ASSESSING MEMBRANE AGEING USING HISTORICAL DATA

by

Md Nurul Afcher Shishir

B.Sc. Engg.(Civil), Bangladesh University of Engineering and Technology, 2018

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Civil Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

January 2024

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

ASSESSING MEMBRANE AGEING USING HISTORICAL DATA

submitted by Md Nurul Afcher Shishir in partial fulfillment of the requirements for

the degree of Master of Applied Science

in Civil Engineering

Examining Committee:

Dr. Pierre R. Bérubé, Professor, Department of Civil Engineering, UBC
Supervisor

Dr. Jongho Lee, Associate Professor, Department of Civil Engineering, UBC
Supervisory Committee Member

Abstract

Historical data from four full-scale ultrafiltration membrane facilities was analyzed to gain insight into changes in membranes' performance over time, commonly referred to as membrane ageing. The analysis indicates that performance factors, such as hydraulically irreversible resistance, clean membrane resistance, total fouling rate, and the extent of resistance reversed during Backwash (BW), increase as membranes age. The rates at which these performance factors increased were substantially different between facilities indicating that membrane ageing is influenced by site-specific raw water and operational conditions. A framework was developed to forecast membrane replacement age using data from the initial years of operation. Adoption of this framework provides evidence-based tools to assist with the operation and long-term financial management of membrane infrastructure.

Lay Summary

Membrane filtration is a popular choice for water treatment due to the efficient removal of harmful pathogens and contaminants. However, membranes undergo changes in their operational and physical properties over time, which is popularly known as membrane aging. In the present study, historical data was acquired from four full-scale ultrafiltration membrane facilities and analyzed to examine the effects of ageing on the various operational properties (referred to as performance factors) of the membrane. The analysis indicates that, in general, as membranes age, they cannot filter as much water (i.e., resistance increases) and must be cleaned more frequently. Using historical data, a forecasting tool was developed to inform membrane replacement, which eventually would help the managers of full-scale ultrafiltration membrane facilities to make better financial plans.

Preface

This thesis is original, unpublished, and independent work by Md Nurul Afcher Shishir. Two conference presentations have resulted from this thesis as listed below.

Conference Presentations:

1. **Shishir, Md Nurul A.**, Robertson, Chanel J., Andrews, Robert C., & Bérubé, Pierre R. (2022). Parametric Evaluation of Historical Data of Ultrafiltration Systems for Drinking Water Treatment. Podium presentation at the Euro Membrane Conference, Sorrento, Italy.
    - This work was carried out by me, with Chanel Robertson executing and modifying the codes for different facilities under the supervision of Dr. Pierre Bérubé and the co-supervision of Dr. Robert Andrews.
2. **Shishir, Md Nurul A**., Robertson, Chanel J., Andrews, Robert C., & Bérubé, Pierre R. (2023). Evaluation of Long-Term Performance of Full-Scale UF Facilities Treating Drinking Water. Podium presentation at the IWA MTC, St. Louis, Missouri, USA.
    - This work was carried out by me, with Chanel Robertson executing and modifying the codes for different facilities under the supervision of Dr. Pierre Bérubé and the co-supervision of Dr. Robert Andrews.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ARIMA	Autoregressive Integrated Moving Average
BW	Backwash
CEB	Chemical Enhanced BW
LMH	Liter per m^2 per hour
MBR	Membrane Bioreactor
MC	Maintenance Cleaning
MF	Microfiltration
NF	Nanofiltration
RC	Recovery Cleaning
RMSE	Root Mean Square Error
RO	Reverse Osmosis
TMP	Trans-membrane Pressure
UF	Ultrafiltration
ZW	ZeeWeed

Acknowledgements

This dissertation is the culmination of the inputs from many great people, to whom the author is deeply grateful. Dr. Pierre Bérubé, conceived the project, supervised it throughout and provided all the financial and mental support needed to complete it.

The author would like to express gratitude to the members of 'Membrane Ageing Research' group - Professor Robert Andrews, Rahul Dutta, Baohui Jia, Jia Bian, Tyler Malkoske, and Malik Awan for sharing their valuable insights.

The author would like to acknowledge undergraduate student Chanel J Robertson, who helped in executing and modifying the codes for different facilities. The author is also grateful to Bakai Toichubekov and all other partner facilities for helping with the data acquisition process.

The author would like to extend heartfelt thanks to his wife, Samia Syeoti Ramim, for her enduring patience, unwavering support, and constant inspiration. The author is grateful to his parents for their unconditional support throughout his life.

To my parents

1 Introduction

Ultrafiltration (UF) is a popular choice for treating surface water. As long as UF membranes are not breached, they are capable of consistently removing particulates/contaminants regardless of potential fluctuation in raw water quality (Crittenden et al., 2012). As UF membranes filter water, material accumulates on the membrane surface or within the pores. This material is typically defined as foulant (U.S.EPA, 2005). In order to remove foulants, UF membrane facilities periodically perform hydraulic and chemical cleanings (Crittenden et al., 2012). Over time, these cleaning events can adversely affect the performance of UF membrane (e.g., increase in resistance, rate of fouling, etc.) and characteristics (e.g., decrease in hydrophilic additive contents, etc.). This progressive deterioration of the performance and characteristics of UF membrane is commonly defined as 'membrane ageing' (Robinson et al., 2016). Adverse effects due to membrane ageing are reported to be primarily responsible for UF membrane replacement (Prulho et al., 2013).

Lab-accelerated ageing studies and field-harvested membrane studies, the commonly used approaches to investigate the impacts of ageing on UF membranes, are either time or resource intensive (Robinson & Bérubé, 2020). These approaches also lack the capability to forecast UF membrane replacement age. In addition, lab-scale studies do not accurately reflect membrane ageing at full-scale facilities (Filho et al., 2021). As a result, limited knowledge exists regarding membrane ageing at full-scale facilities (Yu et al., 2021) and an evidence-based approach to inform UF membrane replacement has yet to be established (Fenu et al., 2012).

Fortunately, most full-scale UF membrane facilities collect extensive operational data. This operational data could potentially be used to investigate the impact of membrane ageing on performance (Robinson & Bérubé, 2020). The present study assessed if historical data from full-scale facilities could potentially be used to quantify membrane ageing and forecast membrane replacement age. Change in resistance to permeate flow was considered as a metric to assess membrane ageing. Historical data from four full-scale UF membrane facilities, operating in Canada, was considered. All facilities that were considered, used the same membrane type (ZeeWeed 1000, Veolia Water Technologies, Canada). However, each facility had different operating protocols. Combined, the historical data from the four full-scale UF membrane facilities represents over 30 years of operational data. A comprehensive statistical methodology was developed to 1) clean up relevant time series resistance data, and 2) extract summative parameters of time series resistance data. Membrane ageing was quantified based on the change in the extracted summative parameters. A framework was also developed to forecast time series resistance so that the replacement age of UF membranes could be estimated. The overall goal of the present study is to provide managers of full-scale UF membrane facilities with evidence-based tools to manage operations and forecast eventual membrane replacement.

2 Literature Review

2.1 Introduction

Membrane filtration employs a porous semipermeable physical barrier (commonly referred to as a membrane) to separate material of interest from a liquid stream (U.S.EPA, 2005). Based on the pore size, membranes can be classified into four categories: Microfiltration (MF, pore size ~ 0.1 µm), Ultrafiltration (UF, pore size ~ 0.01 µm), Nanofiltration (NF, pore size ~ 0.001 µm), and Reverse Osmosis (RO, nonporous) (Crittenden et al., 2012). MF and UF are collectively called low-pressure membranes, and typically used to remove particulate material, while NF and RO are collectively called high-pressure membranes, and typically used to remove soluble materials. For surface freshwater treatment, low-pressure membranes are typically used, because the group of contaminants that is typically of the greatest concern is particulate material (i.e., protozoa, bacteria, and virus) (Crittenden et al., 2012). The present study considers the use of UF membranes, hereafter simply referred to as membranes, for the treatment of surface freshwater.

The throughput of membranes (J) is proportional to the driving force (P) for permeation and inversely proportional to the product of the total membrane resistance (R_t) to the permeate flow, at given time (t), and the viscosity (μ_T) of the permeate flow at a given temperature (T) (U.S.EPA, 2005).

$$J = \frac{P}{R_t * \mu_T} \quad \text{Equation 1}$$

The driving force, commonly referred to as trans-membrane pressure (TMP), is the difference in pressure between the feed and permeate sides of the membrane. As membranes filter water, material, commonly referred to as foulants, is retained on the surface of the membrane or within membrane pores. The process of accumulation of foulants is referred to as fouling. Because of fouling, the resistance to the permeate flow generally increases with time (U.S.EPA, 2005). The total membrane resistance to the permeate flow is the sum of the intrinsic membrane resistance (R_m) and resistance due to fouling (R_f).

$$R_t = R_m + R_f \quad \text{Equation 2}$$

The intrinsic membrane resistance is dependent on the physical properties of membranes. This resistance is generally obtained from the membrane manufacturers or clean water filtration tests with the virgin (i.e., new) membranes. Resistance due to fouling (R_f) is the variable portion of the total membrane resistance, increasing during filtration and decreasing during cleaning as foulants are removed from the membrane. Membrane facilities typically monitor the total membrane resistance continuously or at the start and end of filtration cycles. For simplicity, hereafter, total membrane resistance (R_t) is referred to as resistance (R). Resistance is a commonly used metric for analyzing and comparing filtration and cleaning data for membrane facilities (Baars et al., 2005).

2.2 Membrane Cleaning

Physical and chemical cleaning are typically performed to remove retained foulants on/in the membranes, and as a consequence, decrease the resistance to permeate flow (Crittenden et al., 2012). Backwash (BW), which is achieved by reversing the flow direction through membranes, is commonly used for physical cleaning. BW is typically performed at a relatively high frequency (i.e., every hour); and is usually accompanied by some form of scouring to enhance the removal of foulants from a membrane surface. Scouring is commonly achieved by adding air at the membrane surface. However, some foulants cannot be removed by means of physical cleaning approaches. These foulants are commonly referred to as hydraulically irreversible foulants (Smith et al., 2005).

Chemical cleaning, which is achieved by exposing membranes to one or more chemical cleaning agents, is commonly used to remove hydraulically irreversible foulants (Regula et al., 2014). Two types of chemical cleaning approaches, both achieved by soaking the membranes in chemical cleaning agents, are typically used.

1) Mild chemical cleaning; commonly referred to as Maintenance Cleaning (MC). These are typically performed more frequently (e.g., daily to weekly) with a relatively low soak duration and low concentration of chemical cleaning agents (Wang et al., 2014).
2) Extensive chemical cleaning; commonly referred to as Recovery Cleaning (RC). These are typically performed less frequently (e.g., monthly) with a relatively high soak duration and high concentration of chemical cleaning agents (Wang et al., 2014).

Various chemical cleaning agents are commonly used in practice for both MC and RC, including bases (e.g., NaOH), oxidants (e.g., HOCl), acids (e.g., HCl), acid chelates (e.g., Citric acid), alkaline chelates (e.g., EDTA) and surfactants (Porcelli & Judd, 2010). The type and frequency of chemical cleaning agents to use are highly dependent on the nature of the foulants (Porcelli & Judd, 2010). Eventually, some of the resistance due to fouling cannot be completely removed with chemical cleanings. This resistance is commonly defined as chemically irreversible resistance (Wang et al., 2014). A pictorial representation of the evolution of resistance to permeate flow in a membrane system is presented in Figure 1.

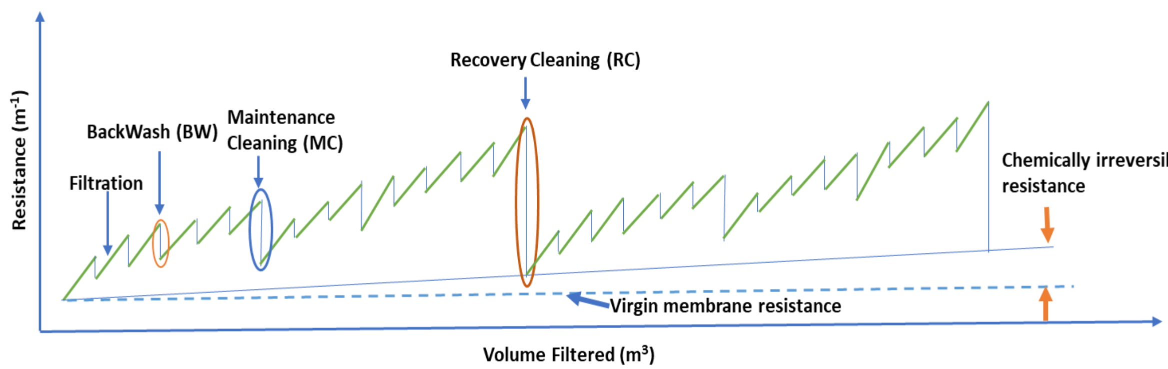


Figure 1: Resistance evolution in membrane during filtration and cleaning

Adapted from (Wang et al., 2014)

The frequencies at which cleanings are performed vary widely from facility to facility. In a survey of over 87 full-scale membrane facilities, Adham et al. (2005) reported BW frequencies of 0.2 - 5 per hour with a median value of 1.3 per hour, average MC frequency of 1 per week, and RC frequencies of 0.02 - 4.2 per month with a median value of 0.3 per month. Due to the dynamic nature of fouling, which changes with time, it is essential to estimate the appropriate frequency and relevant chemical dose required for BW, MC, and RC (Heo et al., 2022). Unfortunately, most full-scale membrane facilities follow manufacturer prescribed protocols for cleaning, which do not comprehensively consider the impact of relevant site-specific water characteristics or operational conditions (Heo et al., 2022).

2.3 Membrane Ageing

The long-term increase in chemically irreversible resistance, along with other changes in membrane performance and characteristics, is commonly referred to as 'membrane ageing'

(Robinson et al., 2016). The two most commonly used approaches to investigate the impacts of ageing on membranes are lab-accelerated ageing studies and laboratory characterization of field-harvested membranes (Robinson & Bérubé, 2020). Few studies have monitored the performance change at full-scale membrane facilities over time (Yu et al., 2020). Note that most studies focusing on membrane ageing considered resistance as a metric to quantify change in performance (Robinson et al., 2016). Robinson & Bérubé (2020) also proposed that resistance after extensive chemical cleaning could potentially be used to benchmark membrane ageing. For this reason, change in resistance with time (i.e., membrane age) is used in the present study as a metric of membrane ageing.

Previous lab-accelerated ageing studies, performed by soaking membranes in a chemical cleaning agent for a specified amount of time, suggested that resistance decreases with membrane age (Abdullah & Bérubé, 2013; Ren et al., 2021). In contrast, previous lab-accelerated ageing studies, performed using cyclic filtration and cleaning, suggested that resistance increases with membrane age (Hajibabania et al., 2012; He et al., 2014). Abdullah and Bérubé (2013) suggested that as membranes age, they become more hydrophobic, attracting more foulants and resulting in a greater rate of increase in resistance. They also reported that the extent of resistance reversed during chemical cleaning decreases as membranes age. Ren et al. (2021), on the other hand, reported no impact of ageing on the extent of resistance reversed during chemical cleaning, and a decrease in the extent of resistance reversed during physical cleaning with membrane age.

Results from field-harvested membrane studies are also inconsistent, likely due to the impact of site-specific water characteristics and operating conditions on membrane ageing. Robinson & Bérubé (2020) reported that resistance and rate of fouling remained unchanged for the initial five years of operation (i.e., age). The authors, however, reported a significant increase in resistance and rate of fouling after five years of operation. Touffet et al. (2015) reported that resistance decreases with membrane age, although they reported an increase in rate of fouling with membrane age.

In a full-scale membrane performance monitoring study, Yu et al., (2020) reported a progressive increase in resistance and rate of fouling over a 7 year time-period. As a result, more frequent cleaning was required to maintain a target flux. The authors, however, noted that the results from different full-scale facilities might produce different results due to impact of site-specific conditions.

Listed below is a summary of the impact of ageing on membrane performance factors reported in the literature.

