



HAL
open science

Forecasting Multicycle Hollow Fiber Ultrafiltration Fouling Using Time Series Analysis

Ghadi Dagher, Aurélien Martin, Jean-Michel Galharret, Laurent Moulin,
Jean-Philippe Croué, Benoit Teychene

► **To cite this version:**

Ghadi Dagher, Aurélien Martin, Jean-Michel Galharret, Laurent Moulin, Jean-Philippe Croué, et al.. Forecasting Multicycle Hollow Fiber Ultrafiltration Fouling Using Time Series Analysis. *Journal of Water Process Engineering*, 2023, 56, pp.104441. 10.1016/j.jwpe.2023.104441 . hal-04255886

HAL Id: hal-04255886

<https://hal.science/hal-04255886v1>

Submitted on 24 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

1 **Forecasting Multicycle Hollow Fiber Ultrafiltration**

2 **Fouling Using Time Series Analysis.**

3 Authors: G. Dagher¹, A. Martin², JM. Galharret³, L. Moulin², JP. Croué¹, B. Teychene^{1,*}

4 1. Institut de Chimie des Milieux et Matériaux de Poitiers (IC2MP - UMR CNRS 7285),
5 Université de Poitiers, CNRS, E-BiCOM, 1 rue Marcel Doré, F-86073 Poitiers, France.

6 2. Eau de Paris, DRDQE, R&D laboratory, 33 Avenue Jean Jaurès, 94200, Ivry-Sur-Seine,
7 France

8 3. Laboratoire de Mathématiques Jean Leray, UMR 6629 CNRS – Nantes Université, 44000
9 Nantes, France

10 **Keywords:** Time series analysis, exponential smoothing, ultrafiltration, multicycle filtration,
11 membrane fouling prediction.

12 **Highlights**

- 13 • Multicycle fouling is predicted using exponential smoothing.
- 14 • Two models best predict fouling based on observed mechanism.
- 15 • Forecasting natural water fouling was successful with small errors.
- 16 • Permeability at full-scale was forecasted for 41 days with an error of 3.2%.
- 17 • A web application was developed to help apply the new fouling prediction method.

18

19 *Corresponding author. Tel: +33549453846. E-mail: benoit.teychene@univ-poitiers.fr

20 **Abstract**

21 This study proposes the application of time series analysis and exponential smoothing (ETS)
22 to predict multicycle membrane fouling with high accuracy and great interpretability. First,
23 lab-scale filtration tests were performed with foulant surrogates to assess the potential of ETS
24 models in predicting multicycle membrane fouling. Two models, ETS(A,A,A) and
25 ETS(A,A,M), were found to be most suitable for describing the variation of transmembrane
26 pressure. The ETS(A,A,A) model showed a slight edge over the ETS(A,A,M) model when
27 reversible fouling was dominant. In contrast, the ETS(A,A,M) performed better when
28 irreversible fouling was prevalent. Then, the effectiveness of ETS models in multicycle
29 forecasting during the filtration of natural water resources was evaluated. Residual mean
30 squared errors between 0.02 bar and 0.15 bar were obtained for horizons going from 2 up to
31 10 filtration cycles. Finally, ETS models were applied to predict the variation in permeability
32 of drinking water treatment plant membrane skid. A three-month dataset was utilized to
33 predict the next 41 days of permeability variation, yielding an average error of 3.2%. Along
34 with that, a web application designed to assist membrane users in utilizing ETS was
35 developed and presented at the end. The web application can be accessed through the
36 following link: <https://fouling-ets.shinyapps.io/shiny-ets-forecasting/>.

37 **1. Introduction**

38 Membrane filtration is a widely used technology for drinking water production and
39 wastewater treatment. Despite its effectiveness, membrane fouling is a significant issue that
40 can impair filtration performances, resulting in increased operating costs, energy
41 requirements, and reduced membrane lifespan [1]. Therefore, understanding and predicting
42 membrane fouling is critical to ensure sustainable membrane filtration, and to facilitate the
43 selection of optimal pretreatment conditions to prevent or mitigate membrane fouling [2–5].

44 In recent decades, there has been a growing interest in predicting and modeling membrane
45 fouling in solid-liquid separation processes [6,7]. Phenomenological models based on
46 hydraulic resistances, transmembrane pressure (TMP), pollutant rejection and other operating
47 parameters have been developed to understand and predict membrane behaviors in solid-
48 liquid separation processes. Charfi et al. [8] developed a mathematical model that predicts
49 multicycle filtration performance based on hydraulic resistances and mass balance of foulant
50 deposition and removal during each cycle. Teychene et al. [9] constructed a model also based
51 on hydraulic resistances to predict the synergistic fouling resulting from the filtration of
52 mixtures comprising particulate matter and NOM. Similarly, Griffiths et al. [10] utilized a
53 network model based on Hermia's models to simulate fouling behavior that allows for the
54 combination of different blocking mechanisms (standard blocking, intermediate blocking...)
55 and fouling caused by particulate adsorption. The study was further extended by Sanaei and
56 Cummings [11], who incorporated the membrane morphology and investigated its impact on
57 separation efficiency and membrane fouling. The models developed in these works and others
58 found in the literature [12] present the advantage of understanding the fouling mechanisms.
59 However, they rely on multiple assumptions about membrane and foulants morphology and
60 structure. Additionally, they require information that is challenging to acquire at an industrial
61 scale which restricts their practical application [11].

62 Therefore, there is a need to develop more reliable and practical approaches that can
63 overcome the shortcomings of traditional models and accurately predict the performance of
64 membrane processes. The rapid development of artificial intelligence (AI) offers an excellent
65 opportunity to achieve this goal [6,7,13]. With the development of the concept of Industry 4.0,
66 long-term predictions of process performance and the integration of AI strategies into system
67 operations are the subject of a large number of efforts. In particular, due to the complexity of
68 membrane fouling phenomena, the scientific community is currently moving towards data-
69 driven modelling [14–26]. Two main approaches are generally used: (i) The first approach
70 involves using various water quality parameters and operating conditions as inputs for
71 artificial neural networks (ANN) or other AI techniques, which in turn provide the flux or
72 pressure variation as outputs. For instance, previous studies have employed parameters such
73 as TMP, crossflow velocity, pH, feed concentration, filtration time, and operating temperature
74 to predict variables like flux decline or pollutant rejection in the filtration of wastewater [17],
75 oily wastewater [14–16], and milk [23]. The resulting models demonstrated generally high
76 accuracy, with R^2 values exceeding 0.99. (ii) The second approach consists of predicting the
77 variation of TMP (or the permeate flux) using past events. As an example Shim et al. [19]
78 developed a Long Short Term Memory (LSTM) model to predict the value of the next
79 permeate flux (J at time t) of the filtration curves using the four previous data points (J at
80 times $t-1$, $t-2$, $t-3$, $t-4$). However, despite the advantages of such models, there is a trade-off in
81 terms of interpretability and understanding of fouling mechanisms.

82 One approach known as time series analysis may offer a solution and present a compromise
83 between the interpretability of mechanistic models and the accuracy of previously mentioned
84 data driven models. Time series analysis involves the examination of data points collected
85 sequentially over time and is usually performed by identifying patterns, trends, and other
86 features that can be used for prediction and decision-making. Similarly to the second

87 approach explained above (i.e. the LSTM model), previous data points are used to predict
88 future values and variations. Time series analysis encompasses a multitude of methods that
89 fall under the umbrella of AI-based techniques. These methods are widely utilized in diverse
90 fields, from economics to chemical processes, and has recently been applied to membrane
91 filtration [24–27]. The underlying concept behind its implementation in membrane filtration is
92 that the TMP or flux are measured sequentially over time, making it feasible to predict future
93 values based on past data. An example of this implementation was demonstrated by Ruby
94 Figueroa et al. [24]. They employed ARIMA (Auto Regressive Integrated Moving Average)
95 models to predict a 4-hour permeate flux decline in a 10-hour dead-end filtration test without
96 backwashing. In another study, Teychene et al. [26] used an ARIMA(2,1,0) model to forecast
97 the irreversible fouling rate of two underground water resources for a period of 6 months,
98 based on a 12-month database. The advantages of employing these time series models are
99 their simple interpretability, lack of requirement for feed solution or operating conditions
100 data, and their sensitivity to different fouling behaviors, indicating a potential association with
101 membrane fouling. Nevertheless, the previous studies that used ARIMA or ANN models were
102 limited by the use of data coming only from single-cycle filtration tests and did not consider
103 TMP fluctuations observed in multicycle filtration. As a result, these models have limited use
104 in long-term forecasting.

