easy to implement and can help technicians in their decision-making process related to the replacement of PVDF modules. Particularly, the seasonal decomposition procedure was applied to divide the obtained data into three different factors: trend, oscillation, and residual. The forecast for replacing the PVDF modules was made based on the trend section found after the seasonal decomposition procedure was undertaken. For the oscillation and residual factors, mechanical prediction methods were applied. The seasonal decomposition results for the obtained data are shown in Fig. 10. The trend section in Fig. 10 can be used for aging estimation of the

PVDF DCMD module under normal operation conditions, that is, inlet pressure of 2–2.5 bar and working temperature of 30–40 °C. The relation obtained for the calculation of module aging is:

$$\varepsilon = -\frac{1}{24}\, C + 80 \qquad (9)$$

Here *C* represents the number of cycles and ε is the estimated normalized flux of the PVDF DCMD module after *C* cycles. Results can be obtained by first estimating the trend by applying a convolution filter to the data and then removing the trend from the series to find the average of this de-trended series for each period, which represents the returned oscillation component. The reversible decrease in module efficiency can be estimated from the oscillation section of Fig. 10. The residual in seasonal decomposition represents what will remain after seasonality and trends are removed. Time series models assume that the data is stationary and only the residual component satisfies the conditions for stationarity. The bounds of noises in each level of operation can be estimated from the residual part of Fig. 9.

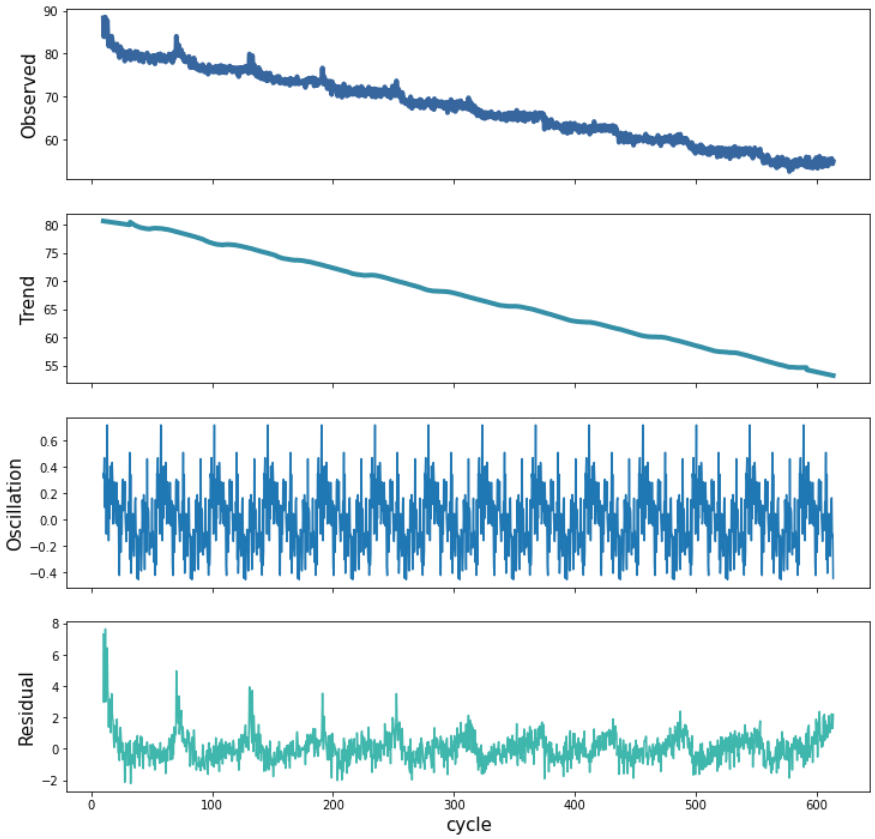


Fig. 10. Decomposition of the reduction rate of the permeate flow over time.

4.1 Deep learning

In the first step, the PVDF DCMD module's degradation data were utilized in an effort to train LSTM. The trained model was used to authenticate the prediction outcomes. Furthermore, as shown in Fig. 11, an important part of the LSTM training process is to bring the window size for training data up to date. In the training process for RNN, the data are also used to verify the network structure one step at a time, whereby the anticipation of the next time step relies on the previous time step and the figure obtained several steps before.