1) Resistance was reported to either increase (Hajibabania et al., 2012; He et al., 2014) or decrease (Abdullah & Bérubé, 2013; Ren et al., 2021).
2) Rate of fouling was reported to increase (Abdullah & Bérubé, 2013; Touffet et al., 2015; Robinson & Bérubé, 2020; Yu et al., 2020).
3) Extent of resistance reversed during chemical cleaning was reported to either decrease (Abdullah & Bérubé, 2013) or remain unchanged (Ren et al., 2021).

Because lab-scale studies are not representative of full-scale membrane ageing (Filho et al., 2021), laboratory characterization of field-harvested membranes is time and resource intensive, and limited monitoring of ageing at full-scale facilities to date, the long-term performance of full-scale membrane facilities is not well understood (Yu et al., 2021).

2.4 Forecasting Membrane Replacement

In a survey of 106 full-scale membrane facilities treating drinking water, Chang et al. (2022) reported that membrane replacement age ranged from 5 to 10+ years; with 37% of the facilities replacing their membranes after approximately 7 years of operation (i.e., age) and 13% of the facilities replacing their membranes after more than 10 years of operation. Change in performance due to ageing was identified to be the primary reason for membrane replacement (Prulho et al., 2013). A number of change in performance triggers for replacing membranes have been reported; these include the followings.

- When resistance increase causes substantial flux decline such that the membrane system cannot meet the demand (Fenu et al., 2012);

- When warranty, provided by the manufacturers, in terms of maximum chemical dose is reached (Cote et al., 2012);
- When resistance can no longer be decreased using chemical cleaning (De Wilde et al., 2007);
- When resistance increase causes substantial increase in operating cost (Fenu et al., 2012);
- When any of the followings changes substantially: resistance, rate of fouling, breach frequency, infrastructure deterioration, and extent of resistance reversed during chemical cleaning (Robinson et al., 2016);
- When progressive increase in resistance creates the need for frequent chemical cleaning (Cote et al., 2012);
- When TMP increase becomes significantly greater than what has historically been observed (Woo et al., 2022).

While these triggers might be beneficial for better management of day-to-day operations, they do not provide insight to predict when a membrane would have to be replaced. Unfortunately, there is no universally accepted change in performance trigger for the replacement of membranes (Fenu et al., 2012).

Classical time series forecasting models are considered in the present study to inform membrane replacement. These models were selected because they are simple and can be easily interfaced with data analysis and control systems commonly available at full-scale membrane facilities. Many classical time series forecasting models such as linear forecasting, moving average, exponential smoothing, autoregressive integrated moving average (ARIMA) etc. are able to make reliable prediction into the future (Hyndman & Athanasopoulos, 2018). Of these classical models, exponential smoothing model is very popular due to its inherent simplicity, and ability to readily combine it with more sophisticated models (Vandeput, 2021).

A few studies have considered the use of forecasting tools to gain insight into membrane replacement age and/or performance. Ayala et al. (2011) performed a linear regression between the permeability decline and operating time of field harvested membranes from six full-scale membrane bioreactor (MBR) facilities. They utilized over six years of field harvested membranes data and estimated the membrane replacement age of approximately 7 years. Utilizing data from the previous one year, Teychene et al. (2018) attempted to predict the irreversible rate of fouling of the next six months using ARIMA time series forecasting model. However, the confidence intervals associated with the ARIMA forecast were very wide, suggesting that a large amount of data is required for a reliable forecast. For both studies, the forecasting timeline was very short (i.e., 6 months to 1 year). In addition, neither of the studies identified the minimum amount of data required to have a reliable forecast of membrane replacement; and nor did they propose a framework to effectively forecast membrane replacement using historical data from full-scale operation.

To the best knowledge of the author, the use of exponential smoothing model in forecasting membrane replacement has not yet been reported. In the present study, both linear and exponential smoothing model were used for forecasting membrane replacement. These two models are simple and can be easily implemented using a commonly available platform (e.g., 'Microsoft Excel').

Linear model involves fitting a linear regression line through an available dataset to estimate a slope and an intercept. The estimated slope and intercept are used to forecast beyond the data used for regression. Exponential smoothing model has three different variants - single, double and triple exponential smoothing (Vandeput, 2021). In the present study, double exponential smoothing (DES) model, otherwise known as Holt's linear trend, was used. In simple terms, the DES model involves estimating a level component, which is the weighted sum of past data, and a trend component, which is the weighted sum of the difference between past data. The estimated level and trend components are then used to forecast beyond the data used for estimating both components (Vandeput, 2021).

2.5 Research Gaps

The following research gaps were identified from the literature review:

1) Lab-accelerated ageing studies and laboratory characterization of field-harvested membranes can provide insight on membrane ageing mechanisms; however, they are either not representative of full-scale performance or time and resource intensive.

2) To date limited studies have been undertaken on performance monitoring of full-scale membrane facilities using historical data.
3) No study has considered the use of historical data from multiple full-scale membrane facilities to gain insight into membrane ageing.
4) No study has considered the use of forecasting tools, for forecasting timeline greater than 1 year, to provide guidance into when membranes should be replaced; and no study has identified how much historical data would be required to generate a relevant forecast.

3 Objectives of the Study

The key research objectives of the present study are listed below.

1) Investigate how performance factors change with membrane age for different facilities by analyzing historical data from full-scale membrane facilities.
2) Develop a framework to forecast membrane replacement age by utilizing historical data from full-scale membrane facilities.

4 Materials and Methods

4.1 Full-scale Data Analysis

4.1.1 Introduction

Four full-scale membrane facilities treating surface freshwater were considered for the present study. The name and location of these facilities are omitted to maintain anonymity. These facilities were selected because they used the same type of membranes for treating surface freshwater (ZeeWeed 1000, Veolia Water Technologies, Canada). Although these facilities use the same type of membranes, the specific membrane chemistry type, the characteristics of the raw water, the pretreatment approaches, the operating set points for permeation, as well as physical and chemical cleaning protocols, differed. A detailed description of the membrane chemistry type and operating protocols used at different facilities is provided in Appendix A.

Time series train-wise historical data from these four full-scale membrane facilities was acquired. To generate a manageable dataset, data from three randomly selected trains from each facility was analyzed. The type of historical data available for each facility differed; therefore, different approaches were used to obtain the required information. Refer to Appendix B for details of the acquired data. Note that data analysis was performed using Python programming language.

4.1.2 Calculation of Resistance

a) Facilities 1, 2, and 4

For facilities 1, 2, and 4, historical temperature and permeability data were available. Resistance (R) was calculated using Equation 3.

$$R = \frac{1}{\mu_T * B_T} \qquad \text{Equation 3}$$

where,

R = Resistance to permeate flow (m^{-1})

B_T = permeability at temperature T ($m^3.m^{-2}.hr^{-1}.bar^{-1}$)

μ_T = dynamic viscosity of water at temperature T (Pa.s)

μ_T in Equation 3 can be calculated from temperature using Equation 4.

$$\mu_T = 1.784 - (0.0575 * T) + (0.0011 * T^2) - (10^{-5} * T^3) \qquad \text{Equation 4}$$

where,

μ_T = dynamic viscosity of water at temperature T (cP)

T = Temperature of water (^{0}C)

b) Facility 3

For facility 3, historical permeate flux, TMP, and temperature data were available. Resistance (R) was calculated using Equation 5.

$$R = \frac{TMP}{J_T * \mu_T} \qquad \text{Equation 5}$$

where,

R = Resistance to permeate flow (m^{-1})

J_T = Flux at temperature T ($m^3.m^{-2}.hr^{-1}$)

TMP = Trans-membrane pressure at temperature T (Nm^{-2})

μ_T = dynamic viscosity of water at temperature T (Pa.s)

4.1.3 Occurrences of BW, MC and RC

I. BW

a) Facilities 1 and 3

For facilities 1 and 3, every minute permeate flow, TMP, and tank water level data were available. BW was considered to have occurred when the following three conditions were met within 5 minutes. A 5 minute period was selected because all relevant conditions associated with a BW occurred within 5 minutes.

1) A tank drain.
2) An interruption in permeate flow.
3) A transition from a negative pressure to a positive pressure in the permeate line.

b) Facilities 2 and 4

For facilities 2 and 4, a BW generated a distinct data tag, which was used to identify the occurrence of a BW.

II. MC and RC

a) Facilities 1 and 3

For facilities 1 and 3, an MC and RC were considered to have occurred when the following four conditions were met within the corresponding reported duration of an MC (approximately 8 – 60 minutes) and RC (> 100 minutes), respectively.

1) Two tank drains.
2) An interruption in permeate flow between the tank drains.
3) A transition from a negative pressure to a positive pressure in the permeate line.
4) Chemical soak between the tank drains (depending on the duration of soak length MC and RC were distinguished).

For facility 1, the chemical cleaning agents used for an MC and RC were identified from the chemical flow data tags during an MC and RC, respectively. While for facility 3, the chemical cleaning agent used for an MC and RC were identified from the operator's log.

b) Facility 2

For facility 2, an MC and RC generated distinct data tags which were used to identify the occurrence of an MC and RC; and the chemical cleaning agents used during an MC and RC, respectively.

c) Facility 4

For facility 4, occurrences of MC were not reported in the operator's log, nor could they be identified from the available data. RC dates and the chemical cleaning agents used were identified from the operator's log. However, the exact time of the RC was not reported. The exact time of an RC was identified when the following three conditions were met within the reported duration of an RC.

1) An interruption in permeate flow.
2) A transition from a negative pressure to a positive pressure in the permeate line.
3) A high resistance drop (i.e., > 5×10^{10} m^{-1}, which was the lowest resistance drop on the dates of the reported RC).

4.1.4 Rate of Change in Resistance

I. Between BW

a) Facilities 1 and 3

For facilities 1 and 3, every minute historical resistance data was available. Typical change in resistance between BW is illustrated in Figure 2.

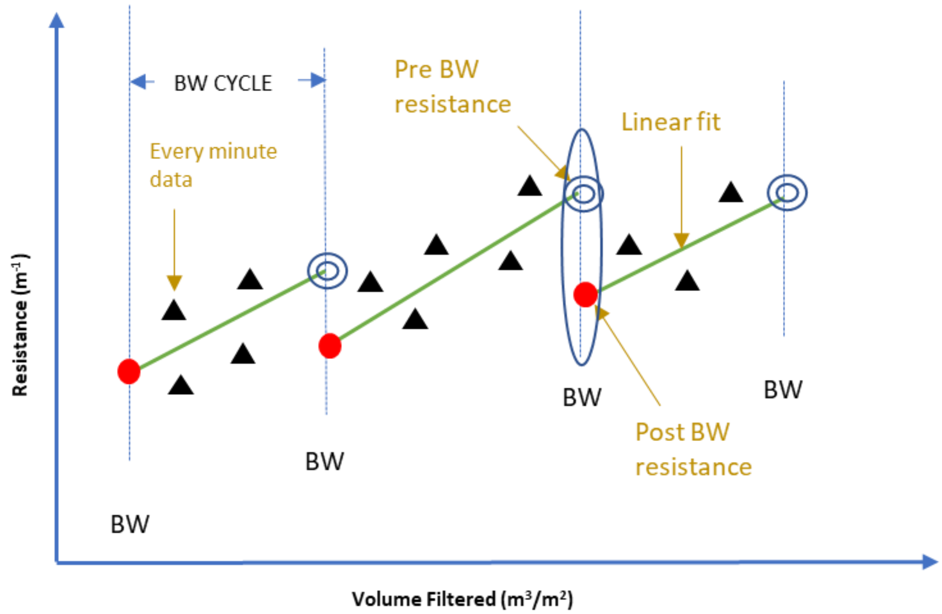


Figure 2: Change in resistance between BW

Triangular symbols correspond to resistance; circular solid and open symbols correspond to the post and pre BW resistance estimated with the intercepts of the linear regression model; the solid lines correspond to the linear regression model fitted to the resistance measurements between two consecutive BW; dashed lines correspond to the occurrences of BW.

From the resistance data between BW, the parameters of interests were:

1) Pre and post BW resistance.

2) Rate of change in resistance between BW.

A linear regression model was applied to every minute resistance data between BW (see Figure 2). The intercepts at times corresponding to BW_n and BW_{n+1} were used as the estimates of pre and post BW resistance (R_{preBW}, R_{postBW}), respectively, and the slope as an estimate of rate of change in resistance between BW. Note that between 2 successive BW, the permeate flow was periodically interrupted due to the low demand. When this occurred, the analysis was performed by excluding the time period when the permeate flow was interrupted.

Also, note that because the permeate flow immediately pre and post BW can be variable for up to 2 minutes (because of pump starts and stops), data for the first two minutes immediately pre and post BW was excluded from the regression analysis. Resistance value that was not within the 95% prediction interval of the fitted linear regression model was considered to be an outlier and was excluded from the analysis. Linear regression was again performed without the excluded data.

A change in resistance between BW that could not be accurately modeled using a linear regression model was excluded from the analysis. The accuracy of the model was quantified using R^2 of the linear regression model. The R^2 "cutoff" was defined as equivalent to a 90% confidence interval of z-values (Equation 6), estimated from the R^2 values as described by Fisher (1921). In total, approximately 15% of the change in resistance between BW were excluded from the analysis.

$$Z = \frac{1}{2} \ln\left(\frac{1 + \sqrt{R^2}}{1 - \sqrt{R^2}}\right) \qquad \text{Equation 6}$$

b) Facilities 2 and 4

For facilities 2 and 4, historical resistance data was only available immediately pre and post BW. For this reason, these two measurements were used as the estimates of pre and post BW resistance (R_{preBW} and R_{postBW}), respectively. Note that because the permeate flow immediately pre and post BW can be variable, the rate of change in resistance between BW was not estimated for facilities 2 and 4.

II. Between MC

The resistance post BW (R_{postBW}) was used to assess the rate of change in resistance between MC. For all facilities, a number of relationships were considered to describe the change in R_{postBW} between MC. However, none could accurately model the change in R_{postBW} between MC.

The accuracy of the model was quantified using R^2 of each relationship. Because of the high variability in the rate of change in R_{postBW} between MC, the impact of MC on resistance was not further considered. Thus, hereafter, chemical cleaning only refers to RC.

III. Between RC

a) Facilities 1, 2, and 3

The resistance post BW (R_{postBW}) was also used to assess the change in resistance between RC. Typical change in R_{postBW} between RC is illustrated in Figure 3.

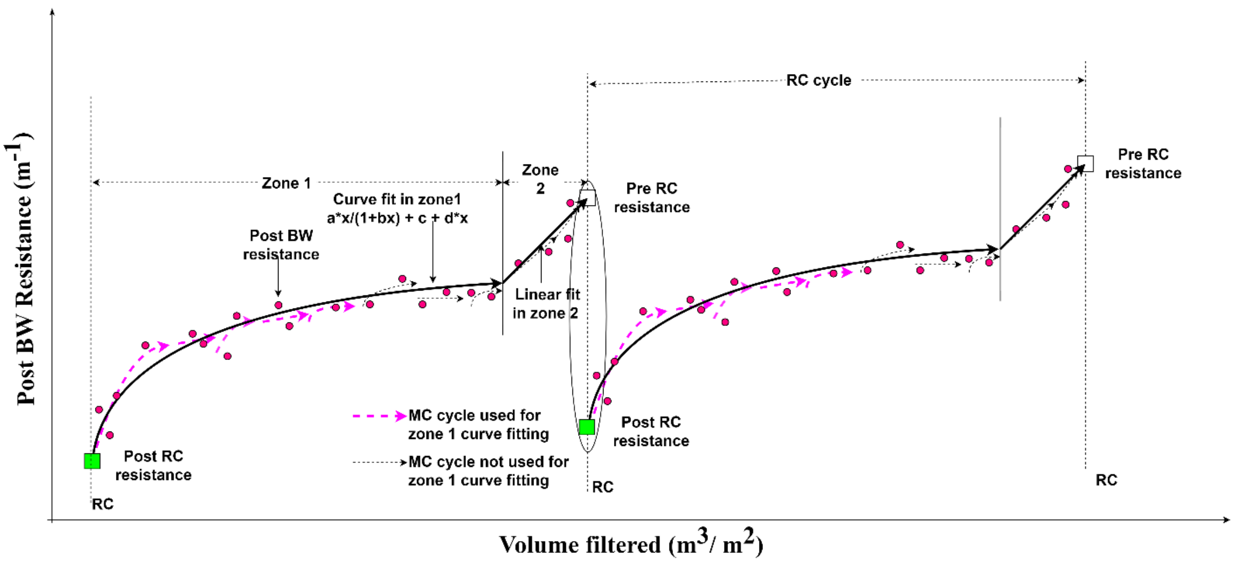


Figure 3: Change in post BW resistance between RC

Circular solid symbols correspond to post BW resistance; unfilled rectangular symbols correspond to the pre RC resistance; filled rectangular symbols correspond to the post RC resistance; dashed vertical lines correspond to the occurrences of RC; dashed pink lines correspond to the mc cycle used for fitting Equation 7 in zone 1; dashed black lines correspond to mc cycle not used for fitting Equation 7; curved solid lines correspond to Equation 7 fitted to the R_{postBW} data in zone 1; linear solid lines correspond to the linear regression model fitted to the R_{postBW} data in zone 2.