105 Therefore, the objective of this study is to expand the application of time series analysis to
106 multicycle filtration and develop an algorithm capable of predicting the performance of
107 multicycle filtration, specifically the TMP evolution. To achieve this, experimental data were
108 acquired by conducting lab-scale multicycle filtration tests on synthetic water mixtures and
109 natural water samples. Then a decomposition of filtration curves into trends and seasonal
110 components was performed to get a better understanding of these curves. Afterwards, a time
111 series analysis method called exponential smoothing (or **ETS** which stands for **Error Trend**

112 Seasonality) was applied on the experimental results. Ultimately, the ETS approach was
113 applied to industrial filtration data. The choice of the time series approach is based on its
114 ability to strike a balance between prediction accuracy and data interpretability, in addition to
115 its ability to be employed on multicycle filtration data. Moreover, ETS models were chosen
116 over ARIMA models because they are more easily generalized and do not require as many
117 assumptions or mathematical constraints (i.e. the presence of unit root [28,29]). In the
118 following, the filtration setup and operating conditions are first described. Then a theoretical
119 description of the ETS models and the developed algorithm are introduced. Finally, the
120 application of these models on multicycle filtration results obtained on various feed water is
121 presented and discussed.

122 **2. Materials and methods**

123 **2.1. Feed water and analytical tests**

124 In this study, synthetic feed waters were prepared by dissolving varying concentrations of
125 bovine serum albumin (BSA) (Sigma Aldrich A7906) (ranging from 0 to 5 mgC/L) and
126 sodium alginate (Sigma Aldrich W201502) (ranging from 0 to 5 mgC/L) mixed or not with
127 bentonite (100 mg/L, d_{50} : 2 μ m, Na-Montmorillonite, SWy-2, Clay Minerals Society) in
128 MilliQ water (pH 6.8, 18 M Ω .cm at 25 °C). A total of 17 combinations of these three
129 constituents were investigated, inducing different fouling behaviors (both reversible and
130 irreversible). Table S1 provides a detailed list of the 17 combinations investigated.
131 Additionally, surface water samples from the Boivre and Clain rivers in Poitiers, as well as
132 two ground water resources, named Nouvet and Noé exploited by Eau de Paris and used as
133 potential feed water for the “Saint- Cloud” and “L’Haÿ-les Roses” drinking water treatment
134 plants (DWTP) (ultrafiltration process), were utilized in this study. Table S2 provides quality
135 measurements for these resources.

136 To assess the quality of the water samples, the total organic carbon (TOC), UV absorbance at
137 254 nm, and turbidity were measured. The TOC analysis was carried out using a TOC-V
138 analyzer from Shimadzu (Japan), while the UV₂₅₄ was measured using a Cary 50 UV-Vis
139 spectrometer from Varian (The Netherlands). Turbidity was measured using a WTW Turb
140 550 IR Lab Turbidity Meter from WTW (Germany). Additionally, the fluorescence excitation
141 and emission matrix (FEEM) was determined using a FluoroMax 4 spectrofluorometer from
142 Horiba (France), with the excitation and emission wavelengths set to a range of 240-500 nm
143 and 290-600 nm, respectively [30].

144 **2.2. Membranes and filtration protocol**

145 Commercial polyethersulfone (PES) hollow fiber membranes (NX Filtration, Netherlands)
146 with a pore size of 0.02 μm , an internal diameter of 0.8 mm and an external diameter of 1.3
147 mm were used to prepare homemade mini-modules made of 4 fibers of 12 cm potted with
148 epoxy resin (Araldite 2011) and a total membrane surface area of approximately 12 cm^2 . The
149 pure water permeability of virgin membranes was equal to $640 \pm 90 \text{ L/h/m}^2/\text{bar}$ (LMHB) at
150 20°C. Inside-out filtration tests, lasting a total of 24 hours, were carried out on an automated
151 pilot system (as described in Touffet et al. [5]) at a flux of 100 L/h/m^2 (LMH) in one-hour
152 cycles interspersed by 2 min of hydraulic backwash at 200 LMH using the produced
153 permeate. The TMP was measured at 3-minute intervals, resulting in 20 measurement points
154 per cycle. The experiments were automatically terminated when the TMP reached its
155 threshold value of 1.5 bars (even if the 24-hour limit had not been reached). Filtration tests
156 were carried out at temperatures ranging from 18 to 22°C. To ensure the stability of the
157 filtration system, MilliQ water was filtered, and repeatability was assessed by repeating twice
158 filtration tests of randomly selected conditions. The obtained results are presented in Figure
159 S1. The filtration system was deemed stable as no TMP buildup was observed during the
160 filtration of MilliQ water for 24h (Figure S1(a)). Moreover, similar TMP variations were

161 observed for the repeated conditions, indicating a high level of repeatability. The comparison
162 of filtration curves revealed residual mean square error (RMSE) values ranging from 0.022 to
163 0.072 bars (Figures S1 (b)-(c)-(d)).

164 **2.3. Fouling indices**

165 Fouling indices are used to quantify and characterize the different fouling behavior observed
166 at lab scale during synthetic suspensions filtration. In this study, fouling indices based on the
167 work of Nguyen et al. [31] were used. These indices have been used extensively in the past
168 and were proven to be good performance indicators of multicycle filtration tests [32–34].

169 In brief, the total fouling index (TFI) and the hydraulically irreversible fouling index (HIFI)
170 were calculated according to equation 1.

$$\frac{1}{J'_s} = 1 + (FI)V_s \quad (1)$$

171 Where FI is the considered fouling index (i.e. TFI or HIFI) in units of m^2/L , V_s the specific
172 filtered volume (L/m^2) defined by the permeate volume divided by the membrane surface area
173 and J'_s is the normalized specific flux defined as described in equation 2.

$$J'_s = \frac{\left(\frac{J}{TMP}\right)_v}{\left(\frac{J}{TMP}\right)_0} \quad (2)$$

174 Thus, under constant flux conditions equation 3 is obtained

$$\frac{TMP_v}{TMP_0} = 1 + (FI)V_s \quad (3)$$

175 Where TMP_v represents the transmembrane pressure corresponding to a permeate volume V
176 and TMP_0 corresponds to the initial TMP at the start of the experiment.

177 When TMP_v and V_s are measured between two hydraulic backwashes, the fouling index FI
178 becomes the TFI and can be determined with linear regression if the variation of TMP_v/TMP_0

179 vs V_s is linear. If it is not linear then the TFI is calculated using the two-point method where
180 only the first and last point of the filtration cycle are used to calculate the slope [32]. HIFI on
181 the other hand can be calculated across multiple filtration cycles with TMP_v and V_s values
182 that corresponds to the start of every filtration cycle. Similarly, to the TFI, linear regression
183 and the two-point method are applied to calculate the HIFI.

184 **2.4. Industrial filtration data**

185 In addition to the laboratory scale filtration tests, industrial filtration data were obtained from
186 one skid (120 modules) on the "L'Hay-les-Roses" DWTP operated by Eau de Paris (maximum
187 production 150000 m³/day). These data were gathered during the period between September
188 2020 and January 2021 and pertain to an ultrafiltration process that involved different types of
189 membrane cleaning. Among the various cleaning techniques, the most significant one was the
190 chemical backwash with sodium hydroxide, followed by sulfuric acid, which was performed
191 once every 12 days. During the entire operation, the permeability was monitored at intervals
192 of 5 seconds and then reduced to one hour to decrease the size of the dataset. The
193 corresponding data are illustrated in Figure S2.

194 **3. Mathematical background and forecasting methodology**

195 **3.1. Data driven model: Exponential smoothing**

196 Exponential smoothing models are forecasting methods for time series first proposed by
197 Brown [35], Holt [36] and Winters [37]. Each model consists of an equation that describes the
198 variation of y_{t+1} (the variable that is forecasted) as a combination of three components and an
199 error. The three components are: (1) l_t the level of the time series at time t , (2) b_t the trend
200 (slope) of the time series at time t , and (3) s_{t+1-m} the seasonal component at time $t+1-m$ where
201 m represents the frequency of the seasonality (i.e. the number of acquired data in one season).