The observed efficiency from the first cycle through to the 420th was implemented to develop the RNN network. The efficiency data from the 421st cycle to the 620th cycle was employed to attest to the validity of the predicted efficiency. For validation purposes, the size of the sliding window

was considered to be 10 cycles while the number of hidden neuron layers was 20 and the delay layer was set at four cycles. The training data were distributed uniformly as the PVDF module temperature was kept homogenous. Additionally, to ensure improved evaluation of the trained model, random sampling was performed on the training and validation datasets. For the uniform distribution of the training data, the heating and cooling temperature of the DCMD module was controlled by a PID controller. The PID controller induced the heating element to operate sporadically when the DCMD unit worked under lower operational conditions, specifically related to pressure and flow rate. This was the core reason the outlet temperature varied considerably. When the temperature of the inlet water was set to the lower limit (room temperature), the PID stopped sending signals to the element and the element therefore stopped heating the water. When a higher inlet temperature was needed or the temperate of the heated water decreased, the PID started sending signals to the heater. The variation in efficiency occurred in the training segment of data as well as the oscillation of efficiency in the test segment. Particularly, these variations intensified after the 500th cycle. The predicted efficiencies based on the developed LSTM network implied a sound agreement with the observed data. As shown in Fig. 12, the values of the MAPE for the PVDF unit's degradation under the cyclic operation temperature between 20°C to 45°C and pressure between 0 to 2.2 bar were 0.0063 and 0.0057. The former value is the MAPE related to the training section and the latter represents the MAPE of the prediction section. The decrease of 0.0006 in the value of the MAPE illustrates the superior functionality of the RNN network in the prediction phase compared to the training phase. The coefficient of determination (R2) for the developed network was approximately 0.87.

A comparison of the network between the prediction and training phases revealed that the efficiency curve for the prediction phase was smoother. The developed RNN network illustrated

456 that higher operating temperature and pressure result in a smoother LSTM-based prediction. In
457 harsher operation conditions, the value of the MAPE was shown to decline for both the prediction
458 and training phases, while the R2 increased to approximately 0.92, which demonstrates a strong
459 agreement between the predicted and training data. Conversely, the predicted efficiency based on
460 the Recursive Mean-Variance (RMV) model did not show a strong agreement, especially between
461 the 500th and 600th cycles. The developed Relevance Vector Machine (RVM) model showed
462 improved behavior in less intense operation conditions where the PVDF degradation was much
463 smoother.

464 As represented and validated by the empirical data, the projected efficiencies based on the LSTM
465 (RNN) short-term prediction are considered accurate. Therefore, the degradation rate of the PVDF
466 modules can be estimated with a high level of confidence based on this RNN. Training data
467 distribution was found to have a severe impact on trained network accuracy and decrease the
468 prediction-related performance of an RNN structure. As a uniform distribution of the obtained data
469 was attempted, it can be concluded that the trained model has significant reliability for the
470 prediction of aging effects on the PVDF DCMD module. The error of the trained LSTM network
471 is plotted in Fig. 13.

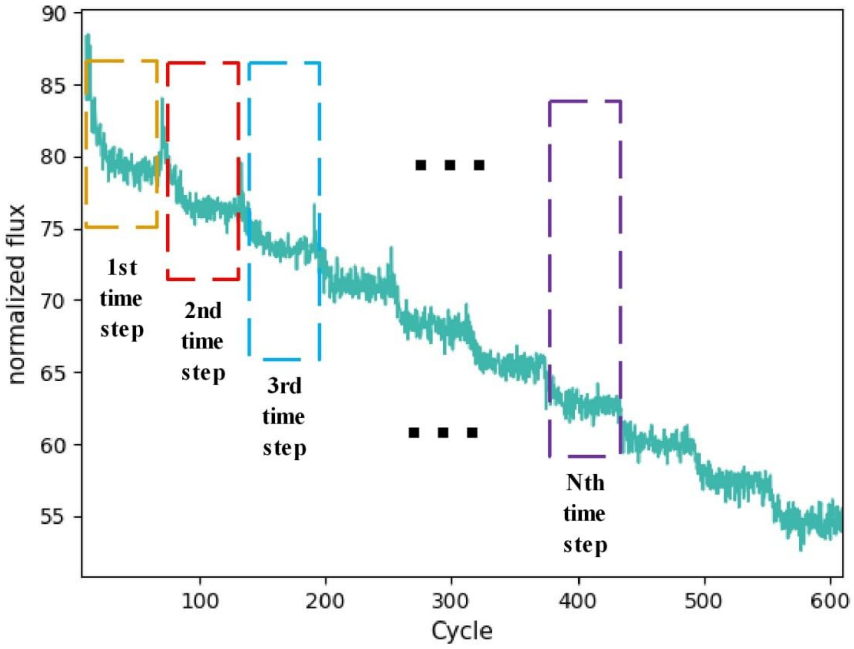


Fig. 11. Sliding windows for LSTM- RNN training process.