From the R_{postBW} data between RC, the parameters of interests were:

1) Pre and post RC resistance.
2) Rate of change in R_{postBW} between RC.

These parameters were estimated using a combination non-linear regression (i.e., Equation 7) and linear regression. The change in R_{postBW} between could be characterized by an initial rapid increase followed by a slower gradual increase in R_{postBW}. In Figure 3, this behavior is illustrated in zone 1. When the interval between RC was relatively long (approximately 10% of the interval between RC), the change in R_{postBW} was observed to again increase rapidly. In Figure 3, this later behavior is illustrated in zone 2.

1) Zone 1: Change in R_{postBW}, in zone 1, could be modeled using non-linear regression relationship presented in Equation 7.

$$y = \frac{a * x}{1 + b * x} + c * x + d \qquad \text{Equation 7}$$

where,

x = volume filtered within each RC cycle/ m^2 membrane area

y = post backwash resistance (m^{-1})

a, b, c and d = empirically estimated parameters

Equation 7 was fitted to the R_{postBW} data corresponding to the first i^{th} number of MC after an RC; and extended to the next RC. The number i was selected to be equal to the minimum number of MC below which zone 2 behavior was not observed. Provided that zone 2 behavior was not observed, intercepts at times corresponding to RC_n and RC_{n+1} were extracted as the estimates of pre and post RC resistance (R_{preRC} and R_{postRC}), respectively, and the rate of change in R_{postBW} between RC was estimated using the coefficients a, b, and c from Equation 7..

Note that R_{postBW} having error outside of 2 standard deviations of errors (actual – modeled R_{postBW}) was considered to be an outlier and excluded from the analysis. Non-linear regression analysis, using Equation 7, was again performed without the excluded data.

Change in R_{postBW} between RC (i.e., RC_n, RC_{n+1}) that could not be accurately modeled using Equation 7 was excluded from the analysis. Because RC is a relatively infrequent cleaning approach compared to BW, the change in R_{postBW} between RC for which Equation 7 could not be accurately fitted were excluded from the analysis. This was identified by the Python 'scipy.curve_fit()' function returning R^2 values that were not greater than zero. Note that less than 5% of the change in R_{postBW} between RC were excluded from the analysis.

2) Zone 2: Start of zone 2 was identified when the actual R_{postBW} remained consistently below or above Equation 7 fitted to the data, for the duration of at least two consecutive MC. In zone 2, the change in R_{postBW} could be described using a linear regression model. Thus, when zone 2 behavior occurred, the change in R_{postBW} between RC was modelled with Equation 7 and a linear regression model as illustrated in Figure 3.

The intercepts at times corresponding to RC_n and RC_{n+1} were extracted as the estimates of pre and post RC resistance (R_{preRC}, R_{postRC}), from zone 2 and zone 1 respectively, and the rate of change in R_{postBW} between RC (RC_n, RC_{n+1}) was quantified using the coefficients a, b, and c from Equation 7 (estimated from zone 1) and the slope of linear regression model (estimated from zone 2) (see Figure 3).

Note that for zone 2, R_{postBW} values that were not within the 95% prediction interval of the fitted linear regression model was considered to be an outlier and were excluded from the analysis. Linear regression was again performed without the excluded data.

Change in R_{postBW} in zone 2 that could not be accurately modeled using a linear regression was excluded from the analysis. Similar to zone 1, this was identified by the Python 'linregress()' function returning R^2 values that were not greater than zero. Note that no change in R_{postBW} in zone 2 were excluded from the analysis for this.

b) Facility 4

For facility 4, zone 2 could not be identified because the occurrence of MC could not be identified. Change in R_{postBW} between RC was modeled using Equation 7. Estimation of change in R_{postBW} between RC, and extraction of R_{preRC} and R_{postRC} were performed according to the discussion presented for zone 1 above. Note that zone 2 behavior was not observed to occur for facility 4.

4.1.5 Quantification of Performance Factors

Of the performance factors that have been reported to change with membrane age (see section 2.3), resistance, rate of fouling, and extent of resistance reversed during cleaning events were considered in the present study. For resistance both hydraulically irreversible resistance following BW and clean membrane resistance following RC were considered. For rate of fouling, rate of change in resistance between BW, defined as total fouling rate, was considered. For the extent of resistance reversed during cleaning events, resistance reversed during both BW, and RC were considered. A summary of performance factors examined for the present study is presented in Table 1.

Table 1: Performace factors examined from full-sale data

Performance factors	**Performance factors quantified based on***
Hydraulically irreversible resistance	Post BW resistance (R_{postBW})
Clean membrane resistance[1]	Post RC resistance (R_{postRC})
Total fouling rate	Slope of the linear regression model fitted to the resistance (R) data between BW
Extent of resistance reversed during BW	Difference between the resistance pre and post BW ($R_{preBW} - R_{postBW}$)
Extent of resistance reversed during RC	Difference between the resistance pre and post RC ($R_{preRC} - R_{postRC}$)

* see section 4.1.4 for the estimation of each of these parameters.

Note that for all performance factors, data outside of 2 standard deviations of the raw data was considered as outlier and excluded from the analysis. For train 2 at facility 3, at the end of 2017, an abrupt drop in resistance was observed, without any occurrence of cleaning event during that period. For this reason, any data before 2018 for that train was excluded from the analysis. For facility 1, at startup, the initial data had higher variability. For this reason, for all trains at facility 1, data corresponding to cumulative permeate volume less than 50 m^3/m^2 was excluded from the analysis.

[1] In the present study, clean membrane refers to the membrane after RC, even though membranes might not be fully cleaned after an RC due to the accumulation of chemically irreversible resistance.

4.2 Forecasting Framework

4.2.1 Introduction

Forecasting was performed using clean membrane resistance data (R_{postRC}). Two time series forecasting models were considered: linear, and DES. For both models, data was divided into two sets – training data and test data. The percentage of all available data considered for training ranged from 5% to 60% at 5% increments to determine the minimum amount of data required for a reliable forecast. The root mean square error (RMSE) between the forecasted R_{postRC} and the actual R_{postRC}, for the test data, was used as a metric of accuracy of the forecast.

Membranes were assumed to have to be replaced when forecasted R_{postRC} was increased to a value corresponding to a 70% decrease in the permeability of a virgin (i.e., new) membrane. A 70% reduction in permeability is currently being used by managers of facility 4 as a trigger for membrane replacement. Note that as the initial few years of data for train 2 at facility 3 was excluded from the analysis (see section 4.1.5), forecasting was not performed for that train. Also, for facility 1, initial R_{postRC} data corresponding to cumulative permeate volume less than 50 m^3/m^2 was excluded from the forecasting (see section 4.1.5). Note that forecasting was performed using Microsoft Excel-2019.

4.2.2 Data Preprocessing for Forecasting

At full-scale facilities, RC is not performed exactly at the same time interval. However, for the DES model, data (both training data and test data) is required to have equal time interval (Wright, 1986). Also, the DES model is more effective without any significant outliers in the training data (Ahmad & Ahmad, 2013). For these reasons, R_{postRC} data needed preprocessing before forecasting could be performed. To have consistent data, for both models, data was preprocessed in the same manner using the following two steps:

1. **Equalizing the temporal spacing of the data:** The training and test data was modified as needed such that the time interval between consecutive data was constant. This was achieved by assigning each R_{postRC} to the first day of the respective month so that the time interval between each R_{postRC} was one month. If a certain month had no R_{postRC}, an equivalent value was interpolated, assuming linear regression through the training and

test data. If a certain month had more than one R_{postRC}, the average of those R_{postRC} was used and assigned at the first day of the respective month.

2. **Replacing outliers in the training data**: Outliers in the training data were identified and replaced by an interpolated value, assuming linear regression through the training data. Outliers were identified as the data lying outside of the 95% prediction interval of that regression line. Note that outliers were replaced only in the training data as the model was trained in the training data only.

4.2.3 Forecasting Using Linear Model

For the linear model, linear regression was performed using the training data following the procedure outlined in Berthouex & Brown (2002) to estimate relevant summative time series parameters (i.e., slope and intercept). For the linear forecast, the linear regression performed in the training data was simply extrapolated to the test data and beyond using the estimated slope and intercept.

The prediction interval for the linear model was constructed using Equation 8 to quantify the uncertainties associated with the forecast, as outlined in Berthouex & Brown (2002).

$$Y_t \pm t_{crit} * s * \sqrt{1 + \frac{1}{n} + \frac{(x_t - \overline{x})^2}{\sum_{i=0}^{n}(x_i - \overline{x})^2}} \quad \text{Equation 8}$$

where,

Y_t = linear forecasted R_{postRC} at time t

t_{crit} = critical t value corresponding to 95% prediction interval

s = variance of the training data

n = number of observations in the training data

x_t = membrane age at time t

$\overline{x}$ = mean membrane age of the training data

4.2.4 Forecasting Using DES Model

For the DES model, the procedure outlined in (Vandeput, 2021) was used. Briefly, each training data is associated with a level (L_t) and a trend (b_t) component. These components were estimated using Equation 9 and Equation 10, respectively.

$$L_t = \alpha * y_t + (1-\alpha) * (L_{t-1} + b_{t-1}) \qquad \text{Equation 9}$$

$$b_t = \beta * (L_t - L_{t-1}) + (1-\beta) * (b_{t-1}) \qquad \text{Equation 10}$$

where,

L_t = estimated level component at time t

L_{t-1} = estimated level component at time t-1

b_t = estimated trend component at time t

b_{t-1} = estimated trend component at time t-1

α & β = smoothing parameters with a range of $0 \leq \alpha$ & $\beta \leq 1$

y_t = R_{postRC} at time t

Level (L_t) and trend (b_t) component for the first training data were initialized as the first R_{postRC} and the difference between first two R_{postRC}, respectively. Note that the trend component can be both positive and negative in Equation 10. The unknowns in Equation 9 & Equation 10 are the smoothing parameters (i.e., α & β), which were estimated by minimizing the RMSE between the actual R_{postRC} and forecasted R_{postRC} using the "Solver" function in Microsoft Excel-2019, in the training data. The Forecasted R_{postRC} for the training data was obtained using Equation 11.

$$F_t = L_t + b_t \qquad \text{Equation 11}$$

where,

F_t = forecasted R_{postRC} at time t

L_t = estimated level component at time t

b_t = estimated trend component at time t

After α & β were estimated (i.e., model was trained), the forecast could be extrapolated into the test data and beyond using Equation 12.

$$F_{t+h|t} = L_{last} + h * b_{average} \qquad \text{Equation 12}$$

where,

h = h-step ahead forecast (i.e., 1, 2, 3…….)

$F_{t+h|t}$ = h step ahead forecasted R_{postRC}

L_{last} = estimated level component of the last training data

$b_{average}$ = average of all positive trend component of the training data[2]

The prediction interval for the DES model was constructed using Equation 13 to quantify the uncertainties associated with the forecast, as outlined in Hyndman & Athanasopoulos (2018).

$$F_{t+h|t} \pm c * \sigma_h \qquad \text{Equation 13}$$

where,

c = 1.96 for 95% prediction interval

σ_h = standard deviation of the forecast

σ_h was calculated using Equation 14.

$$\sigma_h^2 = \sigma^2[1 + (h-1)\left\{\alpha^2 + \alpha\beta h + \frac{\beta^2 h(2h-1)}{6}\right\}] \qquad \text{Equation 14}$$

where,

σ^2 = variance of the training data

[2] If there was no positive trend component in the training data, the absolute average trend component of all training data was used for forecasting.

5 Results and Discussion

5.1 Change in Performance Factors

For all the performance factors considered in the present study (see section 4.1.5), the followings are presented:

a) typical raw time series data and typical average time series data; and
b) the rate of change of the estimated time series data for 3 trains at all facilities.

Raw time series data provides qualitative insight into the variability of the time series data; while estimated time series data, obtained by fitting a linear regression model to the raw data, provides qualitative insight into the overall trend of the time series data. The rate of change of the estimated time series data, obtained from the slope of the linear regression model fitted to the raw data, provides quantitative information on the overall trend of the time series data. Considering results for all facilities in parallel provides insight into the impact of site-specific water characteristics and operational conditions, while considering the results from 3 trains at each facility provides insight into the variability in time series data within a particular facility. Also, for each train at each facility, the coefficient of determination (R^2) provides a measure of how representative the estimated time series data is of the raw time series data.

When relevant, typical performance data is presented. For all performance factors considered, the results are presented with respect to both cumulative filtered and membrane age.

5.1.1 Hydraulically Irreversible Resistance

Typical hydraulically irreversible resistance, quantified based on the post the BW resistance, is presented in Figure 4.

(a) (b)

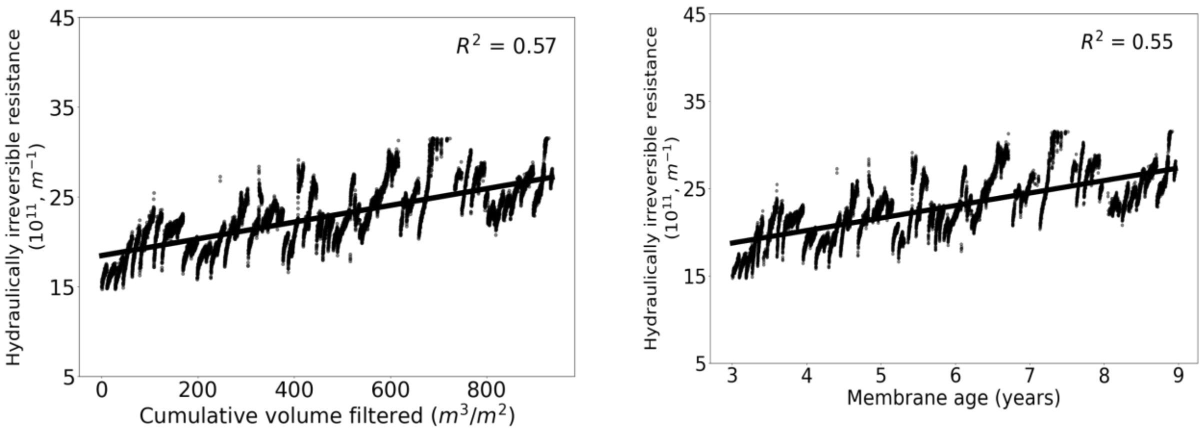


Figure 4: Typical hydraulically irreversible resistance

(a) with respect to cumulative volume filtered; (b) with respect to membrane age. Solid line corresponds to linear regression model fitted to the raw data. 95% confidence interval of the linear regression model is included but cannot be viewed because it overlaps with the solid line. Results presented for train 1 at facility 3.

Overall, the hydraulically irreversible resistance increases with respect to both cumulative volume filtered and membrane age. The abrupt changes in raw hydraulically irreversible resistance data coincide with the occurrences of RC; suggesting that hydraulically irreversible resistance is impacted by the operational conditions of the facilities. The rate of change of the estimated hydraulically irreversible resistance, for three trains at each facility, is presented in Figure 5. Details relating to the fit and slope of the linear regression model to the raw data is presented in Appendix C Table C1.