202 The error ε_t is considered to be normally distributed (i.e. the mean equals to 0 and the
203 variance σ^2 is constant over the considered period).

204 The level and the error components are always present in the model whereas the trend and
205 seasonal components may be either present or absent depending on the considered model
206 [38]. In general, ETS models' taxonomy are defined by three letters: The first letter represents
207 the error component which can be additive (A) or multiplicative (M), the second letter
208 represents the trend which can be absent (N for none present), additive (A), additive and
209 damped (A_d) or multiplicative (M) and the third letter represent the seasonal component
210 which can be absent (N), additive (A) or multiplicative (M). For example, an ETS(A,A,M)
211 model has an additive error, an additive trend and a multiplicative seasonal component. It is
212 written as follows (Equations 4-7):

$$y_{t+1} = (l_t + b_t)s_{t+1-m} + \varepsilon_{t+1} \quad (4)$$

$$l_t = l_{t-1} + b_{t-1} + \alpha\varepsilon_t/s_{t-m} \quad (5)$$

$$b_t = b_{t-1} + \beta\varepsilon_t/s_{t-m} \quad (6)$$

$$s_t = s_{t-m} + \gamma\varepsilon_t/(l_{t-1} + b_{t-1}) \quad (7)$$

213 where $\varepsilon_t = y_t - \hat{y}_t$ and \hat{y}_t being the predicted value at time t.

214 In equation 4, the trend component (b_t) is added to the level component (l_t) whereas the
215 seasonal component (s_{t+1-m}) is multiplied to the other components. Also, the error is added to
216 the equation. The level at time t (l_t) (equation 5), is the sum of the previous level value (l_{t-1}),
217 the value added due to the slope (b_{t-1}) and an error normalized to the seasonal component
218 ($\alpha\varepsilon_t/s_{t-m}$). The other equations (5 and 6) can be understood in a similar fashion. Finally,
219 α, β and γ are three parameters that varies between 0 and 1. They represent the degree of
220 variation between the modeled value and the real value and are estimated using the training
221 dataset. Values that are close to 1 indicate a larger impact of recent observations on future

222 data (i.e. the last recorded values have a strong impact on future variations, indicating local
223 adaptation for these three components). Conversely, values close to 0 show that older values
224 govern the prediction (i.e. the level, the trend and the seasonal components are global). In
225 total, there are 24 ETS models that corresponds to all the different combinations of error,
226 trend and seasonality. More details about exponential smoothing and innovative state space
227 model (i.e. ETS models) can be found in literature [38,39].

228 **3.2. Forecasting algorithm**

229 In order to forecast the TMP evolution in real time and over multiple filtration cycles a
230 specific algorithm was developed (Figure S3). After acquiring TMP data for a minimal of
231 three cycles (in order to set the seasonal component), an ETS model is selected (choice
232 criteria of selected model to be explained later in section 4.3) and applied to forecast multiple
233 cycles into the future. When a new filtration cycle ends, the entire methodology is repeated.
234 This algorithm utilizes past TMP data as input and predicts future TMP variations as output.
235 The choice of the number of observations for input and output data can be tailored as
236 required. The ETS models and the forecasting algorithm were generated via the R software
237 and the “fpp2” package [40]. This package employs the maximum likelihood to estimate the
238 smoothing parameters α , β and γ , and the initial states l_0 , b_0 , s_0 , s_{-1} , \dots , s_{-m+1} . The accuracy of
239 the predictions is assessed through the calculation of R^2 and RMSE.

240 **4. Results and discussion**

241 **4.1. Filtration performances**

242 The experimental conditions were specifically designed to simulate different fouling
243 behaviors (reversible and irreversible fouling) with varying intensities. To achieve this, three
244 fouling surrogates were tested alone and in different mixtures under constant filtration
245 conditions. The normalized TMP variation ($P-P_0$), along with the corresponding fouling
246 indices TFI and HIFI, are reported for the investigated conditions in Figures 1 and 2.

247 The filtration of alginate alone (Figure 1(a)), resulted in high irreversible fouling while
248 reversible fouling was smaller and related to the investigated feed concentration. This finding
249 is consistent with the results reported by Chang et al. 2016 [41] when using the permeate for
250 backwash on PES membranes. Moreover, the maximum TMP of 1.5 bar was reached after
251 19.3 hours of filtration (corresponding to a specific filtered volume of 1930 L/m^2) for 1 mg/L
252 of alginate while higher concentrations of 2 mg/L and 5 mg/L stopped at 11.8 hours (1180
253 L/m^2) and 4.9 hours (490 L/m^2), respectively. Both HIFI and TFI increased with alginate
254 concentration (Figure 2(a)). Specifically, when 1 mg/L of alginate was used, the HIFI and TFI
255 values were $0.003 \text{ m}^2/\text{L}$ and $0.005 \text{ m}^2/\text{L}$, respectively. These values rose to $0.016 \text{ m}^2/\text{L}$ and
256 $0.029 \text{ m}^2/\text{L}$ for 5 mg/L of alginate, respectively. Regarding rejection, alginate was partially
257 rejected by the membrane as TOC rejection varied between 50% and 75%.

258 Upon filtration of BSA alone, a combination of reversible and irreversible fouling was
259 observed with an increase in reversible fouling over time, as illustrated in Figure 1(b). At the
260 start of the filtration process, a sudden rise in normalized TMP variation ($P-P_0$) was observed,
261 which can be attributed to pore blocking and/or adsorption of BSA on PES membranes as
262 seen in previous works [42,43]. The fouling intensity was found to increase with BSA
263 concentration, as indicated by the TMP reaching 1.5 bars after 7.5 hours (750 L/m^2) and 24

264 hours for concentrations of 5 mg/L and 2 mg/L, respectively, while the corresponding TMP
265 value for 1 mg/L was only 0.62 bar after 24 hours of filtration. HIFI and TFI (Figure 2(a))
266 increased as the concentration of BSA increased, from 0.002 m²/L and 0.008 m²/L at 1 mg/L
267 to 0.009 m²/L and 0.027 m²/L at 5 mg/L. Finally, the TOC rejection of BSA averaged around
268 73% whereas the UV₂₅₄ rejection averaged around 50%. The removal of BSA was further
269 confirmed through the FEEM spectra shown in figure S4 (a) and (b).

270 The mixture of alginate and BSA (figure 1(c)) showed a synergic effect as it led to only
271 irreversible fouling that increased with time. This implies that the backwash efficiency is
272 significantly reduced when the two compounds are combined, due to potential adsorption of
273 the foulants on the membrane surface. Similar results on polysaccharide–protein mixtures on
274 ultrafiltration were reported by Susanto et al. 2008 [44] who justified the synergy between
275 alginate and BSA through intermolecular binding. The feed mixture concentration obviously
276 influences the TMP variation. As observed in the 1 mg/L BSA and 1 mg/L alginate mixture,
277 the maximum TMP was achieved after 14.5 h (1450 L/m²), compared to 8.5 h (850 L/m²)
278 when the concentrations of the two components were doubled. This influence is further
279 evident in the two mixtures tested, one containing 2 mg/L BSA and 1 mg/L alginate, and the
280 other containing 1 mg/L BSA and 2 mg/L alginate (figure S5 (a)). Both mixtures exhibited
281 similar TMP variations, reaching 1.5 bars after around 10 hours (1000 L/m²), which falls
282 between the values observed in figure 1(c)). To facilitate further explanations these two
283 conditions, which exhibit the same variation as the curves in figure 1(c), were excluded from
284 the rest of the study. Moreover, HIFI and TFI increased when doubling the concentrations of
285 alginate and BSA from 0.005 m²/L for both of them to 0.011 m²/L and 0.018 m²/L
286 respectively (figure 2(a)). In terms of rejection, TOC and UV₂₅₄ rejections ranged from 50%
287 to 70%. Additionally, the FEEM spectra indicated the elimination of BSA from the mixture,
288 as shown in Figures S6 (a) and (b).