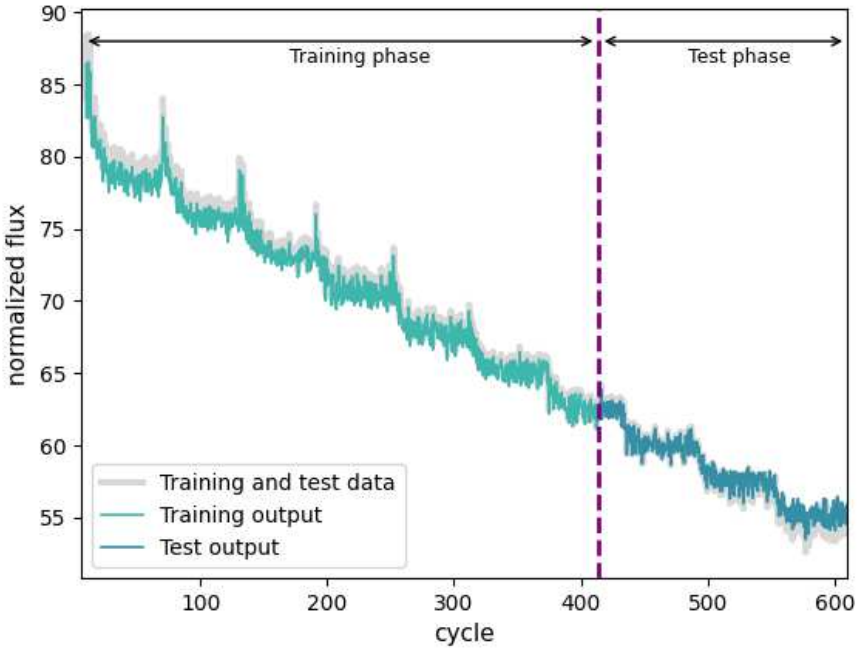


Fig. 12. LSTM-RNN prediction curve for PVDF DCMD aging rate.

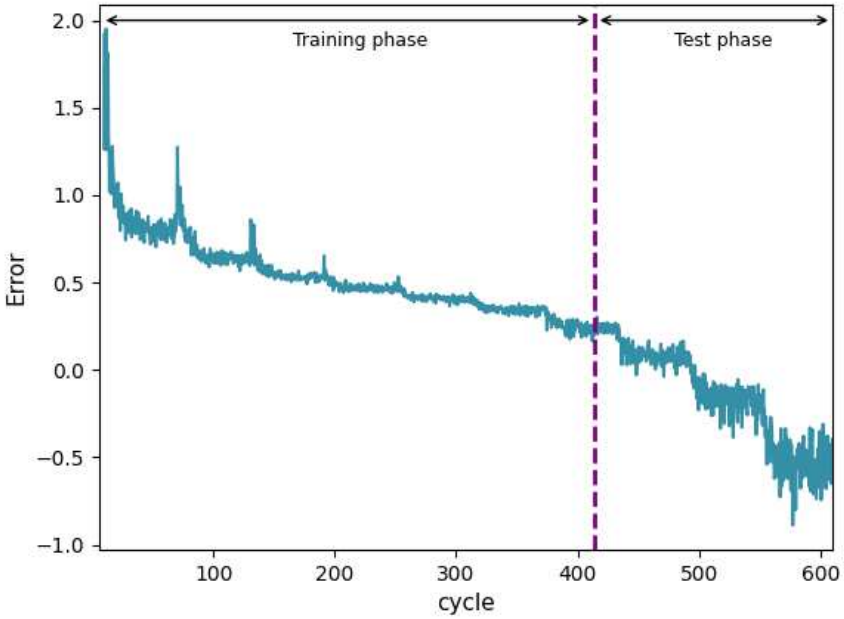


Fig. 13. Error of the LSTM- RNN trained network.