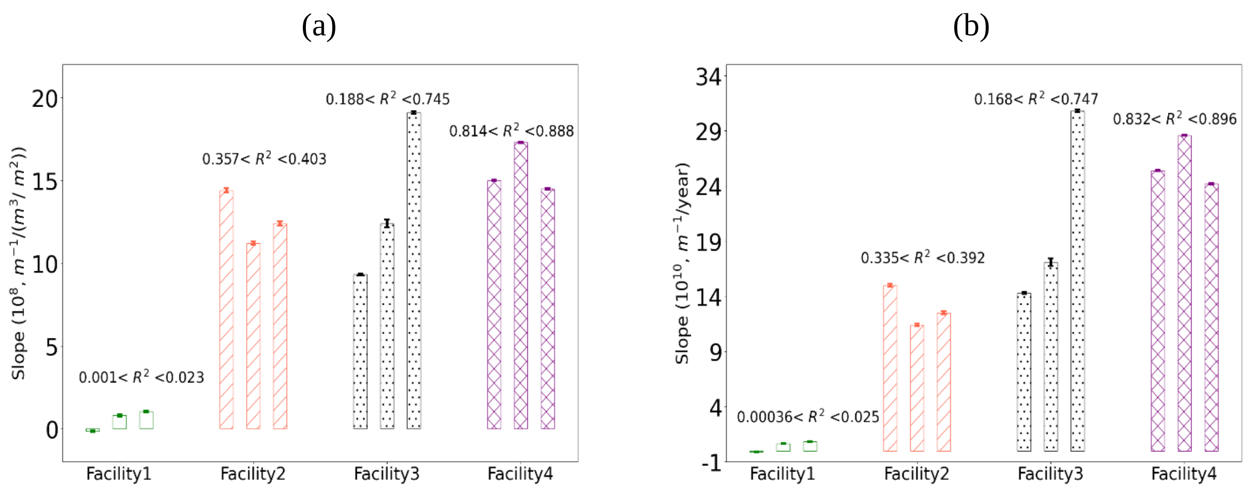


Figure 5: Rate of change of estimated hydraulically irreversible resistance

(a) with respect to cumulative volume filtered; (b) with respect to membrane age. Note that the scale of y axis in both figures are different as the magnitude of change with respect to cumulative volume filtered is different from that of membrane age. Error bar indicates standard error of the estimated slope. Results presented for 3 randomly selected trains at each facility.

For facilities 2, 3, and 4, the rate of change of the estimated hydraulically irreversible resistance is significantly greater than 0 (zero) ($p < 0.05$), indicating substantive membrane ageing, based on this metric, at these facilities. For facility 1, although the rate of change is significantly greater than 0 (zero) ($p < 0.05$) for trains 2 and 3, the magnitudes of the rate of change are relatively low; while for train 1, the rate of change is negative. The relatively low and/or negative rate of change of the estimated hydraulically irreversible resistance suggests limited membrane ageing, based on this metric, at facility 1.

The coefficient of determination (R^2) was substantively different for all facilities. The coefficient of determination was consistently greater than 0.7 for facility 4, approximately ranged from 0.2 – 0.7 for facility 3, approximately 0.4 for facility 2, and less than 0.1 for facility 1. A coefficient of determination of greater than 0.7, range from 0.2 – 0.7, and less than 0.1, indicate that the estimated hydraulically irreversible resistance represents the raw hydraulically irreversible resistance data very well, moderately well, and poorly, respectively. Note that when the rate of change is low, the coefficient of determination is inherently low (Miaou et al., 1996).

In summary, the hydraulically irreversible resistance substantively increased with membrane age for 3 of the 4 facilities. For these facilities, the rate of change of the estimated hydraulically irreversible resistance could generally be approximated with a linear relationship moderately well to very well. For one of the facilities (i.e., facility 1), the hydraulically irreversible resistance was not observed to substantively increase with membrane age.

5.1.2 Clean Membrane Resistance

Typical clean membrane resistance, quantified based on the post RC resistance, is presented in Figure 6.

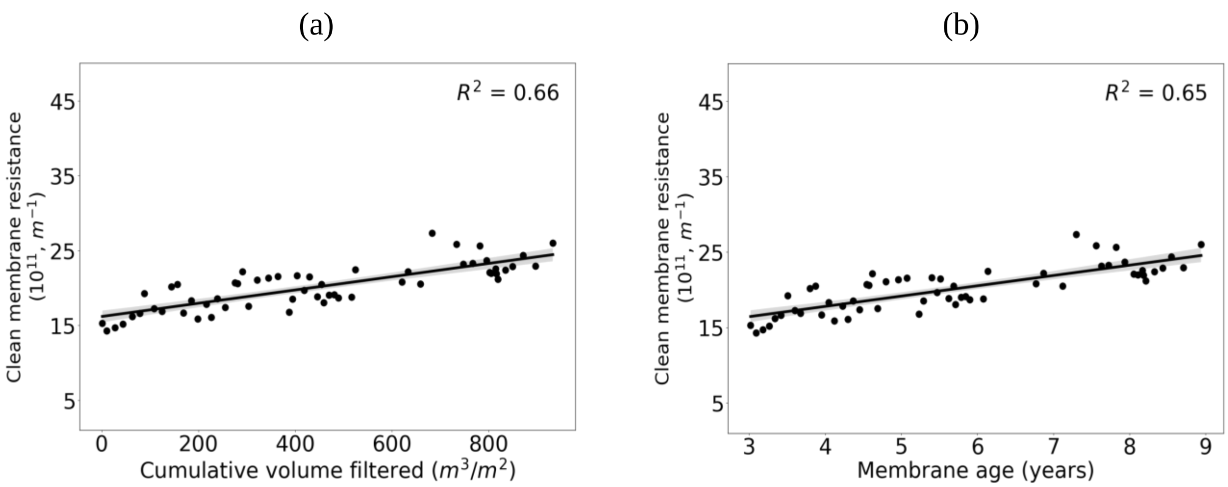


Figure 6: Typical clean membrane resistance

(a) with respect to cumulative volume filtered; (b) with respect to membrane age. Solid line corresponds to linear regression model fitted to the raw data. 95% confidence interval of the linear regression model is indicated by the light black shade. Results presented for train 1 at facility 3.

Overall, the clean membrane resistance increases with respect to both cumulative volume filtered and membrane age. Unlike hydraulically irreversible resistance (see section 5.1.1), there is no abrupt change in raw clean membrane resistance data; suggesting that the clean membrane resistance is relatively unimpacted by the operational conditions of the facilities. The rate of change of the estimated clean membrane resistance, for three trains at each facility, is presented in Figure 7. Details relating to the fit and slope of the linear regression model to the raw data is presented in Appendix C Table C2.

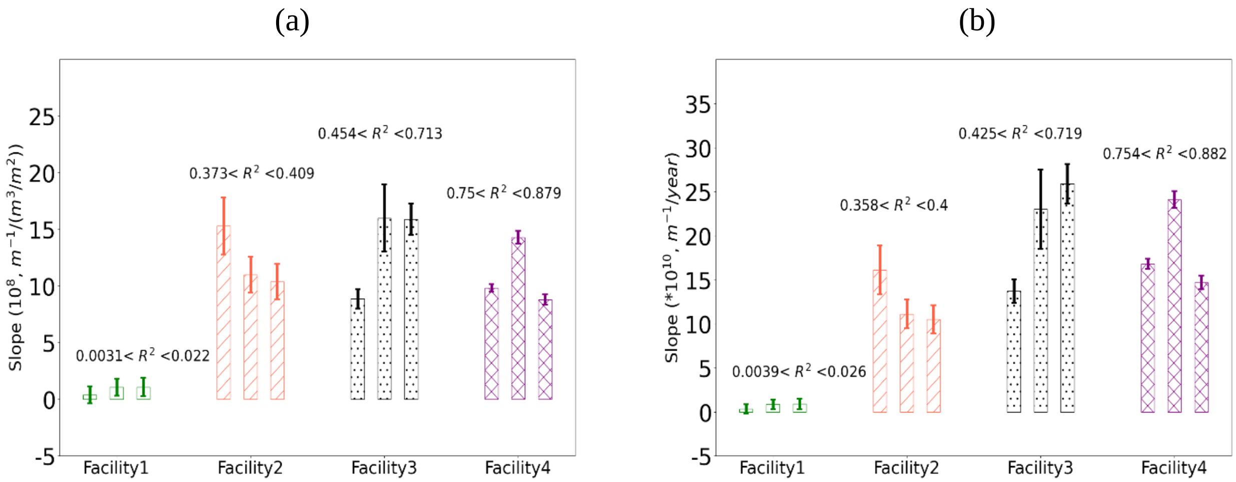


Figure 7: Rate of change of estimated clean membrane resistance

(a) with respect to cumulative volume filtered; (b) with respect to membrane age. Note that the scale of y axis in both figures are different as the magnitude of change with respect to cumulative volume filtered is different from that of membrane age. Error bar indicates standard error of the estimated slope. Results presented for 3 randomly selected trains at each facility.

For facilities 2, 3, and 4, the rate of change of the estimated clean membrane resistance is significantly greater than 0 (zero) ($p < 0.05$), indicating substantive membrane ageing, based on this metric, at these facilities. For facility 1, the rate change of the estimated clean membrane resistance is not significantly greater than 0 (zero) ($p > 0.05$) and the magnitudes of the rate of change are relatively low. The relatively low and insignificant rate of change of the estimated clean membrane resistance suggests limited membrane ageing, based on this metric, at facility 1.

The coefficient of determination (R^2) was substantively different for all facilities. The coefficient was consistently greater than 0.7 for facility 4, approximately ranged from 0.2 – 0.7 for facility 3, approximately 0.4 for facility 2, and less than 0.1 for facility 1. A coefficient of determination of greater than 0.7, range from 0.2 – 0.7, and less than 0.1 indicate that the estimated clean membrane resistance represents the raw clean membrane resistance data very well, moderately well, and poorly, respectively. Note that when the rate of change is low, the coefficient of determination is inherently low (Miaou et al., 1996).

In summary, the clean membrane resistance substantively increased with membrane age for 3 of the 4 facilities. For these facilities, the rate of change of the estimated clean membrane resistance could generally be approximated with a linear relationship moderately well to very well. For one of the facilities (i.e., facility 1), the clean membrane resistance was not observed to substantively increase with membrane age.

5.1.3 Total Fouling Rate

Typical total fouling rate, quantified based on the slope of the linear regression model fitted to the resistance data between BW, is presented in Figure 8.

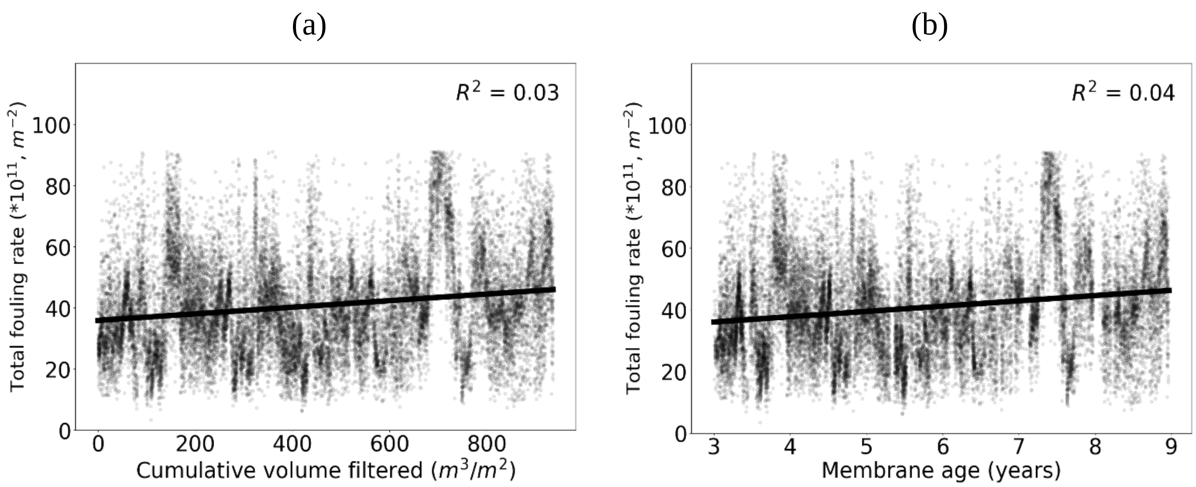


Figure 8: Typical total fouling rate

(a) with respect to cumulative volume filtered; (b) with respect to membrane age. Solid line corresponds to linear regression model fitted to the raw data. 95% confidence interval of the linear regression model is included but cannot be viewed because it overlaps with the solid line. Results presented for train 1 at facility 3.

Overall, the total fouling rate increases with respect to both cumulative volume filtered and membrane age. Raw total fouling rate data is highly scattered; suggesting that the total fouling rate is impacted by the operational conditions of the facilities. The rate of change of the estimated total fouling rate, for three trains at facilities 1 and 3, is presented in Figure 9. Note that for facilities 2 and 4, total fouling rate could not be estimated (see section 4.1.4). Details relating to the fit and slope of the linear regression model to the raw data is presented in Appendix C Table C3.

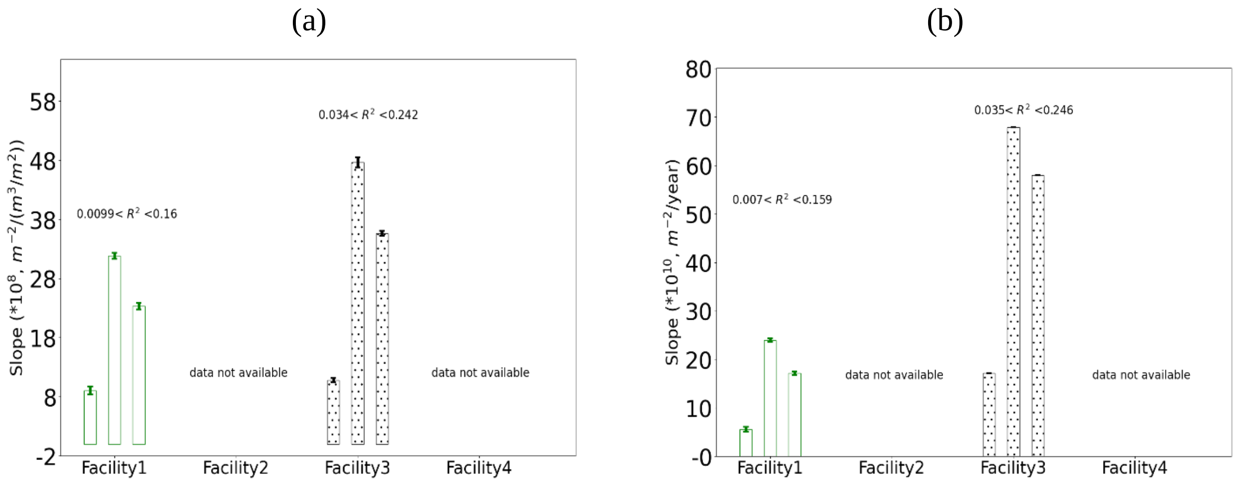


Figure 9: Rate of change of estimated total fouling rate

(a) with respect to cumulative volume filtered; (b) with respect to membrane age. Note that the scale of y axis in both figures are different as the magnitude of change with respect to cumulative volume filtered is different from that of membrane age. Error bar indicates standard error of the estimated slope. Results presented for 3 randomly selected trains at facilities 1 and 3.

For facilities 1 and 3, the rate of change of the estimated total fouling rate is significantly greater than 0 (zero) ($p < 0.05$), indicating substantive membrane ageing, based on this metric, at these facilities. For facility 1, on average, the magnitude of the rate of change is relatively lower than for facility 3. The relatively lower increase in the rate of change of the estimated total fouling rate suggests relatively lower membrane ageing, based on this metric, at facility 1.

The coefficient of determination (R^2) was approximately similar for both facilities. The coefficient of determination was approximately 0.1 for both facilities. A coefficient of determination of 0.1 indicates that the estimated total fouling rate represents the raw total fouling rate data poorly. Note that when the rate of change is low, the coefficient of determination is inherently low (Miaou et al., 1996).