289 The addition of 100 mg/L of bentonite to the previous investigated suspensions (Figures 1(d),
290 1(e) and 1(f)) resulted in mostly reversible fouling with larger TMP variations during
291 filtration. Filtration of bentonite suspension alone did not lead to membrane fouling as seen in
292 Figure S5 (b) and had the lowest TFI and HIFI (Figure 2(a)). Moreover, the mixture of
293 bentonite with alginate showed smaller irreversible fouling compared to alginate alone,
294 whereas the addition of bentonite to BSA strongly reduced total fouling resistance and led to
295 highly reversible fouling. The effect of bentonite may be attributed to the adsorption of
296 alginate and BSA on suspended clay particles. This phenomenon has been reported in
297 previous papers [45–47], and it is believed to reduce the adsorption of alginate and BSA on
298 the membrane surface, thereby reducing irreversible fouling. The addition of bentonite had a
299 clear impact on both HIFI and TFI, as demonstrated in Figure 2(a). The HIFI values were
300 notably reduced, and were significantly smaller than the TFI values. The impact of increasing
301 concentrations of alginate and BSA was mostly reflected in the increase of TFI (figure 2(a)).
302 For instance, when the concentration of alginate increased from 2 mg/L to 5 mg/L, the TFI
303 increased from 0.039 m²/L to 0.080 m²/L. Similarly, for the BSA and bentonite mixture, the
304 TFI increased from 0.010 m²/L to 0.015 m²/L for BSA concentrations of 2 mg/L and 5 mg/L.
305 Finally, when the concentrations of alginate and BSA were doubled (2 mg/L alginate + 2
306 mg/L BSA + 100 mg/L bentonite), TFI increased from 0.023 m²/L to 0.040 m²/L. Regarding
307 membrane rejection, the permeate turbidity consistently remained below 0.5 NTU, which
308 corresponds to rejection rates of over 99% since 100 mg/L of bentonite lead to an average
309 turbidity of 50 NTU. Also, the presence of bentonite did not appear to have any major effect
310 on the removal of TOC and UV₂₅₄, as their values were similar to those discussed previously.

311

312

313 Upon closer analysis of TFI and HIFI, the filtration curves can be divided into two categories.
314 In the absence of bentonite (Figure 2(a)), HIFI tends to increase with increasing
315 concentrations and is relatively large compared to TFI, indicating that irreversible fouling is
316 the major contributor to the global TMP variation. However, with the addition of bentonite,
317 HIFI is no longer affected by foulant concentrations and becomes relatively insignificant
318 compared to TFI, indicating that reversible fouling becomes much more dominant.
319 Furthermore, the boxplot in Figure 2(b) demonstrates that TFI tends to increase over
320 successive filtration cycles in the absence of bentonite. This is illustrated by the relatively
321 larger length of the boxplots and a wider dispersion of the orange dot points, which represent
322 individual TFI values for every cycle. Conversely, in the presence of bentonite, TFI values
323 remain more constant, which is evident by the smaller length of the boxplots.

324 To summarize, the filtration curves obtained from the experiments without bentonite can be
325 classified as the first category. In this category, irreversible fouling plays a significant role in
326 total fouling. Due to this, total fouling increases gradually with every filtration cycle. On the
327 other hand, the second category corresponds to the filtration of samples with bentonite where
328 the total fouling is predominantly reversible and remains constant, while irreversible fouling
329 is minimal.

330 **4.2. Time series analysis of filtration curves**

331 As mentioned earlier, since the filtration curves (TMP evolution data) are time dependent,
332 they can be analyzed as time series. Nonetheless, in order to insure the time dependency of
333 TMP data, it is necessary to verify the presence of a correlation between the TMP at a specific
334 time and certain previous TMP values. To do so, the autocorrelation function (ACF) is
335 calculated. Figure 3 presents an example for the curve obtained from the filtration of 2 mg/L
336 alginate and 100 mg/L bentonite, where the correlations between any TMP and the previous
337 200 TMP measurements (equivalent to 10 filtration cycles) are calculated. The highest ACF

338 value of 0.88 was observed between TMP values at time t and $t-1$ (lag = 1). This correlation
339 decreased until it reached the red dotted lines after 8 filtration cycles, below which the
340 correlation was no longer significantly different from zero. Thus, for this filtration curve,
341 TMP values were dependent on the variation of TMP from the last 8 cycles. Furthermore, it is
342 evident that the ACF decrease is not gradual. Instead, there are cyclic spikes in ACF values at
343 multiples of 20 lags, indicating that there is a stronger correlation between TMP values that
344 are obtained at the same time in the previous filtration cycles. Similar results were found for
345 the other curves and some examples are presented in the figure S7. In general, when data
346 exhibit a trend, autocorrelations for small lags tend to be large and positive as observations
347 that are close in time are also similar in size. Hence, ACF plots of trended time series
348 typically have slowly decreasing positive values as lags increase. When data are seasonal, the
349 autocorrelations are higher for the seasonal lags, which are multiples of the seasonal
350 frequency (20 in this case), than for other lags. If the data exhibit both a trend and a seasonal
351 cycle, a combination of these effects is observed. Therefore, the ACF plots support the idea of
352 treating filtration curves as time series with both trend and seasonal cycles.

353 These components can be visualized through various decomposition techniques. In figure 4,
354 the trend and seasonal components as well as the remainder component are obtained using a
355 classical additive decomposition method. The trend component (Figure 4(b)) which reflects
356 the general movement of the data is seen to be increasing. Its variation is linked to irreversible
357 fouling. The seasonal component (Figure 4(c)) reflects the repeatable pattern between various
358 seasons. In the case of multicycle filtration, it is related to the increase in TMP between the
359 start of every cycle and its end. Finally, the remainder component (Figure 4(d)) represents the
360 unexplained portion of the curve i.e. the remaining data after removing the trend and seasonal
361 components.

362 As mentioned in the introduction, time series analysis contributes to improving the
363 interpretability and getting a better understanding of filtration curves. Tools such as ACF
364 plots and time series decomposition play pivotal roles in achieving this goal. For instance,
365 when examining the ACF of a filtration curve, significant correlations at specific lags signal
366 robust dependencies between values at those particular time points. In the examples given in
367 this section, strong correlations predominantly aligned with the frequency of hydraulic
368 backwash, underscoring the pivotal role of backwashes in shaping curve patterns. However, if
369 another influencing factor, such as water quality, exhibits consistent variation at a specific
370 frequency, the ACF plot can reveal its impact distinct from backwash effects. High
371 correlations at lags unrelated to backwash frequency indicate the presence of an additional
372 factor, potentially unknown, influencing the filtration curve's variation. Similarly, through
373 time series decomposition, it's possible to get a better understanding, by distinguishing
374 between the impacts of seasonal occurrences, the overall trend of the curve, irregularities and
375 anomalies, and the unexplained variability or noise in the data.

376 **4.3. ETS modeling of TMP during multicycle filtration tests**

377 After confirming the feasibility of handling filtration curves as times series data with trends
378 and seasonal components, it becomes possible to apply time series modeling techniques on
379 filtration curves. One of the most common methods is the previously explained exponential
380 smoothing. At this stage, it was crucial to assess which ETS models have the ability to
381 effectively model and forecast filtration curves. To achieve this, the filtration curves were
382 divided into training and testing datasets, with the former comprising the initial 75% of the
383 data points. Viable models were then tested on the training portion of each curve, and
384 subsequently used to forecast the testing data, with the corresponding RMSE being calculated.
385 The viability of ETS models was decided by restricting the choice to certain models that are
386 relevant to multicycle filtration TMP data. Firstly, models containing a multiplicative trend

387 were removed as they generally provide poor forecasts [38]. Secondly, models with a damped
388 trend were eliminated since they are typically used when the trend is expected to become a
389 flat line in the long term. However, since it is known that the TMP will most likely continue
390 to increase during filtration, the possibility of a damped trend can be dismissed. Models
391 having no trends are eliminated for the same reason. Finally, models with no seasonality were
392 removed as the presence of seasonal variations in filtration curves was confirmed in the
393 previous section (4.2).