5 Conclusion

This study proposed a novel data-driven aging correlation and recurrent neural network - long short-term memory model based on the operation condition for PVDF DCMD modules. Based on the obtained data, the reduction in the permeate flux of PVDF DCMD modules was found to consist of two core parts: a short-term aging effect (reversible decrease in normalized flux) and a long-term aging effect (irreversible decrease in the normalized flux due to irreversible damage). The short-term aging was mainly due to the increase in the module's temperature and concentration. This aging effect was found to be reversible through chemical cleaning. The long-term aging effect was shown to be an irreversible process, predominantly because of the damage induced to the unit due to the purification and cleaning process. The rate of the short-term aging effect was found to be much higher than the rate of the long-term aging effect. Both short- and long-term aging effects can be estimated by the obtained model. These results were found under different DCMD unit operating conditions. The trained recurrent neural network - long short-term memory model achieved a mean absolute error percentage of 0.57% for the aging rate. The ability to predict the aging rate provides significant practical information, demonstrating the model's utility in optimizing operation conditions and enhancing the module's lifetime performance. The aging rate of the membrane can be approximated as a 0.17% decrease in normalized flux per cycle in the initial stages, with additional irreversible damage of about 4% after each cleaning cycle. By using the trained deep learning model and specifying the working temperature, inlet pressure, and inlet flow rate, operators can accurately estimate the aging rate of a PVDF DCMD module as well as its cleaning efficiency. Further, a simple algebraic relation was obtained for the approximate primary estimation of a PVDF DCMD module condition.

Appendix A

The output activation and gate activation functions can be stated as:

$$\sigma(z) = y = \frac{1}{1+e^{-z}}, \sigma'(z) = y(1-y) \tag{9}$$

$$\tanh(z) = y = \frac{e^{z}-e^{-z}}{e^{z}+e^{-z}}, \tanh'(z) = 1-y^{2} \tag{10}$$

Characterizing work ° as multiplied by the components:

$$a^{\circ}b = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \\ a_n \end{bmatrix} \circ \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \dots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 b_1 \\ a_2 b_2 \\ a_3 b_3 \\ \dots \\ a_n b_n \end{bmatrix} \tag{11}$$

$$a^{\circ}X = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \\ a_n \end{bmatrix} \circ \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2n} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3n} \\ & & \dots & & \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{nn} \end{bmatrix} = \begin{bmatrix} a_1 x_{11} & a_1 x_{12} & a_1 x_{13} & \dots & a_1 x_{1n} \\ a_2 x_{21} & a_2 x_{22} & a_2 x_{23} & \dots & a_2 x_{2n} \\ a_3 x_{31} & a_3 x_{32} & a_3 x_{33} & \dots & a_3 x_{3n} \\ & & \dots & & \\ a_n x_{n1} & a_n x_{n2} & a_n x_{n3} & \dots & a_n x_{nn} \end{bmatrix} \tag{12}$$

The error term at time step t (δ_t) is expressed as:

$$\delta_t = \frac{\partial E}{\partial h_t} \tag{13}$$

The error reverse passing along backpropagation, prompts to the error term at time step t−1 (δ_{t-1}) as:

$$\delta_{t-1} = {\delta_{0,t}}^{T} \frac{\partial net_{0,t}}{\partial h_{t-1}} + {\delta_{f,t}}^{T} \frac{\partial net_{f,t}}{\partial h_{t-1}} + {\delta_{i,t}}^{T} \frac{\partial net_{i,t}}{\partial h_{t-1}} + {\delta_{\tilde{C},t}}^{T} \frac{\partial net_{\tilde{C},t}}{\partial h_{t-1}} = {\delta_{0,t}}^{T} \boldsymbol{W}_0 + {\delta_{f,t}}^{T} \boldsymbol{W}_f + {\delta_{i,t}}^{T} \boldsymbol{W}_i +$$

$${\delta_{\tilde{C},t}}^{T} \boldsymbol{W}_{\tilde{C}} \tag{14}$$

Where,

$${\delta_{0,t}}^{T} = {\delta_t}^{T} \circ \tanh(c_t) \circ o_t \circ (1-o_t) \tag{15}$$