In summary, the total fouling rate increased substantively with membrane age for facilities 1 and 3. For both facilities, the rate of change of the estimated total fouling rate could generally be approximated with a linear relationship poorly. For facilities 2 and 4, total fouling rate could not be quantified (see section 4.1.4).

5.1.4 Extent of Resistance Reversed during BW

Typical extent of resistance reversed during BW, quantified based on the difference between resistance pre and post BW ($R_{preBW} - R_{postBW}$), is presented in Figure 10. Note that the extent of resistance reversed during BW was normalized with respect to the volume of permeate filtered before each BW; to account for the fact that the extent of resistance incurred between BW is expected to be proportional to the volume of permeate filtered.

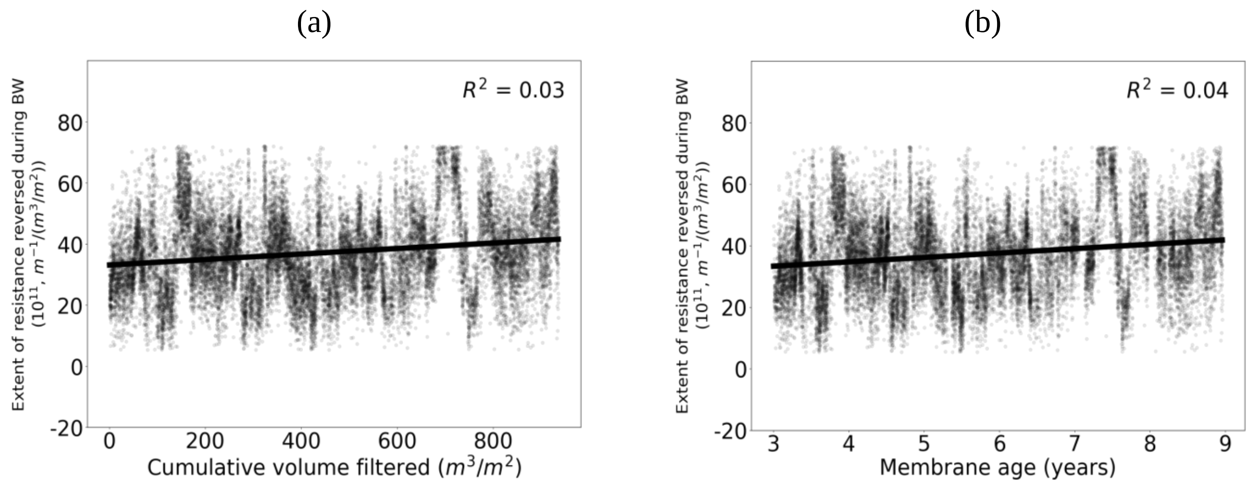


Figure 10: Typical extent of resistance reversed during BW

(a) with respect to cumulative volume filtered; (b) with respect to membrane age. Solid line corresponds to linear regression model fitted to the raw data. 95% confidence interval of the linear regression model is included but cannot be viewed because it overlaps with the solid line. Results presented for train 1 at facility 3. Note that the extent of resistance reversed during BW was normalized with respect to volume of permeate filtered before each BW.

Overall, the extent of resistance reversed during BW increases with respect to both cumulative volume filtered and membrane age. Similar to the total fouling rate (see section 5.1.3), raw extent of resistance reversed during BW data is highly scattered; suggesting that the extent of resistance reversed during BW is also impacted by the operational conditions of the facilities. The rate of change of the estimated extent of resistance reversed during BW, for three trains at each facility, is presented in Figure 11. Details relating to the fit and slope of the linear regression model to the raw data is presented Appendix C Table C4.

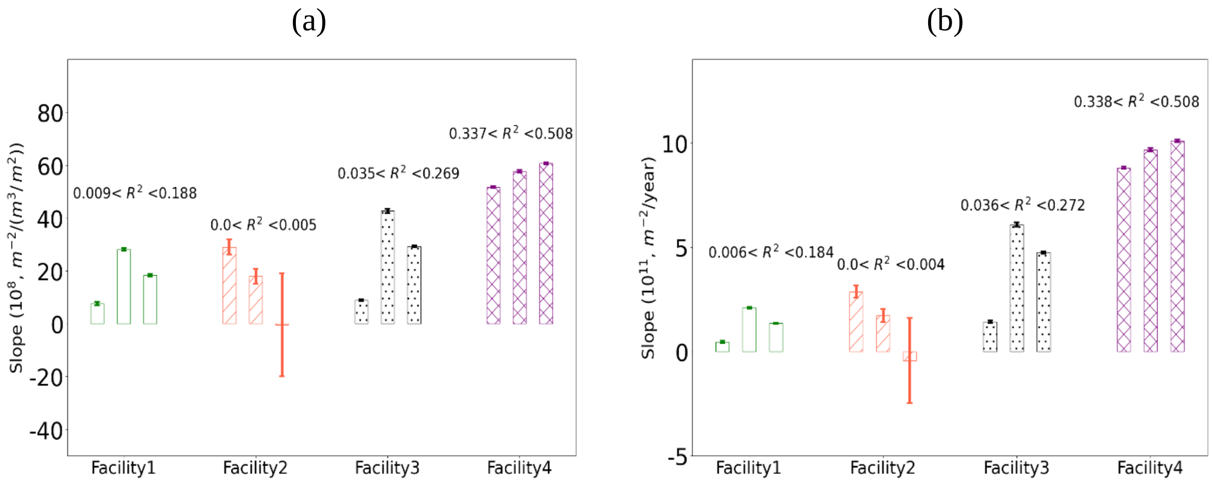


Figure 11: Rate of change of estimated extent of resistance reversed during BW

(a) with respect to cumulative volume filtered; (b) with respect to membrane age. Note that the scale of y axis in both figures are different as the magnitude of change with respect to cumulative volume filtered is different from that of age. Error bar indicates standard error of the estimated slope. Results presented for 3 randomly selected trains at each facility.

For facility 4, the rate of change of the estimated extent of resistance reversed during BW is significantly greater than 0 (zero) ($p < 0.05$), indicating substantive membrane ageing, based on this metric, at this facility. For facilities 1, 2, and 3, although the rate of change of the estimated extent of resistance reversed during BW is significantly greater than 0 (zero) ($p < 0.05$) (except train 3 at facility 2), the magnitudes of the rate of change are relatively lower than for facility 4. For train 3 at facility 2, the rate of change is relatively low and negative, and associated with relatively high error. The relatively lower increase in the extent of resistance reversed during BW suggests relatively lower membrane ageing, based on this metric, at facilities 1, 2 and 3 than that of at facility 4.

The coefficient of determination (R^2) was substantively different for all facilities. The coefficient of determination ranged from 0.2 – 0.7 for facility 4, approximately 0.1 for facilities 1 and 3, and consistently less than 0.1 for facility 2. A coefficient of determination of 0.2 – 0.7, approximately 0.1 to less than 0.1 indicate that the estimated extent of resistance reversed during BW represents the raw extent of resistance reversed during BW data moderately well, and poorly, respectively.

Note that when the rate of change is low, the coefficient of determination is inherently low (Miaou et al., 1996).

In summary, the extent of resistance reversed during BW increased with membrane age for all four facilities. For these facilities, the rate of change of the estimated extent of resistance reversed during BW could generally be approximated with a linear relationship poor to moderately well.

5.1.5 Extent of Resistance Reversed during RC

Typical extent of resistance reversed during RC, quantified based on the difference between resistance pre and post RC ($R_{preRC} – R_{postRC}$), is presented in Figure 12.

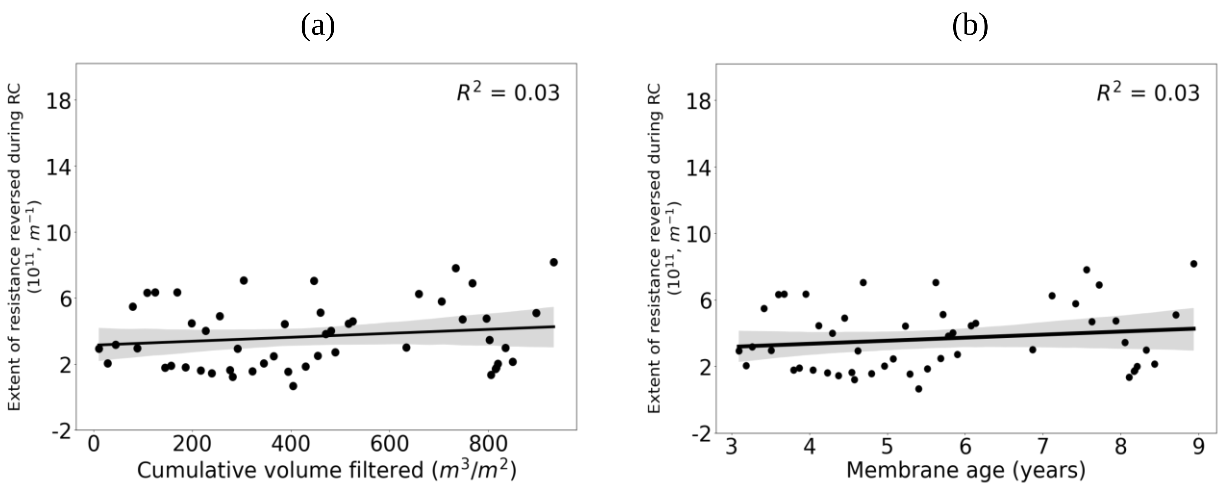


Figure 12: Typical extent of resistance reversed during RC

(a) with respect to cumulative volume filtered; (b) with respect to membrane age. Solid line corresponds to linear regression model fitted to the raw data. 95% confidence interval of the linear regression model is indicated by the light black shade. Results presented for train 1 at facility 3.

Overall, the extent of resistance reversed during RC does not increase with respect to both cumulative volume filtered and membrane age. Similar to clean membrane resistance (see section 5.1.2), there is no abrupt change in raw extent of resistance reversed during RC data; suggesting that the extent of resistance reversed during RC is relatively unimpacted by the operational conditions of the facilities. The rate of change of the estimated extent of resistance reversed during RC, for three trains at each facility, is presented in Figure 13. Details relating to the fit and slope of the linear regression model to the raw data is presented in Appendix C Table C5.

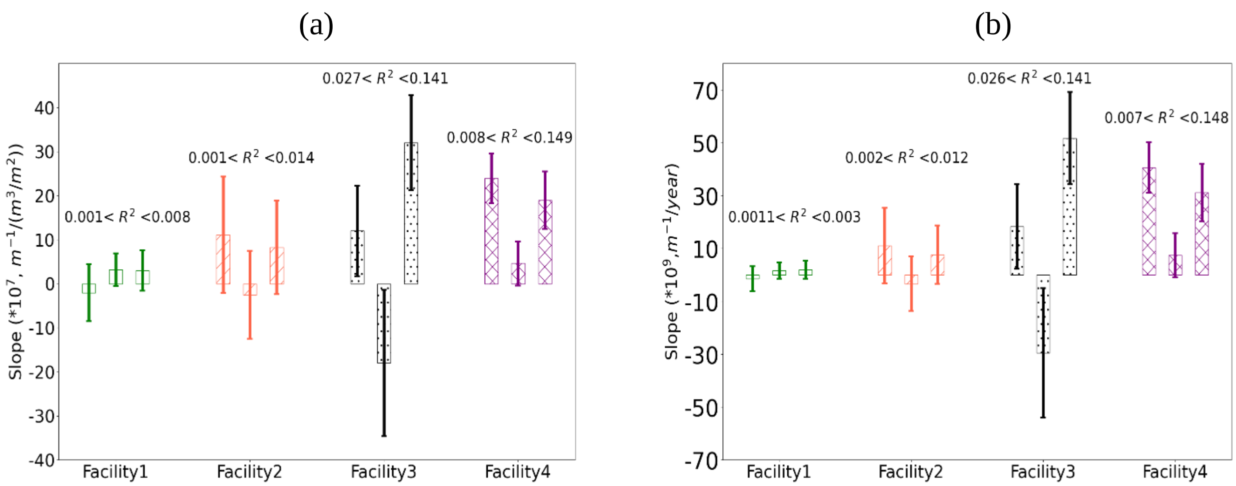


Figure 13: Rate of change of estimated extent of resistance reversed during RC

(a) with respect to cumulative volume filtered; (b) with respect to membrane age. Note that the scale of y axis in both figures are different as the magnitude of change with respect to cumulative volume filtered is different from that of age. Error bar indicates standard error of the estimated slope. Results presented for 3 randomly selected trains at each facility.

For facilities 1, 2, and 3, the rate change of the estimated extent of resistance reversed during RC is not significantly greater than 0 (zero) ($p > 0.05$) (except train 3 at facility 3). For train 3 at facility 3, the rate of change is significantly greater than zero 0 (zero) ($p < 0.05$). The insignificant and relatively inconsistent rate of change of the estimated extent of resistance reversed during RC suggests limited membrane ageing, based on this metric, at facilities 1, 2, and 3. For facility 4, the rate of change is significantly greater than 0 (zero) ($p < 0.05$), indicating membrane ageing, based on this metric, at this facility.

The coefficient of determination (R^2) was different for different facilities. The coefficient of determination (R^2) was approximately 0.1 for facilities 3 and 4, and less than 0.1 for facilities 1 and 2. A coefficient of determination of approximately 0.1 to less than 0.1 indicates that the estimated extent of resistance reversed during RC represents the raw extent of resistance reversed during RC data poorly. Note that when the rate of change is low, the coefficient of determination is inherently low (Miaou et al., 1996).

In summary, the extent of resistance reversed during RC did not increase with membrane age for 3 of the 4 facilities. For these facilities, the rate of change of the estimated extent of resistance reversed during RC could generally be approximated with a linear relationship poorly. For one of the facilities (i.e., facility 4), the extent of resistance reversed during RC was observed to increase with membrane age.

5.1.6 Summary of Change in Performance Factors

Of the performance factors considered for the present study, all except the extent of resistance reversed during RC generally increased with membrane age. A summary of the change in performance factors with membrane age is presented in Table 2.

Table 2: Change in performance factors with membrane age

Performance factors	**Increase with age (yes/no)**	**Extent to which the estimated data represents raw data?**
Hydraulically irreversible resistance	Yes	Moderate to very well
Clean membrane resistance	Yes	Moderate to very well
Total fouling rate	Yes	Poor
Extent of resistance reversed during BW	Yes	Poor to moderate
Extent of resistance reversed during RC	No	Poor

Of these performance factors, the clean membrane resistance (i.e., post RC resistance) was selected as the most suitable metric for monitoring and forecasting full-scale membrane performance based on the following outcomes:

- Unlike other performance factors, raw clean membrane resistance data did not have any abrupt changes suggesting that it is relatively independent of the operational conditions of the facilities;
- A substantive increase in clean membrane resistance, with respect to membrane age, was generally observed at the different facilities; and

- The rate of change of the estimated clean membrane resistance could generally be approximated with a linear relationship moderately well to very well.

Note that the clean membrane resistance was also previously suggested as the benchmark metric for monitoring membrane ageing (Robinson & Bérubé, 2020); and used by managers of facility 4 to inform membrane replacement.

5.2 Forecasting Membrane Performance

5.2.1 Minimum Data Required for Reliable Forecasting

Typical forecasting results are presented in Figure 14.