394 Thus, the remaining four models (ETS(A,A,A), ETS(A,A,M), ETS(M,A,A) and
395 ETS(M,A,M)) were tested and compared. The corresponding RMSEs are given in Figure S8.
396 As shown, several models are capable of producing accurate forecasts for each curve. Some
397 curves, such as those involving mixtures of alginate, BSA, and bentonite, can be forecasted
398 with small errors by all models while other curves had varying RMSE values depending on
399 the selected models, such as the examples of 2 mg/L of BSA or 2 mg/L of alginate and 100
400 mg/L of bentonite. Figure S9, S10 and S11 display the different forecasts for these curves.

401 The selection of models can be made easier by narrowing down the options to just two
402 models, which can generally describe most of the possible TMP variations. These two models
403 are the ETS(A,A,M) and ETS(A,A,A) models presented respectively in equations 4 and 8.

$$\text{ETS(A,A,A)} : \quad y_{t+1} = l_t + b_t + s_{t+1-m} + \varepsilon_{t+1} \quad (8)$$

404 Both models, ETS(A,A,A) and ETS(A,A,M), share similar components as they both include
405 an additive error and an additive trend which are commonly found in time series [38]. The
406 additive trend is denoted by the "+" sign between l_t and b_t , while the additive error is
407 represented by the "+" sign between ε_t and the remaining equation. However, the main
408 difference between the two models is the type of seasonal component used, which can be
409 either additive or multiplicative. Mathematically, a multiplicative seasonality suggests a

410 relationship between the seasonal component s_t and the overall trend of the curve represented
411 by $l_t + b_t$. This relationship indicates that the seasonality of the curve will increase or decrease
412 in amplitude based on the overall trend. Conversely, an additive seasonality indicates that the
413 overall seasonality is independent of the values of $l_t + b_t$.

414 As previously mentioned, the filtration curves can be divided into two categories: the first
415 category with TFI values that increase with the accumulation of irreversible fouling, and the
416 second category where TFI values remain constant regardless of the irreversible fouling. By
417 referring to the theoretical explanation of ETS equations, it is possible to assign the
418 ETS(A,A,M) model to the filtration curves of the first category since there is a relationship
419 between the curve's trend (i.e., the evolution of irreversible fouling) and the seasonal variation
420 represented by the TFI. Similarly, the ETS(A,A,A) can be assigned to the curves of the
421 second category where total fouling is independent of irreversible fouling. The remaining two
422 models (ETS(M,A,A) and ETS(M,A,M)) can still be used, nonetheless, they were disregarded
423 as they do not present any additional information and have larger confidence intervals for
424 their forecasts compared to the ETS(A,A,A) and ETS(A,A,M) models. For that, when dealing
425 with new, unseen data, it is advisable to narrow down the model options to ETS(A,A,A) and
426 ETS(A,A,M) models only. The versatility of ETS models in adapting to fouling behavior
427 represents an additional aspect of interpretability achieved through time series analysis. In
428 cases where fouling behavior is not as evident as the examples illustrated in this study, both
429 ETS models can be tested. The model with superior performance provides insight into the
430 dominant fouling behavior.

431 Lastly, it's important to note that apart from ETS models, ARIMA models were also
432 examined. In general, good forecasts were achieved with ARIMA, which isn't surprising since
433 many ETS and ARIMA models share an equivalence relationship [38]. Nonetheless, ARIMA
434 models encounter certain mathematical constraints, particularly concerning stationarity [28,

435 29]. These limitations make it challenging to apply a single ARIMA model universally across
436 all filtration curves. For these reasons, the decision was made to focus solely on ETS models
437 in this study and exclude ARIMA models.

438 **4.4. Multicycle forecasting of natural water filtration**

439 Based on these findings, it is now feasible to apply the forecasting algorithm to predict the
440 variation in TMP during the filtration of natural water. Figure 5 demonstrates this application
441 using Nouvet resource filtration as an example. Specifically, after 6 hours of filtration (Figure
442 5(a)), an ETS(A,A,M) model was developed from the recorded data. The training data is
443 represented in black, and the model fit is depicted in red. Utilizing this model, 6-hour
444 forecasts were generated (depicted in yellow) with 95% prediction intervals. These forecasts
445 were then compared to filtration data recorded between 6 and 12 hours (shown in grey),
446 serving as testing data. This comparison yielded an RMSE of 0.09 bar and an R^2 of 0.83 as
447 displayed in figure 5(b).

448 The same analysis was replicated, and ETS(A,A,M) models were created for training data
449 corresponding to 12 and 18 hours of filtration (Figure 5(c) and (e)). Similarly, forecasts were
450 produced for the subsequent 6 hours and compared with testing data spanning the same
451 period. Remarkably low forecasting errors persisted even after 6 hours, with RMSE values of
452 0.07 and 0.08 bar and R^2 values of 0.91 and 0.82 observed at 12 and 18 hours, respectively
453 (Figure 5(d) and (f)).

454 It's important to highlight that both ETS(A,A,M) and ETS(A,A,A) models were evaluated, but
455 ETS(A,A,M) was chosen due to its superior performance with lower errors. This selection
456 aligns with the earlier discussion, since the filtration curve displayed signs of irreversible
457 fouling and an escalating total fouling.

458 This implementation can be viewed as a real-time forecasting application during filtration on
459 an industrial scale. ETS models are established using available data to predict future TMP
460 variations. The testing data here simulates unseen future data that would be acquired if no
461 intervention is made. Moreover, unlike the earlier example where models and forecasts were
462 generated at fixed intervals (6, 12, and 18 hours), in real application, these models and
463 corresponding forecasts will be generated at the conclusion of each new filtration cycle. As
464 these models can be generated within seconds, real-time forecasting of TMP data becomes
465 feasible.

466 In addition to that, the current analysis involved a 6-cycle filtration horizon. However, it is
467 feasible to generate forecasts with shorter or longer horizons. Typically, shorter forecast
468 horizons exhibit lower errors and have narrower prediction intervals. Conversely, as the
469 forecast horizon increases, errors tend to increase while prediction intervals grow rapidly
470 wider. The filtration curve depicted in figure 5 did not exhibit this phenomenon; however, it
471 can be distinctly evident in other scenarios, as illustrated in figure S12.

472 An assessment was conducted on the impact of increasing horizons on the RMSE for all four
473 natural water resources. To this end, forecasts were generated for horizons spanning from 2 to
474 10 filtration cycles and their corresponding RMSE were computed. Figure S13 depicts the
475 resulting RMSE for different forecasts using models created from 6, 12, and 18 filtration
476 cycles. Consistent with expectations, the error in the forecasts increased with the horizon,
477 although remaining below 0.15 bar even for 10-cycle forecasts. Nevertheless, it is
478 recommended to employ relatively short forecasting horizons (not exceeding six filtration
479 cycles in our specific filtration conditions) based on the observed escalation in RMSE with
480 increasing horizons.

481 **4.5. Bridging the gap: from laboratory to industrial-scale filtration**

482 Having established the feasibility of multicycle filtration on laboratory-scale filtration curves,
483 the scope must now be broadened to include the application of ETS to industrial data. This
484 section will provide a brief overview of an industrial filtration data application, followed by
485 an introduction to a forecasting tool that simplifies the use of ETS with new data.

486 **4.5.1. Application on industrial filtration data**

487 The application of ETS forecasting was previously demonstrated on TMP seasonal data,
488 which consisted of filtration cycles separated by hydraulic backwashes. However, this
489 methodology can also be applied to permeability data and chemical cleaning cycles. In this
490 regard, the ultrafiltration industrial data presented earlier was used for ETS forecasting. This
491 dataset presents a significant upgrade compared to lab-scale curves, as it includes 120
492 industrial filtration modules, making it inherently noisier. Additionally, fluctuations in water
493 quality over the four-month period contribute to much more variations in the curve.
494 Specifically, during this period, the TOC varied between 0.2 mgC/L and 0.49 mgC/L, the
495 UV_{254nm} ranged between 0.005 cm^{-1} and 0.017 cm^{-1} and the turbidity varied between 0.11
496 NTU and 7.58 NTU.