$${\delta_{f,t}}^{T} = {\delta_t}^{T} \circ o_t \circ (1-\tanh(c_t)^2) \circ c_{t-1} \circ f_t \circ (1-f_t) \tag{16}$$

$${\delta_{i,t}}^{T} = {\delta_t}^{T} \circ o_t \circ (1-\tanh(c_t)^2) \circ \tilde{c}_t \circ i_t \circ (1-i_t) \tag{17}$$

$$\delta_{0,t}{}^{T} = \delta_t{}^{T} \circ o_t \circ (1 - \tanh(c_t)^2) \circ i_t \circ (1 - \tilde{c}_t{}^2) \quad (18)$$

Therefore, to any time step k, the equation is obtained for passing forward the error term as:

$$\delta_k{}^{T} = \prod_{j=k}^{t-1} \delta_{0,j}{}^{T} \boldsymbol{W}_0 + \delta_{f,j}{}^{T} \boldsymbol{W}_f + \delta_{i,j}{}^{T} \boldsymbol{W}_i + \delta_{\tilde{C},j}{}^{T} \boldsymbol{W}_{\tilde{C}} \quad (19)$$

Then, at that point, the weight inclination at time step t is accomplished as:

$$\frac{\partial E}{\partial \boldsymbol{W}_{0,t}} = \frac{\partial E}{\partial net_{0,t}} \frac{\partial net_{0,t}}{\partial \boldsymbol{W}_{0,t}} = \delta_{0,t} h_{t-1}{}^{T} \quad (20)$$

$$\frac{\partial E}{\partial \boldsymbol{W}_{f,t}} = \frac{\partial E}{\partial net_{f,t}} \frac{\partial net_{f,t}}{\partial \boldsymbol{W}_{f,t}} = \delta_{f,t} h_{t-1}{}^{T} \quad (21)$$

$$\frac{\partial E}{\partial \boldsymbol{W}_{i,t}} = \frac{\partial E}{\partial net_{i,t}} \frac{\partial net_{i,t}}{\partial \boldsymbol{W}_{i,t}} = \delta_{i,t} h_{t-1}{}^{T} \quad (22)$$

$$\frac{\partial E}{\partial \boldsymbol{W}_{\tilde{C},t}} = \frac{\partial E}{\partial net_{\tilde{C},t}} \frac{\partial net_{\tilde{C},t}}{\partial \boldsymbol{W}_{\tilde{C},t}} = \delta_{\tilde{C},t} h_{t-1}{}^{T} \quad (23)$$

The ultimate gradient for weight matrixes is stated as:

$$\frac{\partial E}{\partial \boldsymbol{W}} = \sum_{j=1}^{t} \delta_j \, h_{t-1}{}^{T} \quad (24)$$

Similarly, the gradient for bias vectors is achieved as:

$$\frac{\partial E}{\partial \boldsymbol{b}_{0,t}} = \frac{\partial E}{\partial net_{0,t}} \frac{\partial net_{0,t}}{\partial \boldsymbol{b}_{0,t}} = \delta_{0,t} \quad (25)$$

$$\frac{\partial E}{\partial \boldsymbol{b}_{f,t}} = \frac{\partial E}{\partial net_{f,t}} \frac{\partial net_{f,t}}{\partial \boldsymbol{b}_{f,t}} = \delta_{f,t} \quad (26)$$

$$\frac{\partial E}{\partial \boldsymbol{b}_{i,t}} = \frac{\partial E}{\partial net_{i,t}} \frac{\partial net_{i,t}}{\partial \boldsymbol{b}_{i,t}} = \delta_{i,t} \quad (27)$$

$$\frac{\partial E}{\partial \boldsymbol{b}_{\tilde{C},t}} = \frac{\partial E}{\partial net_{\tilde{C},t}} \frac{\partial net_{\tilde{C},t}}{\partial \boldsymbol{b}_{\tilde{C},t}} = \delta_{\tilde{C},t} \quad (28)$$

The ultimate gradient is obtained for bias vector as:

$$\frac{\partial E}{\partial \boldsymbol{b}} = \sum_{j=1}^{t} \delta_j \quad (29)$$