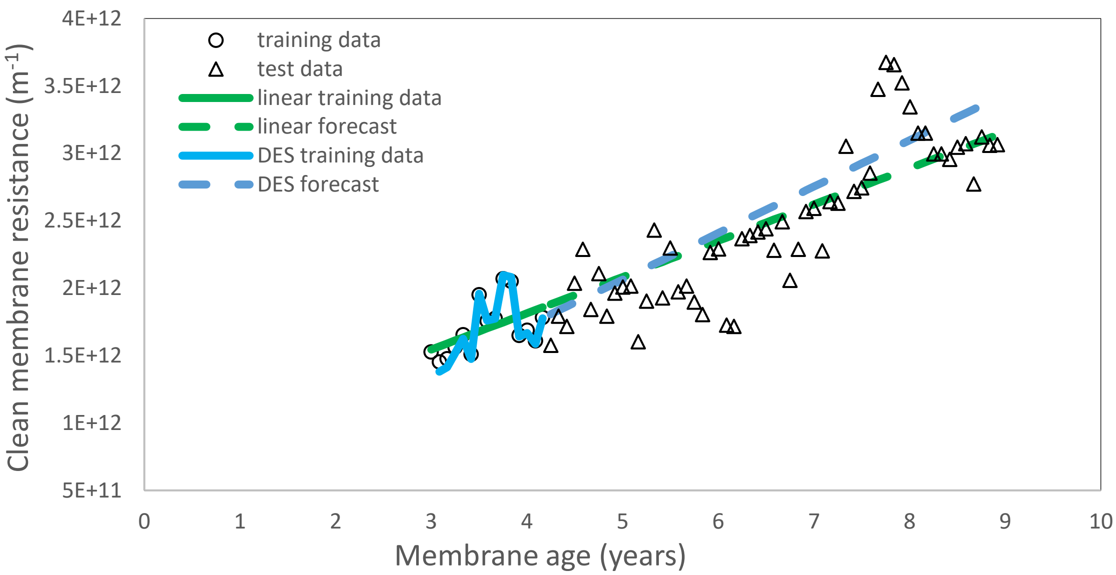


Figure 14: Typical forecasting results

Results presented for 20% of the total data used for training for train 3 at facility 3.

The error associated with the forecasting was quantified using the RMSE between the forecasted clean membrane resistance and actual clean membrane resistance for the test data. The RMSE between the forecasted and actual clean membrane resistance for each facility, when considering a range of fractions of the total data used for training, is presented in Figure 15. Note that forecasting was not performed for train 2 at facility 3 because of the exclusion of initial few years of data (see section 4.2.1).

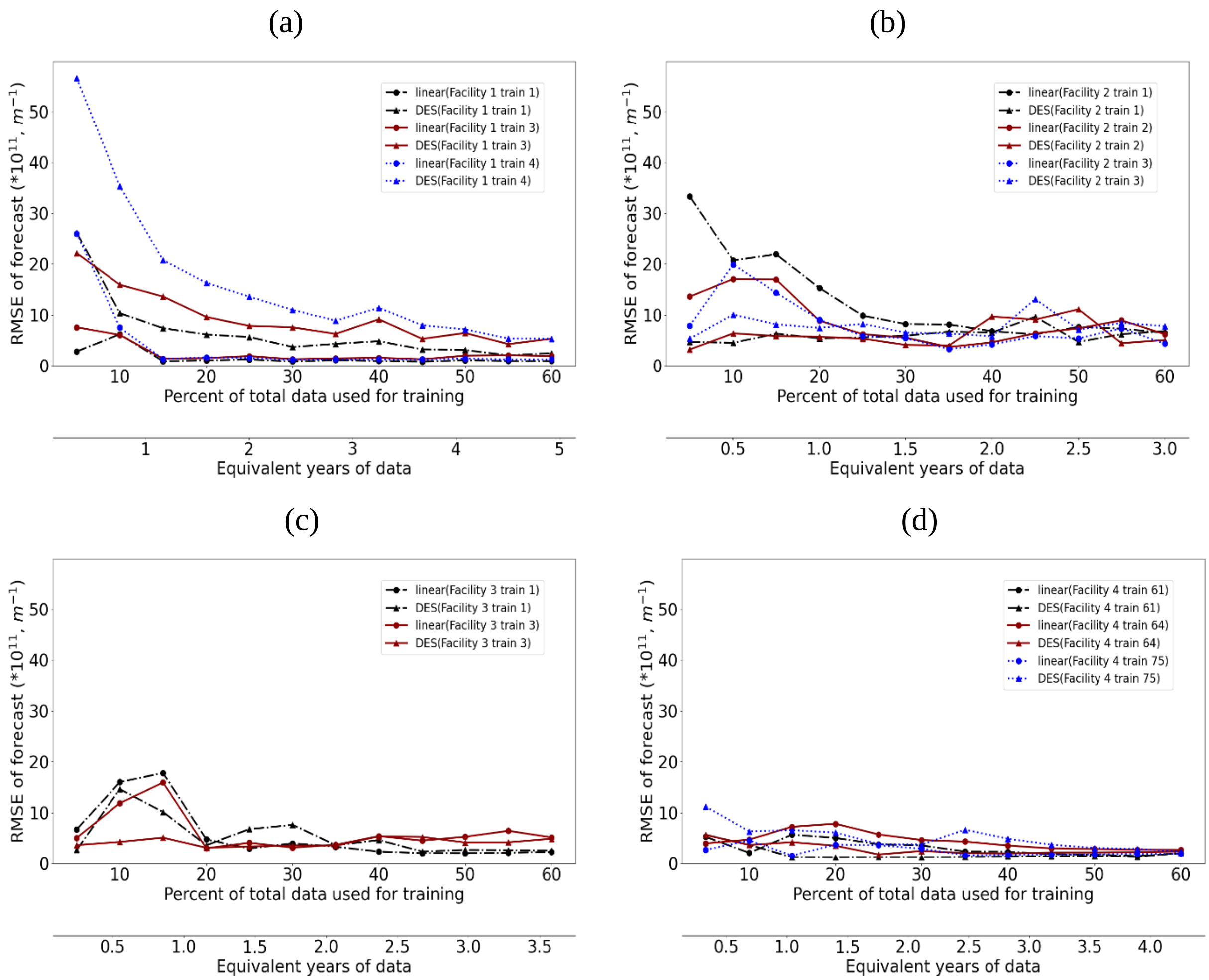


Figure 15: RMSE of the test data between forecasted and actual clean membrane resistance. (a) Facility 1; (b) Facility 2; (c) Facility 3; (d) Facility 4. Note that equivalent years of data on the x axis is different for different facilities due to the different data range reported from each facility.

The overall trends in the RMSE between the forecasted and actual clean membrane resistance, with respect to the fraction of the total data used for training, differed between different facilities and different trains at each facility. However, for all facilities and trains, the RMSE for both models was lower when 20% or more of the total data was used for training. 20% of total data corresponds approximately to the initial 1.5 years of the total data. For this reason, the initial 1.5 years of the total data was selected as the minimum to have a reliable forecast. Note that as the fraction of total data used for training increases, the RMSE of both models decreases and tends to converge.

5.2.2 Choosing between Linear and DES Forecasting Model

The mean RMSE of both models, for each train at each facility, was estimated by averaging the RMSE of the test data; associated with 20% to 60% of the total data used for training at 5% increment. Pairwise comparison was used to assess if the mean RMSE of both models differed, following the procedure in Appendix D. The estimated mean RMSE, and the outcomes of the comparison are summarized in Table 3.

Table 3: Mean RMSE of the linear and the DES model comparison

Facility/train	**Mean RMSE for forecasting model (m^{-1})**		**Different (yes/no)**
	Linear	**DES**	
Facility 1/train 1	*1.11e+11*	4.06e+11	yes
Facility 1/train 3	*1.78e+11*	6.97e+11	yes
Facility 1/train 4	*1.47e+11*	9.76e+11	yes
Facility 2/train 1	8.53e+11	*6.5e+11*	no
Facility 2/train 2	*6.53e+11*	6.61e+11	no
Facility 2/train 3	*5.76e+11*	7.97e+11	yes
Facility 3/train 1	*2.99e+11*	4.17e+11	no
Facility 3/train 2	Not considered	Not considered	-
Facility 3/train 3	4.64e+11	*4.28e+11*	no
Facility 4/train 61	2.80e+11	*1.53e+11*	yes
Facility 4/train 64	4.27e+11	*2.42e+11*	yes
Facility 4/train 75	*2.45e+11*	4.26e+11	yes

(note: the lowest mean RMSE for each train is italicized)

For facility 1, on average, the linear model provides a more accurate forecast. Recall from section 5.1.2 that facility 1 had no substantive increase in clean membrane resistance (i.e., limited membrane ageing). For facility 4, on average, the DES model provides a more accurate forecast. Recall from section 5.1.2 that facility 4 had a substantive increase in clean membrane resistance (i.e., substantive membrane ageing). For facilities 2 and 3, on average, both the linear and the DES model provide forecasts with similar accuracies. Recall from section 5.1.2 that facilities 2 and 3 also had a substantive increase in clean membrane resistance (i.e., substantive membrane ageing).

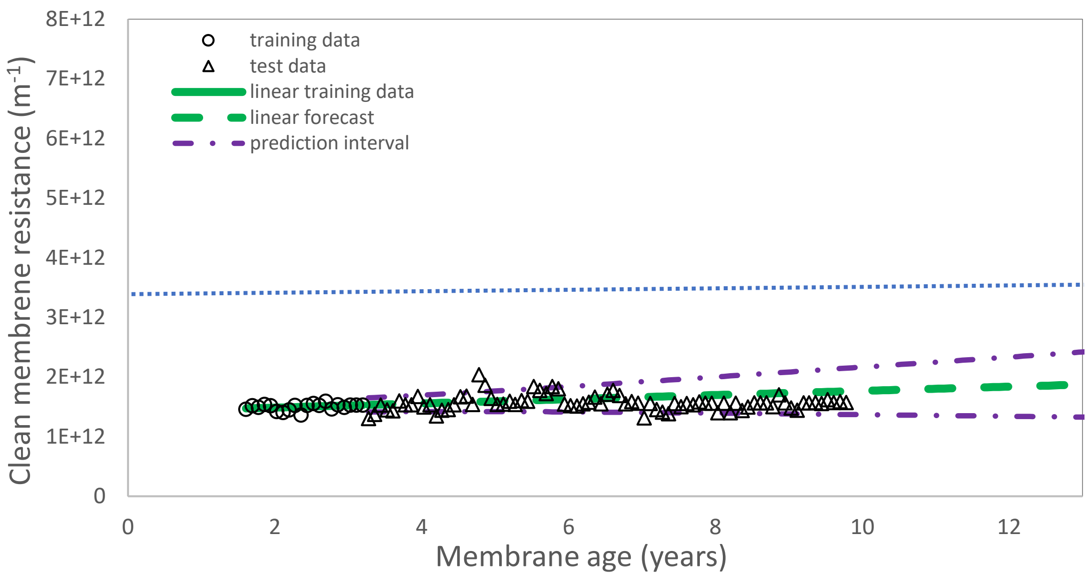
The above results suggest the following:

1. Facilities for which a substantive increase in clean membrane resistance is observed (i.e., substantive membrane ageing), the DES model provides a better or an equivalent forecast of the clean membrane resistance than the linear model.
2. However, facilities for which no substantive increase in clean membrane resistance is observed (i.e., limited membrane ageing), the linear model provides a better forecast of the clean membrane resistance than the DES model.

5.2.3 Forecasting Membrane Replacement

The resistance of a virgin (i.e., new) membrane was estimated based on the measured resistance when original or replacement modules were initially operated. The lowest average measured resistance (for all facilities) was approximately $1x10^{12}$ m^{-1} (lowest value ranged from $8x10^{11}$ to $1.4x10^{12}$ m^{-1}), which was considered as the virgin membrane resistance. A 70% reduction in permeability was selected as the trigger for membrane replacement (see section 4.2.1). A 70% reduction in permeability corresponds to a 3.3 fold increase in resistance to a value of $3.3x10^{12}$ m^{-1}. The forecast for the different facilities was extrapolated until the forecasted resistance was equal to $3.3x10^{12}$ m^{-1}. For facility 1, the linear model was used and for facilities 2, 3, and 4, the DES model was used for forecasting (see section 5.2.2). The results for the forecasting, for one typical train at each facility, are presented in Figure 16.

(a)



(b)

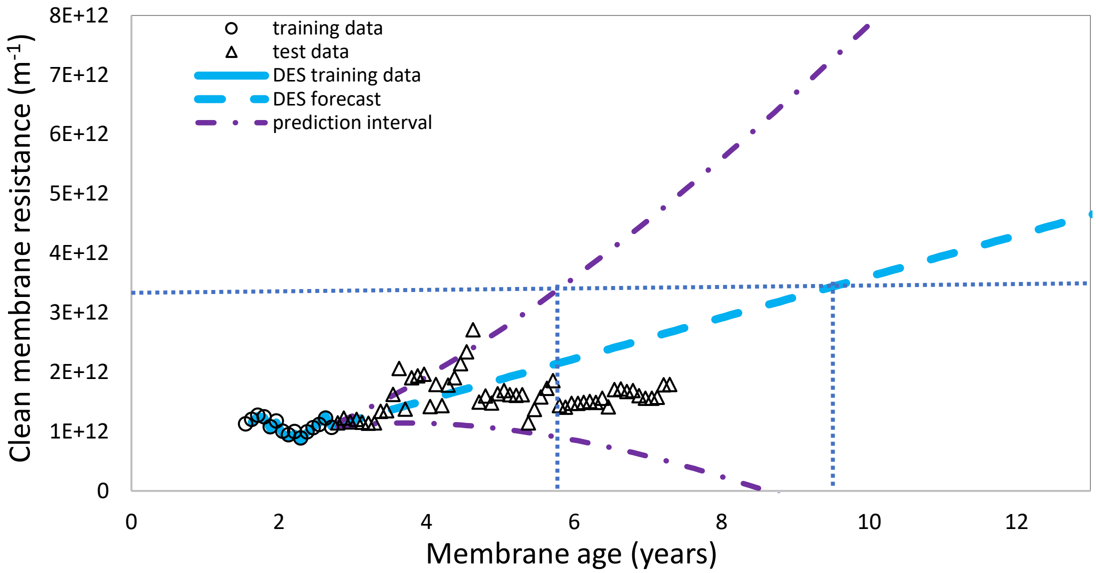
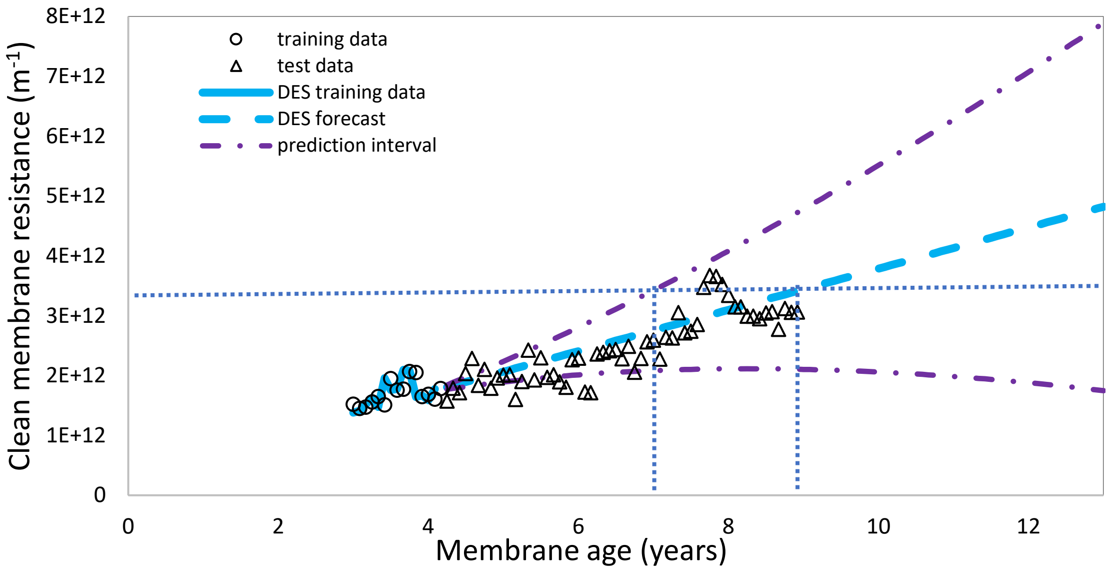


Figure 16: Forecasting membrane replacement

(a) Facility 1; (b) Facility 2; (c) Facility 3; (d) Facility 4. The horizontal dotted line corresponds to a reduction in permeability of 70% compared to that of a virgin membrane and the vertical dotted line corresponds to the membrane operational age when a 70% reduction in permeability is achieved. Forecasting results are based on 20% of the total data being used for training. Results presented for trains 3, 2, 3, and 61 at facilities 1, 2, 3, and 4 respectively.