497 As illustrated in figure 6, the model was trained using five filtration cycles of 12 days
498 measured between 19/09/2020 and 25/11/2020 (the seasonality was confirmed by an ACF
499 plot (data not shown). The ETS(A,A,M) model was selected, although both models gave
500 similar results. Forecasts were generated until 05/01/2021, which was 41 days later, and the
501 results demonstrated excellent forecasting ability with an RMSE of 29.9 LMHB which
502 corresponds to a mean average percentage error (MAPE) equal to 3.2%, as depicted in figure
503 S14. During the seventh cycle, a discrepancy is noticeable between the forecasted and actual
504 permeability values. The forecasts failed to capture the stabilization of permeability observed
505 in the filtration cycle starting on 08/12/2020. These disparities could be attributed to

506 fluctuations in water quality. As a result, ongoing efforts are focused on integrating water
507 quality parameters into the forecasts to more accurately account for such variations.

508 Finally, it is important to note that the original data was preprocessed by reducing the number
509 of data points and smoothed by a moving average of order 9 prior to the application of ETS
510 forecasting to reduce the noise in the data.

511 **4.5.2. Online tool for ETS forecasting**

512 The ability to forecast TMP (or permeability) variations for multiple filtration cycles presents
513 several opportunities for industrial applications that could reduce fouling and prolong
514 membrane life. One such application could be the planning of preventive membranes
515 cleaning, which could be performed when forecasts exceed a predetermined TMP threshold.
516 To facilitate this process, a user-friendly web application was developed to help membrane
517 users apply ETS forecasting to their filtration data. This application allows users to upload
518 their data, create ETS models and customized forecasts, and download their results. The web
519 application and tutorial are accessible via the following link: [https://fouling-
520 ets.shinyapps.io/shiny-ets-forecasting/](https://fouling-ets.shinyapps.io/shiny-ets-forecasting/).

521 **5. Conclusion**

522 In this study, time series analysis was employed to analyze the variation of TMP during
523 multicycle filtration. It was first demonstrated that filtration curves can be decomposed into
524 trends and seasonal components, which enhances their comprehension. Then, TMP values
525 were forecasted using exponential smoothing by utilizing past variations. The results
526 indicated the prevalence of two distinct models: ETS(A,A,A) was attributed to filtration
527 curves dominated by reversible fouling, whereas ETS(A,A,M) was found to be more suitable
528 to forecast curves with irreversible fouling.

529 The models were applied on filtration curves obtained from lab-scale filtration of natural
530 water resources, resulting in errors ranging from 0.02 bar to 0.15 bar depending on the
531 forecast horizon. These models were then used to forecast the decrease in permeability of full-
532 scale industrial membrane skid (120 modules), using three months of filtration data to predict
533 41 days of permeability variation with an RMSE of 29.9 LMHB (MAPE of 3.2%). Finally, a
534 user-friendly web application was developed to assist membrane users in utilizing ETS
535 forecasting. Works are currently being done to implement the methodology and use the
536 resulting forecasts as a preventive technique to effectively adjust feed water pretreatment
537 conditions and optimize cleaning frequency.

538

539 **Acknowledgement**

540 The authors acknowledge financial support from the European Union (ERDF) and Région
541 Nouvelle Aquitaine. This work pertains to the French government program “Investissement
542 d’Avenir” (EUR INTREE, reference ANR-18-EURE-0010). The authors would also like to
543 thank NX Filtration for providing the membranes and the staff of Eau de Paris in charge of
544 water sampling and operating the ultrafiltration DWTPs.

545 **Author Contribution**

546 **G. Dagher:** Methodology, Experimentation, Analysis, Writing. **A. Martin:** Industrial dataset
547 access, Review and Editing. **JM. Galharret, L. Moulin and JP. Croué:** Review and Editing.
548 **B. Teychene:** Conceptualization, Resources, Methodology, Supervision, Review and Editing.

549

550

551

552

553

554 **6. References**

555 [1] R. H. Peiris, M. Jaklewicz, H. Budman, R. L. Legge, et C. Moresoli, Assessing the
556 role of feed water constituents in irreversible membrane fouling of pilot-scale ultrafiltration
557 drinking water treatment systems, *Water Res.*, 47, no 10, (2013), doi:
558 10.1016/j.watres.2013.03.015.

559 [2] A. Touffet, J. Baron, B. Welte, M. Joyeux, B. Teychene, et H. Gallard, Impact of
560 pretreatment conditions and chemical ageing on ultrafiltration membrane performances.
561 Diagnostic of a coagulation/adsorption/filtration process, *J. Membr. Sci.*, 489, 284- 291,
562 (2015), doi: 10.1016/j.memsci.2015.04.043.

563 [3] A. Touffet, Impact de la qualité de la ressource, des prétraitements et des lavages
564 chimiques, These de doctorat, Poitiers, 2014. Consulté le: 26 février 2021. [En ligne].
565 Disponible sur: <https://www.theses.fr/2014POIT2340>

566 [4] C. D. Peters, T. Rantissi, V. Gitis, et N. Hankins, Retention of natural organic matter
567 by ultrafiltration and the mitigation of membrane fouling through pre-treatment, membrane
568 enhancement, and cleaning - A review, *J. Water Process Eng.*, 44, 102374, (2021), doi:
569 10.1016/j.jwpe.2021.102374.

570 [5] W. Gao et al., Membrane fouling control in ultrafiltration technology for drinking
571 water production: A review, *Desalination*, 272, no 1, Art. no 1, (2011), doi:
572 10.1016/j.desal.2011.01.051.

- 573 [6] M. Bagheri, A. Akbari, et S. A. Mirbagheri, Advanced control of membrane fouling in
574 filtration systems using artificial intelligence and machine learning techniques: A critical
575 review, *Process Saf. Environ. Prot.*, 123, 229- 252, (2019), doi: 10.1016/j.psep.2019.01.013.
- 576 [7] C. Niu, X. Li, R. Dai, et Z. Wang, Artificial intelligence-incorporated membrane
577 fouling prediction for membrane-based processes in the past 20 years: A critical review,
578 *Water Res.*, 216, 118299, (2022), doi: 10.1016/j.watres.2022.118299.
- 579 [8] A. Charfi, H. Jang, et J. Kim, Modelling tool to assess membrane regeneration by
580 periodical hydraulic cleaning and fouling control in pressurized membrane process for surface
581 water treatment, *Environ. Earth Sci.*, 78, no 1, Art. no 1, (2019), doi: 10.1007/s12665-018-
582 8011-4.
- 583 [9] B. Teychene, G. Collet, et H. Gallard, Modeling of combined particles and natural
584 organic matter fouling of ultrafiltration membrane, *J. Membr. Sci.*, vol. 505, (2016), doi:
585 10.1016/j.memsci.2016.01.039
- 586 [10] I. M. Griffiths, A. Kumar, et S. Stewart, A combined network model for membrane
587 fouling, *J. Colloid Interface Sci.*, 432, 10- 18, (2014), doi: 10.1016/j.jcis.2014.06.021.
- 588 [11] Sanaei et L. J. Cummings, Flow and fouling in membrane filters: effects of membrane
589 morphology, *J. Fluid Mech.*, 818, 744- 771, (2017), doi: 10.1017/jfm.2017.102.
- 590 [12] E. Iritani, A Review on Modeling of Pore-Blocking Behaviors of Membranes During
591 Pressurized Membrane Filtration, *Dry. Technol.*, 31, no 2, 146- 162, (2013), doi:
592 10.1080/07373937.2012.683123.
- 593 [13] M.-J. Corbatón-Báguena, M.-C. Vincent-Vela, J.-M. Gozávez-Zafrilla, S. Álvarez-
594 Blanco, J. Lora-García, et D. Catalán-Martínez, Comparison between artificial neural

595 networks and Hermia's models to assess ultrafiltration performance, *SePurif. Technol.*, 170,
596 434- 444, (2016), doi: 10.1016/j.seppur.2016.07.007.

597 [14] R. Badrnezhad et B. Mirza, Modeling and optimization of cross-flow ultrafiltration
598 using hybrid neural network-genetic algorithm approach, *J. Ind. Eng. Chem.*, 20, no 2, Art. no
599 2, (2014), doi: 10.1016/j.jiec.2013.05.012.

600 [15] R. Soleimani, N. A. Shoushtari, B. Mirza, et A. Salahi, Experimental investigation,
601 modeling and optimization of membrane separation using artificial neural network and multi-
602 objective optimization using genetic algorithm, *Chem. Eng. Res. Des.*, 91, no 5, Art. no 5,
603 (2013), doi: 10.1016/j.cherd.2012.08.004.