Nomenclature

Abbreviation	Meaning
ANNs	Artificial Neural Networks
ARIMA	Integrated Autoregressive Moving Average
BPTT	Backpropagation Through Time
LSTM	Long Short-Term Memory
MAPE	Mean Absolute Error Percentage
MD	Membrane Distillation
PID	Proportional Integral Derivative
PVDF	Polyvinylidene Fluoride
RO	Reverse Osmosis
RMV	Recursive Mean Variance
RVM	Relevance Vector Machine
RNNs	Recurrent Neural Networks
DCMD	Direct Contact Membrane Distillation
VAR	Vector Autoregressions
WHO	World Health Organization

Symbol	Description
δ	Error for each neuron
W	Self-recurrent connection's weight matrix
net_i	Neuron's weighted input

β_W and β_f	Norms of the upper bound
$f'(net_i)$	Diagonal matrix diag
X_t	Current time step's input
h_t	Result of the present time steps
C_t	Present time step's memory
b_f , b_i, b_o and b_C	Bias vectors
°	Multiplication by the elements
R2	Coefficient of determination
X(t)	Amount of actual water salinity measured
$\widehat{X}(t)$	Amount of water salinity predicted by the degradation model
$\overline{X}(t)$	Average amount of water salinity data
N	Number of values of salinity measured

547

548

549

550

551

552

553

6 Statements and Declarations

6.1 Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

6.2 Authors' Contributions

Amir Hossein Zabihi Tari, Amir H. Keshavarzzadeh, and Pouria Ahmadi contributed to the conception and development of the study and to interpretation of the results. Amir Hossein Zabihi Tari prepared the first draft of the manuscript. All authors reviewed the manuscript and read and approved the final version.

6.3 Ethical Approval

Not applicable. This study did not involve human participants or animals.

6.4 Consent to Participate

Not applicable. This study did not involve human participants.

6.5 Consent to Publish

Not applicable. The manuscript does not contain identifiable information from individual participants.

6.6 Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

6.7 Data Availability Statement

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

7 References

Abdullah SZ, Bérubé PR (2013) Assessing the effects of sodium hypochlorite exposure on the characteristics of PVDF based membranes. Water Res 47(14):5392–5399. https://doi.org/10.1016/j.watres.2013.06.018

Abdullah SZ, Bérubé PR (2018) Filtration and cleaning performances of PVDF membranes aged with exposure to sodium hypochlorite. Sep Purif Technol 195:253–259. https://doi.org/10.1016/j.seppur.2017.12.004

Chen Q, Zhang W, Lou Y (2020) Forecasting stock prices using a hybrid deep learning model integrating attention mechanism, multi-layer perceptron, and bidirectional long-short term memory neural network. IEEE Access 8:117365–117376. https://doi.org/10.1109/ACCESS.2020.3004284

Choi H, Cho K, Bengio Y (2018) Fine-grained attention mechanism for neural machine translation. Neurocomputing 284:171–176. https://doi.org/10.1016/j.neucom.2018.01.007

585 Rafael de Medeiros Paulino (2019) Modelling of drinking water treatment and disinfection by-product
586 formation with artificial neural networks. Master of Applied Science thesis, University of British
587 Columbia, Vancouver, Canada. https://doi.org/10.14288/1.0380442
588 Delattre J, Rabaud B, Bréhant A, Glucina K, Sollogoub C, Thominette F (2012) Ageing of hollow fiber
589 membranes in polyvinylidene fluoride (PVDF) used in water treatment. Procedia Eng 44:764–767.
590 https://doi.org/10.1016/j.proeng.2012.08.562
591 Dürrenmatt DJ, Gujer W (2012) Data-driven modeling approaches to support wastewater treatment plant
592 operation. Environ Model Softw 30:47–56. https://doi.org/10.1016/j.envsoft.2011.11.007
593 Fane AG (2007) Sustainability and membrane processing of wastewater for reuse. Desalination 202(1–
594 3):53–58. https://doi.org/10.1016/j.desal.2005.12.038
595 Feng J, Feng X, Chen J, Cao X, Zhang X, Jiao L, Yu T (2020) Generative adversarial networks based on
596 collaborative learning and attention mechanism for hyperspectral image classification. Remote Sens
597 12(7):1149. https://doi.org/10.3390/rs12071149
598 Gers FA, Schraudolph NN, Schmidhuber J (2002) Learning precise timing with LSTM recurrent networks.
599 J Mach Learn Res 3:115–143.
600 Hajibabania S, Antony A, Leslie G, Le-Clech P (2012) Relative impact of fouling and cleaning on PVDF
601 membrane hydraulic performances. Sep Purif Technol 90:204–212.
602 https://doi.org/10.1016/j.seppur.2012.03.001
603 Hashim NA, Liu Y, Li K (2011) Stability of PVDF hollow fibre membranes in sodium hydroxide aqueous
604 solution. Chem Eng Sci 66(8):1565–1575. https://doi.org/10.1016/j.ces.2010.12.019
605 Huang PY, Liu F, Shiang SR, Oh J, Dyer C (2016) Attention-based multimodal neural machine translation.
606 In: Proceedings of the First Conference on Machine Translation, Volume 2: Shared Task Papers, Berlin,
607 Germany, pp 639–645. Association for Computational Linguistics. https://doi.org/10.18653/v1/W16-
608 2360
609 Kalla S (2021) Use of membrane distillation for oily wastewater treatment – a review. J Environ Chem Eng
610 9(1):104641. https://doi.org/10.1016/j.jece.2020.104641