(c)



(d)

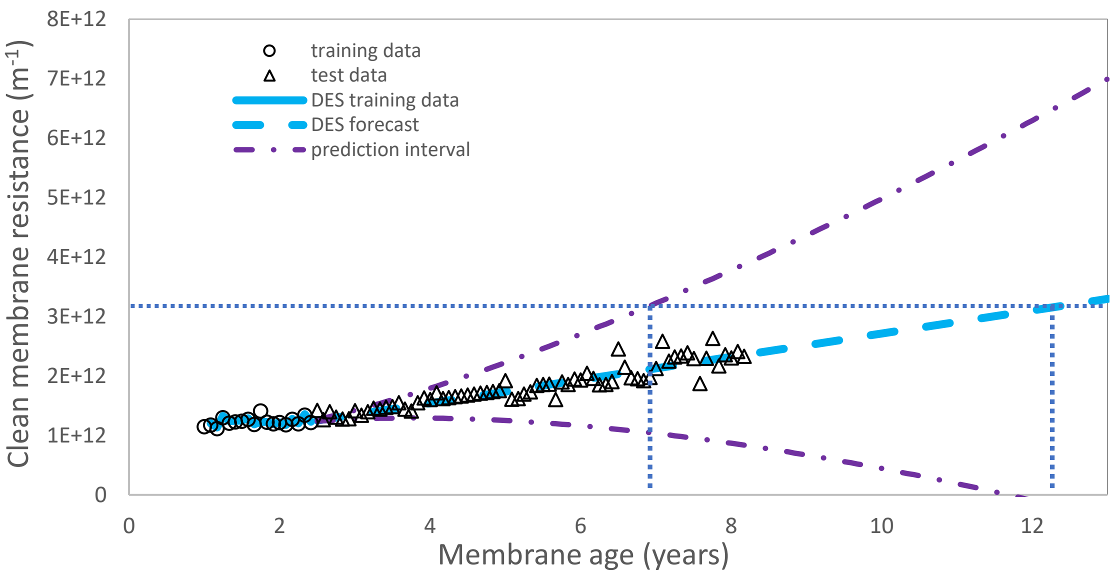


Figure 16: Forecasting membrane replacement (continued)

(a) Facility 1; (b) Facility 2; (c) Facility 3; (d) Facility 4. The horizontal dotted line corresponds to a reduction in permeability of 70% compared to that of a virgin membrane and the vertical dotted line corresponds to the membrane operational age when a 70% reduction in permeability is achieved. Forecasting results are based on 20% of the total data being used for training. Results presented for trains 3, 2, 3, and 61 at facilities 1, 2, 3, and 4 respectively.

For train 3 at facility 1, the linear model forecasts a replacement age of 54 (forecast interval from 23 to infinite) years. Note that in Figure 16a, the replacement age range is beyond the boundary of both axes and cannot be viewed. For train 2 at facility 2, train 3 at facility 3, and train 61 at facility 4, the DES model suggests a replacement age of 9 (forecast interval from 5 to infinite), 8 (forecast interval from 6 to infinite), and 12 (forecast interval from 7 to infinite) years, respectively (see Figure 16b – d). Note that for all trains at all facilities, the upper forecast interval was consistently infinite. Estimated replacement age, for all trains at each facility, are listed in Table 4.

Table 4: Membrane replacement age

Facility	**Train**	**Forecasting model**	**Replacement age per train (years)**	**Average replacement age per facility (years)**
Facility 1	Train 1	Linear	Could not be determined*	54 (23 – infinite)
	Train 3	Linear	54 (23 – infinite)	
	Train 4	Linear	Could not be determined*	
Facility 2	Train 1	DES	18 (13 – infinite)	11 (4 – infinite)
	Train 2	DES	9 (5 – infinite)	
	Train 3	DES	7 (4 – infinite)	
Facility 3	Train 1	DES	11 (6 – infinite)	9 (6 – infinite)
	Train 2	Not considered	Not considered	
	Train 3	DES	8 (6 – infinite)	
Facility 4	Train 61	DES	12 (7 – infinite)	11 (5 – infinite)
	Train 64	DES	13 (6 – infinite)	
	Train 75	DES	8 (5 – infinite)	

*Infinite replacement age was forecasted. Note that all the values are rounded down to the nearest integer to provide a conservative forecast of the replacement age.

For facilities 2, 3 and 4 (i.e., facilities with substantive membrane ageing), the forecast predicts average replacement ages of 9 to 11 years, which are consistent with the typical membrane replacement age (Chang et al., 2022). Note that facility 3, is currently replacing its membranes after approximately 9 years of operation. Based on the data from the initial 1.5 years of operation (i.e., equivalent to the initial 20% of the total data), the forecasting model would have been able to 'predict' membrane replacement at 9 years of operation. For facility 1, the forecast predicts that the membranes should be replaced at approximately 54 years of operation. This timeline is much greater than the typical range. For this facility, the decline in other performance factors that are not linked to the clean membrane resistance (e.g., such as fiber breakages), are likely to govern membrane replacement age. It should be noted that the long timeline to replacement (i.e., 54 years) indicates that the performance of facility 1, in terms of clean membrane replacement, is very good.

6 Conclusions

From the present study, followings can be concluded:

1) Of the performance factors examined, the hydraulically irreversible resistance, the clean membrane resistance, the total fouling rate and the extent of resistance reversed during BW were observed to increase with membrane age; while the extent of resistance reversed during RC was not observed to increase with membrane age. Among these performance factors, the clean membrane resistance (i.e., post RC resistance) was identified as the most suitable metric to monitor membrane ageing and to forecast membrane performance.
2) The rates at which these performance factors changed were substantially different at different facilities indicating that the rates of change are greatly impacted by site-specific conditions.
3) Utilizing the clean membrane resistance data from approximately the initial 1.5 years of years of operation, it is possible to make reliable forecast of membrane replacement age. The forecast is expected to enable the managers of full-scale membrane facilities to make evidence-based decisions and better manage their finances.
4) For facilities with substantive membrane ageing, the DES model outperforms the linear model in forecasting membrane replacement age; while for facilities with limited membrane ageing, the linear model performs better in forecasting membrane replacement age.

7 Recommendations

The present study recommends the followings:

1) The present study was focused exclusively on historical resistance data from full-scale membrane facilities. The study did not attempt to identify the causes of observed changes in historical resistance data. Future studies should attempt to identify the causes of observed changes in historical resistance data.
2) The forecasting models (i.e., linear and DES) used in the present study are influenced by the variability of the historical resistance data. Care should be taken to collect accurate historical data.

References

Abdullah, S. Z., & Bérubé, P. R. (2013). Assessing the effects of sodium hypochlorite exposure on the characteristics of PVDF based membranes. Water Research, 47(14), 5392–5399. https://doi.org/10.1016/j.watres.2013.06.018

Adham, S., Chiu, K.-P., Gramith, K. and Oppenheimer, J. (2005). Development of a Microfiltration and Ultrafiltration Knowledge Base, Denver, Co.: AWWARF

Ahmad, W. K. A., & Ahmad, S. (2013). Arima model and exponential smoothing method: A comparison. AIP Conference Proceedings, 1522, 1312–1321. https://doi.org/10.1063/1.4801282

Ayala, D. F., Ferre, V., & Judd, S. J. (2011). Membrane life estimation in full-scale immersed membrane bioreactors. Journal of Membrane Science, 378(1–2), 95–100. https://doi.org/10.1016/j.memsci.2011.03.013

Baars, E. T., Heijman, S. G. J., & Bosklopper, T. G. J. (2005). Red alert on transmembrane pressure (TMP). Desalination, 179(1-3 SPEC. ISS.), 125–130. https://doi.org/10.1016/j.desal.2004.11.061

Berthouex, P. M. (Paul M., & Brown, L. C. (2002). Statistics for environmental engineers. Lewis Publishers

Chang, H., Zhu, Y., Yu, H., Qu, F., Zhou, Z., Li, X., Yang, Y., Tang, X., & Liang, H. (2022). Long-term operation of ultrafiltration membrane in full-scale drinking water treatment plants in China: Characteristics of membrane performance. Desalination, 543. https://doi.org/10.1016/j.desal.2022.116122

Cote, P., Alam, Z., & Penny, J. (2012). Hollow fiber membrane life in membrane bioreactors (MBR). Desalination, 288, 145–151. https://doi.org/10.1016/j.desal.2011.12.026

Crittenden, J. C. (John C., & Montgomery Watson Harza (Firm). (2012). MWH's water treatment : principles and design. John Wiley & Sons

De Wilde, W., Thoeye, C., & De Gueldre, G. (2007). Membrane life expectancy assessment after 3 years of MBR operation at WWTP Schilde. In Proceedings of the 4th International Water Association Conference on Membranes for Water and Wastewater Treatment, Harrogate, UK (pp. 15-17)

Fenu, A., De Wilde, W., Gaertner, M., Weemaes, M., de Gueldre, G., & Van De Steene, B. (2012). Elaborating the membrane life concept in a full scale hollow-fibers MBR. Journal of Membrane Science, 421–422, 349–354. https://doi.org/10.1016/j.memsci.2012.08.001

Filho, M., Mailler, R., Rocher, V., Fayolle, Y., & Causserand, C. (2021). Comprehensive study of supported PVDF membrane ageing in MBR: A direct comparison between changes at bench scale and full scale. Separation and Purification Technology, 279. https://doi.org/10.1016/j.seppur.2021.119695

Fisher, R., 1921. On the "Probable Error" of a Coefficient of Correlation Deduced from a Small Sample. Metron, Volume 1, pp. 3-32

Hajibabania, S., Antony, A., Leslie, G., & Le-Clech, P. (2012). Relative impact of fouling and cleaning on PVDF membrane hydraulic performances. Separation and Purification Technology, 90, 204–212. https://doi.org/10.1016/j.seppur.2012.03.001

He, Y., Sharma, J., Bogati, R., Liao, B. Q., Goodwin, C., & Marshall, K. (2014). Impacts of Aging and Chemical Cleaning on the Properties and Performance of Ultrafiltration Membranes in Potable Water Treatment. Separation Science and Technology, 49(9), 1317–1325. https://doi.org/10.1080/01496395.2014.882359

Heo, S. K., Nam, K. J., Woo, T. Y., & Yoo, C. K. (2022). Digitally-transformed early-warning protocol for membrane cleaning based on a fouling-cumulative sum chart: Application to a full-scale MBR plant. Journal of Membrane Science, 643. https://doi.org/10.1016/j.memsci.2021.120080

Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice. Otexts

Miaou, S., Lu, A., & Lum, H. S. (1996). Pitfalls of using R 2 to evaluate goodness of fit of accident prediction models. Transportation Research Record, 1542(1), 6-13. https://doi.org/10.1177/0361198196154200102

Porcelli, N., & Judd, S. (2010). Chemical cleaning of potable water membranes: A review. Separation and Purification Technology, 71(2), 137-143. https://doi.org/10.1016/j.seppur.2009.12.007

Prulho, R., Therias, S., Rivaton, A., & Gardette, J. L. (2013). Ageing of polyethersulfone/polyvinylpyrrolidone blends in contact with bleach water. Polymer Degradation and Stability, 98(6), 1164–1172. https://doi.org/10.1016/j.polymdegradstab.2013.03.011

Regula, C., Carretier, E., Wyart, Y., Gésan-Guiziou, G., Vincent, A., Boudot, D., & Moulin, P. (2014). Chemical cleaning/disinfection and ageing of organic UF membranes: A review. Water Research (Oxford), 56, 325-365. https://doi.org/10.1016/j.watres.2014.02.050

Ren, L., Yu, S., Yang, H., Li, L., Cai, L., Xia, Q., Shi, Z., & Liu, G. (2021). Chemical cleaning reagent of sodium hypochlorite eroding polyvinylidene fluoride ultrafiltration membranes: Aging pathway, performance decay and molecular mechanism. Journal of Membrane Science, 625. https://doi.org/10.1016/j.memsci.2021.119141

Robinson, S., Abdullah, S. Z., Bérubé, P., & Le-Clech, P. (2016). Ageing of membranes for water treatment: Linking changes to performance. Journal of Membrane Science, 503, 177-187. https://doi.org/10.1016/j.memsci.2015.12.033

Robinson, S., & Bérubé, P. R. (2020). Membrane ageing in full-scale water treatment plants. Water Research, 169. https://doi.org/10.1016/j.watres.2019.115212

Smith, P. J., Vigneswaran, S., Ngo, H. H., Ben-Aim, R., & Nguyen, H. (2005). Design of a generic control system for 49eriod49ing back flush durations in a submerged membrane hybrid reactor. Journal of Membrane Science, 255(1–2), 99–106. https://doi.org/10.1016/j.memsci.2005.01.026

Teychene, B., Touffet, A., Baron, J., Welte, B., Joyeux, M., & Gallard, H. (2018). Predicting of ultrafiltration performances by advanced data analysis. Water Research, 129, 365–374. https://doi.org/10.1016/j.watres.2017.11.023

Touffet, A., Baron, J., Welte, B., Joyeux, M., Teychene, B., & Gallard, H. (2015). Impact of pretreatment conditions and chemical ageing on ultrafiltration membrane performances. Diagnostic of a coagulation/adsorption/filtration process. Journal of Membrane Science, 489, 284–291. https://doi.org/10.1016/j.memsci.2015.04.043

U.S.EPA. (2005). Membrane filtration guidance manual. Note on the membrane filtration guidance manual purpose

Vandeput, N. (2021). Data science for supply chain forecasting. Walter de Gruyter GmbH & Co KG

Wang, Z., Ma, J., Tang, C. Y., Kimura, K., Wang, Q., & Han, X. (2014). Membrane cleaning in membrane bioreactors: A review. In Journal of Membrane Science (Vol. 468, pp. 276–307). Elsevier. https://doi.org/10.1016/j.memsci.2014.05.060

Woo, T. Y., Nam, K. J., Heo, S. K., Lim, J. Y., Kim, S. Y., & Yoo, C. K. (2022). Predictive maintenance system for membrane replacement time detection using AI-based functional profile monitoring: Application to a full-scale MBR plant. Journal of Membrane Science, 649. https://doi.org/10.1016/j.memsci.2022.120400

Wright, D. J. (1986). Forecasting data published at irregular time intervals using an extension of Holt’s method. Management science, 32(4), 499-510

Yu, H., Chang, H., Li, X., Zhou, Z., Song, W., Ji, H., & Liang, H. (2021). Long-term fouling evolution of polyvinyl chloride ultrafiltration membranes in a hybrid short-length sedimentation/

ultrafiltration process for drinking water production. Journal of Membrane Science, 630. https://doi.org/10.1016/j.memsci.2021.119320

Yu, H., Li, X., Chang, H., Zhou, Z., Zhang, T., Yang, Y., Li, G., Ji, H., Cai, C., & Liang, H. (2020). Performance of hollow fiber ultrafiltration membrane in a full-scale drinking water treatment plant in China: A systematic evaluation during 7-year operation. Journal of Membrane Science, 613. https://doi.org/10.1016/j.memsci.2020.118469

Appendix A

Designed operational characteristics of the different facilities are summarized in Table A1 – Table A4.

Table A1: Capacity, source water and pretreatment approaches of different facilities

	Facility 1	**Facility 2**	**Facility 3**	**Facility 4**
Source (Lake/river etc.)	Lake	River	Lake	Lake
Max capacity	65.2 MLD	116 MLD	20.9 MLD	400 MLD
Avg Capacity	39.1MLD	83 MLD	8-16 MLD	200 MLD
Pretreatment (Pre chlorination/ ozonation etc.)	Pre-chlorination, Ozonation, UV	Pre-chlorination	Pre-chlorination	--
Pretreatment chemical used and dose	--	Sodium Hydroxide pH adjustment: 7.5	Sodium Hydroxide pH adjustment: 7.5	Pre-chlorination Dechlorination with sodium bisulfite; pH adjustment with sodium hydroxide Ozonation
Coagulant	PAX-XL	Aluminium Chlorohydrate (ACH)	Aluminium Chlorohydrate (ACH)	No coagulation and flocculation
Coagulant addition	Inline	Inline	Rapid mix tank	
Rapid mixing G (/s)	2500	--	--	

	Facility 1	Facility 2	Facility 3	Facility 4
Coagulant dose	2 mg/L	Flow and UVA proportional; average~1-2 mg/L (as high as 13mg/L during 52eriod of high UVA)	Flow proportional typically 4-5 mg/L	No coagulation and flocculation
Flocculant tank details	Consist of a dynamic mixer and two flocculation trains with a total storage volume of 1,312 m^3	3 floc trains, each with 2 in series	2 in series for every train, only the first being mixed	
Flocculation G (/s)	50	At 100% - 78.65/s operating at 50% (39.3/s) both tanks at same mixing speed	75% of mixer speed capacity (will send info on mixer power)	
Flocculation HRT (minutes)	10 to 15	31.24 total at average flow (38 MLD)	--	

Note: -- indicates that data not available.