604 [16] B. Rahmanian, Application of experimental design approach and artificial neural
605 network (ANN) for the determination of potential micellar-enhanced ultrafiltration process.
606 *Journal of Hazardous materials*, (2011), vol. 187, no 1-3, 67-74,
607 <https://doi.org/10.1016/j.jhazmat.2010.11.135>

608 [17] S. A. Mirbagheri, M. Bagheri, Z. Bagheri, et A. M. Kamarkhani, Evaluation and
609 prediction of membrane fouling in a submerged membrane bioreactor with simultaneous
610 upward and downward aeration using artificial neural network-genetic algorithm, *Process Saf.*
611 *Environ. Prot.*, 96, 111- 124, (2015), doi: 10.1016/j.psep.2015.03.015.

612 [18] S. Ghandehari, M. M. Montazer-Rahmati, et M. Asghari, A comparison between semi-
613 theoretical and empirical modeling of cross-flow microfiltration using ANN, *Desalination*,
614 277, no 1, Art. no 1, (2011), doi: 10.1016/j.desal.2011.04.057.

615 [19] J. Shim, S. Park, et K. H. Cho, Deep learning model for simulating influence of natural
616 organic matter in nanofiltration, *Water Res.*, 197, 117070, (2021), doi:
617 10.1016/j.watres.2021.117070.

- 618 [20] F. Chen et al., Pilot-scale investigation of drinking water ultrafiltration membrane
619 fouling rates using advanced data analysis techniques, *Water Res.*, 48, 508- 518, (2014), doi:
620 10.1016/j.watres.2013.10.007.
- 621 [21] F. Harrou, T. Cheng, Y. Sun, T. Leiknes, et N. Ghaffour, A Data-Driven Soft Sensor
622 to Forecast Energy Consumption in Wastewater Treatment Plants: A Case Study, *IEEE Sens.*
623 *J.*, 21, no 4, Art. no 4, (2021), doi: 10.1109/JSEN.2020.3030584.
- 624 [22] T. Cheng, F. Harrou, F. Kadri, Y. Sun, et T. Leiknes, Forecasting of Wastewater
625 Treatment Plant Key Features Using Deep Learning-Based Models: A Case Study, *IEEE*
626 *Access*, 8, 184475- 184485, (2020), doi: 10.1109/ACCESS.2020.3030820.
- 627 [23] M. A. Razavi, A. Mortazavi, et M. Mousavi, Dynamic modelling of milk ultrafiltration
628 by artificial neural network, *J. Membr. Sci.*, 220, no 1, Art. no 1, (2003), doi: 10.1016/S0376-
629 7388(03)00211-4.
- 630 [24] R. Ruby-Figueroa, J. Saavedra, N. Bahamonde, et A. Cassano, Permeate flux
631 prediction in the ultrafiltration of fruit juices by ARIMA models, *J. Membr. Sci.*, 524,
632 108- 116, (2017), doi: 10.1016/j.memsci.2016.11.034.
- 633 [25] J. Wu, Improving Membrane Filtration Performance Through Time Series Analysis.
634 (2021). doi: 10.21203/rs.3.rs-1090774/v1.
- 635 [26] B. Teychene, A. Touffet, J. Baron, B. Welte, M. Joyeux, et H. Gallard, Predicting of
636 ultrafiltration performances by advanced data analysis, *Water Res.*, 129, 365- 374, févr. 2018,
637 doi: 10.1016/j.watres.2017.11.023.
- 638 [27] B. Deng, Y. Deng, M. Liu, Y. Chen, Q. Wu, et H. Guo, Integrated models for
639 prediction and global factors sensitivity analysis of ultrafiltration (UF) membrane fouling:

640 statistics and machine learning approach, *SePurif. Technol.*, 313, 123326, (2023), doi:
641 10.1016/j.seppur.2023.123326.

642 [28] R. J. Hyndman et Y. Khandakar, Automatic Time Series Forecasting: The forecast
643 Package for R, *J. Stat. Softw.*, 27, 1- 22, (2008), doi: 10.18637/jss.v027.i03.

644 [29] D. Kwiatkowski, C. B. Phillips, Schmidt, et Y. Shin, Testing the null hypothesis of
645 stationarity against the alternative of a unit root: How sure are we that economic time series
646 have a unit root?, *J. Econom.*, 54, no 1, Art. no 1, (1992), doi: 10.1016/0304-4076(92)90104-
647 Y.

648 [30] G. Coble, Characterization of marine and terrestrial DOM in seawater using
649 excitation-emission matrix spectroscopy », *Mar. Chem.*, 51, no 4, 325- 346, (1996), doi:
650 10.1016/0304-4203(95)00062-3.

651 [31] A. H. Nguyen, J. E. Tobiason, et K. J. Howe, Fouling indices for low pressure hollow
652 fiber membrane performance assessment, *Water Res.*, 45, no 8, Art. no 8, (2011), doi:
653 10.1016/j.watres.2011.02.020.

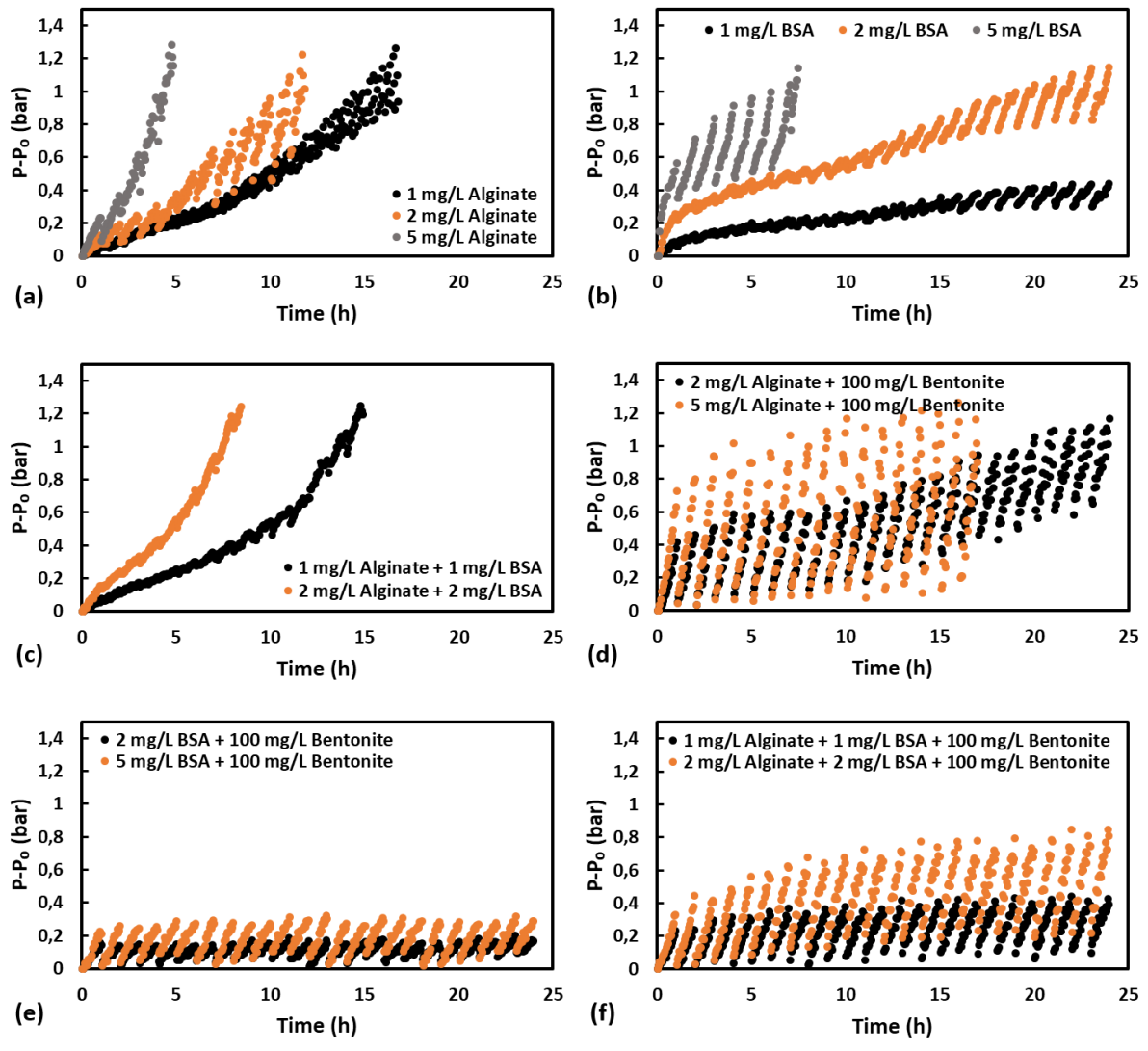
654 [32] M. Zupančič, D. Novak, J. Diaci, et I. Golobič, An evaluation of industrial
655 ultrafiltration systems for surface water using fouling indices as a performance indicator,
656 *Desalination*, 344, 321- 328, (2014), doi: 10.1016/j.desal.2014.04.002.