611 Kimura K, Yamamura H, Watanabe Y (2006) Irreversible fouling in MF/UF membranes caused by natural
612 organic matters (NOMs) isolated from different origins. Sep Sci Technol 41(7):1331–1344.
613 https://doi.org/10.1080/01496390600634665
614 Levitsky I, Duek A, Arkhangelsky E, Pinchev D, Kadoshian T, Shetrit H, Naim R, Gitis V (2011)
615 Understanding the oxidative cleaning of UF membranes. J Membr Sci 377(1–2):206–213.
616 https://doi.org/10.1016/j.memsci.2011.04.046
617 Li L, Rong S, Wang R, Yu S (2021) Recent advances in artificial intelligence and machine learning for
618 nonlinear relationship analysis and process control in drinking water treatment: A review. Chem Eng J
619 405:126673. https://doi.org/10.1016/j.cej.2020.126673
620 Ma R, Yang T, Breaz E, Li Z, Briois P, Gao F (2018) Data-driven proton exchange membrane fuel cell
621 degradation predication through deep learning method. Appl Energy 231:102–115.
622 https://doi.org/10.1016/j.apenergy.2018.09.111
623 Ma W, Yang Q, Wu Y, Zhao W, Zhang X (2019) Double-branch multi-attention mechanism network for
624 hyperspectral image classification. Remote Sens 11(11):1307. https://doi.org/10.3390/rs11111307
625 Micale G, Rizzuti L, Cipollina A (eds) (2009) Seawater desalination: conventional and renewable energy
626 processes. Green Energy and Technology. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-
627 642-01150-4
628 Nagy HM, Watanabe K, Hirano M (2002) Prediction of sediment load concentration in rivers using artificial
629 neural network model. J Hydraul Eng 128(6):588–595. https://doi.org/10.1061/(ASCE)0733-
630 9429(2002)128:6(588)
631 Nunes SP, Peinemann KV (eds) (2001) Membrane technology in the chemical industry. Wiley-VCH,
632 Weinheim. https://doi.org/10.1002/3527600388
633 Pessoa de Amorim MT, Afonso Ramos IR (2006) Control of irreversible fouling by application of dynamic
634 membranes. Desalination 192(1–3):63–67. https://doi.org/10.1016/j.desal.2005.10.011
635 Porcelli N, Judd S (2010) Chemical cleaning of potable water membranes: A review. Sep Purif Technol
636 71(2):137–143. https://doi.org/10.1016/j.seppur.2009.12.007