Table A2: Primary membrane train details of different facilities

	Facility 1	**Facility 2**	**Facility 3**	**Facility 4**
Membrane type*	ZW1000B	ZW1000A	ZW1000A	ZW1000A
Number of trains	7 installed trains, 5 operating trains	7	3	12 (2 banks of 6 trains each)
Casettes/train	4	5	4	8
Casette type	v3	v3.3	V3.2	96M
Modules/casette (specify blank as well)		84/96 (4 blank stacks of 3 modules)	48/60 (4 blank stacks of 3 modules)	87/96 (3 blank stacks of 3 modules)
Modules/casette	90	84	48	87
Module type	v3	v4	v3	v4
Area/module (m^2)	46.5	41.8	41.8	51.1
Modules/train	360	420	192	696
Total membrane area (m^2)	66960	122892	24076.8	426787.2
Average Flux (LMH)	24.33	28.14	13.84	19.52

*1000B older membrane chemistry; no longer in use; 1000A current membrane chemistry

Table A3: Maintenance cleaning details of different facilities

	Facility 1	**Facility 2**	**Facility 3**	**Facility 4**
Maintenance cleaning	1 time per day, 100 mg/L Sodium hypochlorite	**Type A**: 2 times per month, 100 mg/L sodium hypochlorite, 15 mins soak after hydraulic backwash (no CEB), **Type B**: 1 x per month, 500mg/L Citric acid , 1220 mg/L hydrochloric acid - target pH of 2, 15 mins soak after hydraulic backwash (no CEB)	3 to 5 times per week (only one type), 100 ppm hypo (no citric acid), 1 hr soak after hydraulic backwash (no CEB), Chemical solution reused once and then discharged (either for primary of secondary train)	Sodium hypo soak of 15 mins at 250 mg/L, usually in 1-2 days

Table A4: Chemical cleaning details (Recovery cleaning) of different facilities

	Facility 1	**Facility 2**	**Facility 3**	**Facility 4**
Recovery cleaning	Two 6-hour recovery clean cycles occur every 42 days, **First:** a high pH cycle commences using 500 mg/L sodium hypochlorite; **Second**: a low pH cycle using 2000 mg/L citric acid; increase the water temperature to 40°C.	**Type A**: 1 time per month, 500 mg/L Sodium hypochlorite, 5 hr soak, Usually not heated. 2 x per year do heated cleans for 1 month (to 40C).	**Type A**: 1 time per month (one of Type A in one month then one of type B in following month), 1000 ppm hypo, 5 hr soak, Initial solution heated to 20C, if temp falls to < 18C, heater goes on and permeate recirculated, Added to tank, recirculated from permeate, waste from tank	**Type A**: Sodium hypo soak of 5-6 hrs at 500 mg/L, monthly, uses heated water all year round
		Type B: 2 x per year, 2000 mg/L Citric acid, 915mg/L Hydrochloric acid (target pH 2), 5 hr soak, Heated to 40°C,	**Type B:** Monthly (one of Type A and B), 500 ppm citric, 5 hr soak, Initial solution heated to 20C, if temp falls to < 18C, heater goes on and permeate recirculated, Added to tank, recirculated from permeate, waste from tank	**Type B:** Citric acid soak for 5-6 hr at 2000 mg/L + pH adjustment using sulphuric acid. The use of sulphuric acid has stopped since 2019, monthly

Appendix B

Table B1: Details and frequency of collected data.

Parameters \ Facilities	Facility 1	Facility 2	Facility 3	Facility 4
Permeate flow	Every minute	Every fifteen minute	Every minute	Average of every 5 minutes
Tank water level	Every minute	Every fifteen minute	Every minute	Not available
Temperature	Every minute	Before and after cleaning events	Every minute	Before cleaning events
Temperature corrected permeability	Every minute	Before and after cleaning events	Every minute (no data available before 2020)	Before and after cleaning events
TMP	Every minute	Before and after cleaning events	Every minute	Average of every 5 minutes
Turbidity	Every minute	Every fifteen minute	Every minute	Not available
Chemical flow rate (i.e., hypo and citric tags)	Every minute	Not available	Not available	Not available
Occurrences of MC	Not available (but can be calculated from data)	Exact time	Date but not exact time	Not available
Occurrences of RC	Not available (but can be calculated from data)	Exact time with cleaning agent types	Date with cleaning agent types but not exact time	Date with cleaning agent types but not exact time

Table B2: Reported data range for all facilities

Facilities	Data range (year)	Age of membrane trains (in years)
Facility 1	2013-2022	1-10
Facility 2	2015-2022	1-8
Facility 3	2016-2021	3-8
Facility 4	2014-2022	1-9

Appendix C

Table C1: Summary of linear regression analysis of historical hydraulically irreversible resistance.

Facilities	Cumulative volume filtered as independent variable				Membrane age as independent variable			
	Slope	**Slope std error**	**R^2**	**p-value**	**Slope**	**Slope std error**	**R^2**	**p-value**
Facility 1 train 1	-1.38e+07	5.07e+06	0.001	1.00e-11	-8.44e+08	3.69e+08	0.00036	2.23e-2
Facility 1 train 3	8.02e+07	5.86e+06	0.011	2.06e-42	6.58e+09	4.43e+08	0.014	1.08e-49
Facility 1 train 4	1.05e+08	5.35e+06	0.023	3.99e-84	8.25e+09	4.03e+08	0.025	2.95e-92
Facility 2 train 1	1.44e+09	1.38e+07	0.371	< 2e-16	1.50e+11	1.49e+09	0.355	< 2e-16
Facility 2 train 2	1.12e+09	1.02e+07	0.403	< 2e-16	1.14e+11	1.05e+09	0.392	< 2e-16
Facility 2 train 3	1.24e+09	1.26e+07	0.357	< 2e-16	1.25e+11	1.33e+09	0.335	< 2e-16
Facility 3 train 1	9.32e+08	5.58e+06	0.571	< 2e-16	1.43e+11	8.84e+08	0.555	< 2e-16
Facility 3 train 2	1.24e+09	2.31e+07	0.188	< 2e-16	1.71e+11	3.43e+09	0.168	< 2e-16
Facility 3 train 3	1.91e+09	7.49e+06	0.745	< 2e-16	3.08e+11	1.20e+09	0.747	< 2e-16
Facility 4 train 61	1.50e+09	3.65e+06	0.867	< 2e-16	2.54e+11	5.78e+08	0.877	< 2e-16
Facility 4 train 64	1.73e+09	3.88e+06	0.888	< 2e-16	2.86e+11	6.02e+08	0.896	< 2e-16
Facility 4 train 75	1.45e+09	4.36e+06	0.814	< 2e-16	2.42e+11	6.69e+08	0.832	< 2e-16

Table C2: Summary of linear regression analysis of historical clean membrane resistance.

Facilities	**Cumulative volume filtered as independent variable**				**Membrane Age as independent variable**			
	Slope	**Slope std error**	**R^2**	**p-value**	**Slope**	**Slope std error**	**R^2**	**p-value**
Facility 1 train 1	3.94e+07	7.54e+07	0.0031	0.602	3.22e+09	5.47e+09	0.0039	0.558
Facility 1 train 3	1.06e+08	7.31e+07	0.022	0.152	8.63e+09	5.51e+09	0.026	0.120
Facility 1 train 4	1.09e+08	7.98e+07	0.019	0.174	8.92e+09	6.00e+09	0.023	0.140
Facility 2 train 1	1.53e+09	2.53e+08	0.373	1.05e-7	1.61e+11	2.76e+10	0.358	2.27e-7
Facility 2 train 2	1.10e+09	1.58e+08	0.409	1.44e-9	1.11e+11	1.63e+10	0.4	2.47e-9
Facility 2 train 3	1.04e+09	1.56e+08	0.39	4.57e-9	1.05e+11	1.63e+10	0.374	1.17e-8
Facility 3 train1	8.86e+08	8.56e+07	0.661	1.60e-14	1.37e+11	1.35e+10	0.651	3.52e-14
Facility 3 train2	1.60e+09	2.97e+08	0.454	4.92e-6	2.30e+11	4.52e+10	0.425	1.22e-5
Facility 3 train3	1.59e+09	1.40e+08	0.713	9.95e-16	2.59e+11	2.24e+10	0.719	6.11e-16
Facility 4 train 61	9.81e+08	3.36e+07	0.879	7.45e-56	1.68e+11	5.66e+09	0.882	1.17e-56
Facility 4 train 64	1.43e+09	5.66e+07	0.844	1.80e-49	2.41e+11	9.39e+09	0.848	5.02e-50
Facility 4 train 75	8.79e+08	4.65e+07	0.75	1.28e-37	1.47e+11	7.68e+09	0.754	5.33e-38

Table C3: Summary of linear regression analysis of historical total fouling rate

Facilities	cumulative volume filtered as independent variable				Membrane Age as independent variable			
	Slope	**Slope std error**	**R^2**	**p-value**	**Slope**	**Slope std error**	**R^2**	**p-value**
Facility 1 train 1	9.02e+08	6.57e+07	0.0099	1.24e-42	5.57e+10	4.80e+09	0.007	5.26e-31
Facility 1 train 3	3.18e+09	5.13e+07	0.16	< 2e-16	2.39e+11	3.88e+09	0.159	< 2e-16
Facility 1 train 4	2.33e+09	5.23e+07	0.088	< 2e-16	1.71e+11	3.95e+09	0.083	< 2e-16
Facility 3 train 1	1.08e+09	3.70e+07	0.034	4.57e-183	1.71e+11	5.75e+09	0.035	3.81e-192
Facility 3 train 2	4.76e+09	8.64e+07	0.181	< 2e-16	6.78e+11	1.28e+10	0.17	< 2e-16
Facility 3 train 3	3.56e+09	3.95e+07	0.242	< 2e-16	5.79e+11	6.35e+09	0.246	< 2e-16
Facility 4 train 61	Data not available							
Facility 4 train 64								
Facility 4 train 75								
Facility 2 train 1								
Facility 2 train 2								
Facility 2 train 3								

Table C4: Summary of linear regression analysis of historical extent of resistance reversed during BW

Facilities	Cumulative volume filtered as independent variable				Membrane Age as independent variable			
	Slope	**Slope std error**	**R^2**	**p-value**	**Slope**	**Slope std error**	**R^2**	**p-value**
Facility 1 train 1	7.72e+08	6.77e+07	0.009	5.99e-30	4.58e+10	4.94e+09	0.006	2.00e-20
Facility 1 train 3	2.81e+09	4.62e+07	0.188	< 2e-16	2.10e+11	3.50e+09	0.184	< 2e-16
Facility 1 train 4	1.84e+09	4.92e+07	0.08	4.94e-295	1.35e+11	3.72e+09	0.075	1.16e-277
Facility 2 train 1	2.90e+09	2.88e+08	0.005	7.78e-24	2.87e+11	3.07e+10	0.004	8.38e-21
Facility 2 train 2	1.80e+09	2.90e+08	0.002	5.56e-10	1.73e+11	2.99e+10	0.002	7.60e-9
Facility 2 train 3	-4.76e+07	1.95e+09	0.0	9.80e-1	-4.26e+10	2.03e+11	0.0	8.33e-1
Facility 3 train 1	9.01e+08	3.25e+07	0.035	5.95e-166	1.43e+11	5.05e+09	0.036	6.74e-173
Facility 3 train 2	4.27e+09	7.46e+07	0.21	< 2e-16	6.09e+11	1.10e+10	0.198	< 2e-16
Facility 3 train 3	2.92e+09	3.25e+07	0.269	< 2e-16	4.75e+11	5.24e+09	0.272	< 2e-16
Facility 4 train 61	5.16e+09	3.23e+07	0.508	< 2e-16	8.81e+11	5.51e+09	0.508	< 2e-16
Facility 4 train 64	5.77e+09	5.26e+07	0.337	< 2e-16	9.67e+11	8.81e+09	0.338	< 2e-16
Facility 4 train 75	6.07e+09	3.97e+07	0.495	< 2e-16	1.01e+12	6.55e+09	0.501	< 2e-16

Table C5: Summary of linear regression analysis of historical extent of resistance reversed during RC

Facilities	Cumulative volume filtered as independent variable				Membrane Age as independent variable			
	Slope	**Slope std error**	**R^2**	**p-value**	**Slope**	**Slope std error**	**R^2**	**p-value**
Facility 1 train 1	-2.04e+07	6.46e+07	0.001	7.53e-1	-1.43e+09	4.70e+09	0.0011	7.62e-1
Facility 1 train 3	3.17e+07	3.67e+07	0.008	3.89e-1	1.58e+09	3.10e+09	0.003	6.11e-1
Facility 1 train 4	2.98e+07	4.55e+07	0.005	5.15e-1	1.93e+09	3.43e+09	0.003	0.003
Facility 2 train 1	1.11e+08	1.32e+08	0.014	4.05e-1	1.10e+10	1.43e+10	0.012	4.44e-1
Facility 2 train 2	-2.55e+07	9.98e+07	0.001	7.99e-1	-3.38e+09	1.03e+10	0.002	7.43e-1
Facility 2 train 3	8.22e+07	1.06e+08	0.01	4.42e-1	7.64e+09	1.10e+10	0.008	4.90e-1
Facility 3 train 1	1.20e+08	1.03e+08	0.027	2.52e-1	1.83e+10	1.60e+10	0.026	2.56e-1
Facility 3 train 2	-1.80e+08	1.66e+08	0.035	2.88e-1	-2.95e+10	2.45e+10	0.043	2.38e-1
Facility 3 train 3	3.20e+08	1.08e+08	0.141	4.74e-3	5.17e+10	1.75e+10	0.141	4.71e-3
Facility 4 train 61	2.39e+08	5.64e+07	0.149	4.76e-5	4.06e+10	9.59e+09	0.148	5.02e-5
Facility 4 train 64	4.62e+07	5.00e+07	0.008	3.58e-1	7.44e+09	8.39e+09	0.007	3.77e-1
Facility 4 train 75	1.90e+08	6.53e+07	0.069	4.40e-3	3.12e+10	1.09e+10	0.067	4.94e-3

Appendix D

The RMSE of the test data of each model, for each train at each facility, corresponding to 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, and 60% of the total data used for training constituted a dataset. The mean RMSE of each model was calculated by averaging the RMSE of each dataset. For example, at train 1 facility 3, the RMSE of the test data corresponding to 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, and 60% of the total data used for training are 4.86e+11, 3.09e+11, 4.09e+11, 3.43e+11, 2.46e+11, 2.17e+11, 2.15e+11, 2.24e+11, and 2.39e+11, respectively for the linear model with a mean RMSE of *2.99e+11*; while for the DES model the RMSE of the test data are 3.72e+11, 6.87e+11, 7.71e+11, 3.72e+11, 4.73e+11, 2.48e+11, 2.79e+11, 2.74e+11 and 2.68e+11, respectively with a mean RMSE of *4.17e+11*.

Mean RMSE of both models (i.e., 2.99e+11 and 4.17e+11), for each train at each facility, was then compared to assess if they were statistically different using the following approach:

- First normality assumption of t test for both datasets (RMSE of the DES and the linear model) was checked. If assumption was met, Welch t test was performed.
- If normality assumption was violated, log10 transformation was made to both datasets. If, in both transformed datasets, normality assumption was met, Welch t test was done on the transformed dataset.
- If normality assumption was still not met, non-parametric MannWhitney U test was performed.