637 Puspitasari V, Granville A, Le-Clech P, Chen V (2010) Cleaning and ageing effect of sodium hypochlorite
638 on polyvinylidene fluoride (PVDF) membrane. Sep Purif Technol 72(3):301–308.
639 https://doi.org/10.1016/j.seppur.2010.03.001
640 Qiu J, Wang B, Zhou C (2020) Forecasting stock prices with long-short term memory neural network based
641 on attention mechanism. PLoS One 15(1):e0227222. https://doi.org/10.1371/journal.pone.0227222
642 Regula C, Carretier E, Wyart Y, Sergent M, Gésan-Guiziou G, Ferry D, Vincent A, Boudot D, Moulin P
643 (2013) Ageing of ultrafiltration membranes in contact with sodium hypochlorite and commercial
644 oxidant: Experimental designs as a new ageing protocol. Sep Purif Technol 103:119–138.
645 https://doi.org/10.1016/j.seppur.2012.10.010
646 Ribeiro Pinela S, Larasati A, Meulepas RJW, Gagliano MC, Kleerebezem R, Bruning H, Rijnaarts HHM
647 (2024) Ultrafiltration (UF) and biological oxygen-dosed activated carbon (BODAC) filtration to
648 prevent fouling of reverse osmosis (RO) membranes: A mass balance analysis. J Water Process Eng
649 57:104648. https://doi.org/10.1016/j.jwpe.2023.104648
650 Robinson S, Abdullah SZ, Bérubé P, Le-Clech P (2016) Aging of membranes for water treatment: Linking
651 changes to performance. J Membr Sci 503:177–187. https://doi.org/10.1016/j.memsci.2015.12.033
652 Schwantes R, Chavan K, Winter D, Felsmann C, Pfafferott J (2018) Techno-economic comparison of
653 membrane distillation and MVC in a zero liquid discharge application. Desalination 428:50–68.
654 https://doi.org/10.1016/j.desal.2017.11.026
655 Shahrim NAA, Abounahia NM, El-Sayed AMA, Saleem H, Zaidi SJ (2023) An overview on the progress
656 in produced water desalination by membrane-based technology. J Water Process Eng 51:103479.
657 https://doi.org/10.1016/j.jwpe.2022.103479
658 Shannon MA, Bohn PW, Elimelech M, Georgiadis JG, Mariñas BJ, Mayes AM (2008) Science and
659 technology for water purification in the coming decades. Nature 452(7185):301–310.
660 https://doi.org/10.1038/nature06599
661 Shao Y, Lan J, Liang Y, Hu J (2021) Residual networks with multi-attention mechanism for hyperspectral
662 image classification. Arab J Geosci 14(4):252. https://doi.org/10.1007/s12517-021-06516-6

Shih SY, Sun FK, Lee HY (2019) Temporal pattern attention for multivariate time series forecasting. Mach Learn 108(8–9):1421–1441. https://doi.org/10.1007/s10994-019-05815-0

Song X, Liu Y, Xue L, Wang J, Zhang J, Wang J, Jiang L, Cheng Z (2020) Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. J Pet Sci Eng 186:106682. https://doi.org/10.1016/j.petrol.2019.106682

Tari AHZ, Khosravi M, Maleki Dastjerdi S, Khoshnevisan A, Ahmadi P (2023) Multi-objective optimization and transient analysis of an off-grid building with water desalination and waste heat recovery units. Sustain Energy Technol Assess 59:103406. https://doi.org/10.1016/j.seta.2023.103406

Werbos PJ (1990) Backpropagation through time: what it does and how to do it. Proc IEEE 78(10):1550–1560. https://doi.org/10.1109/5.58337

Zaytar MA, El Amrani C (2016) Sequence to sequence weather forecasting with long short-term memory recurrent neural networks. Int J Comput Appl 143(11):7–11. https://doi.org/10.5120/ijca2016910497

Zhang B, Xiong D, Xie J, Su J (2020) Neural machine translation with GRU-gated attention model. IEEE Trans Neural Netw Learn Syst 31(11):4688–4698. https://doi.org/10.1109/TNNLS.2019.2957276

Zhang D, Lindholm G, Ratnaweera H (2018) Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring. J Hydrol 556:409–418. https://doi.org/10.1016/j.jhydrol.2017.11.018

Zhang G, Patuwo BE, Hu MY (1998) Forecasting with artificial neural networks: The state of the art. Int J Forecast 14(1):35–62. https://doi.org/10.1016/S0169-2070(97)00044-7

Zhang H, Zhao X (2023) Enhanced anti-wetting methods of hydrophobic membrane for membrane distillation. Adv Sci 10(23):2300598. https://doi.org/10.1002/advs.202300598

Zhang J, Man KF (1998) Time series prediction using RNN in multi-dimension embedding phase space. In: Proceedings of the 1998 IEEE International Conference on Systems, Man, and Cybernetics (SMC'98), San Diego, CA, USA, vol 2, pp 1868–1873. IEEE. https://doi.org/10.1109/ICSMC.1998.728168