657 [33] M. Vera et al., Monitoring UF membrane performance treating surface-groundwater
658 blends: Limitations of FEEM-PARAFAC on the assessment of the organic matter role, *Chem.*
659 *Eng. J.*, 317, 961- 971, (2017), doi: 10.1016/j.cej.2017.02.081.

660 [34] C. Ayache et al., Impact of effluent organic matter on low-pressure membrane fouling
661 in tertiary treatment, *Water Res.*, 47, no 8, Art. no 8, (2013), doi:
662 10.1016/j.watres.2013.01.043.

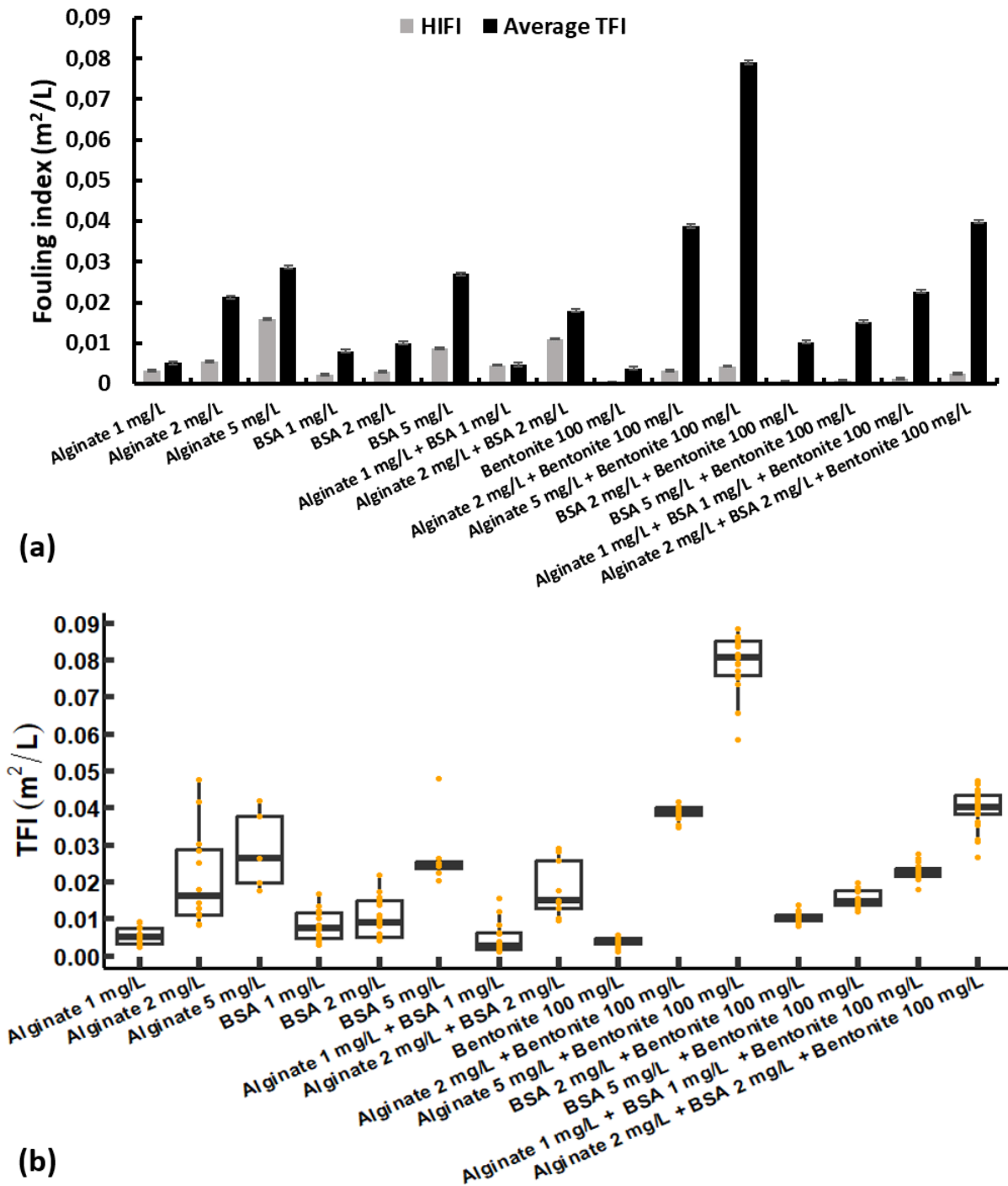
- 663 [35] R. G. Brown, Statistical forecasting for inventory control. McGraw-Hill, (1959).
- 664 [36] C. C. Holt, Forecasting seasonals and trends by exponentially weighted moving
665 averages, *Int. J. Forecast.*, 20, no 1, 5- 10, (2004), doi: 10.1016/j.ijforecast.2003.09.015.
- 666 [37] R. Winters, Forecasting Sales by Exponentially Weighted Moving Averages, *Manag.*
667 *Sci.*, 6, no 3, 324- 342, (1960), doi: 10.1287/mnsc.6.3.324.
- 668 [38] R. J. Hyndman et G. Athanasopoulos, *Forecasting: Principles and Practice* (2nd ed),
669 2nd edition. OTexts: Melbourne, Australia, 2018. Consulté le: 8 mars 2021. [En ligne].
670 Disponible sur: <https://Otexts.com/fpp2/>
- 671 [39] R. Hyndman, A. B. Koehler, J. K. Ord, et R. D. Snyder, *Forecasting with Exponential*
672 *Smoothing: The State Space Approach*. Berlin Heidelberg: Springer-Verlag, 2008. doi:
673 10.1007/978-3-540-71918-2.
- 674 [40] R. Hyndman et RStudio, fpp2: Data for “Forecasting: Principles and Practice” (2nd
675 Edition). 9 septembre 2020. Consulté le: 27 octobre 2021. [En ligne]. Disponible sur:
676 <https://CRAN.R-project.org/package=fpp2>
- 677 [41] H. Chang et al., Role of backwash water composition in alleviating ultrafiltration
678 membrane fouling by sodium alginate and the effectiveness of salt backwashing, *J. Membr.*
679 *Sci.*, 499, 429- 441, févr. 2016, doi: 10.1016/j.memsci.2015.10.062.
- 680 [42] J. Wang, J. Tian, S. Gao, W. Shi, et F. Cui, Dopamine triggered one step
681 polymerization and codeposition of reactive surfactant on PES membrane surface for
682 antifouling modification, *SePurif. Technol.*, 249, 117148, (2020), doi:
683 10.1016/j.seppur.2020.117148.

- 684 [43] D. Liu, T. Wang, et C. He, Antifouling polyethersulfone membrane blended with a
685 dual-mode amphiphilic copolymer, *J. Mater. Sci.*, 51, no 16, Art. no 16, (2016), doi:
686 10.1007/s10853-016-9904-9.
- 687 [44] H. Susanto, H. Arafat, E. M. L. Janssen, et M. Ulbricht, Ultrafiltration of
688 polysaccharide–protein mixtures: Elucidation of fouling mechanisms and fouling control by
689 membrane surface modification, *SePurif. Technol.*, 63, no 3, Art. no 3, (2008), doi:
690 10.1016/j.seppur.2008.06.017.
- 691 [45] W. H. Blade et R. Boulton, Adsorption of Protein by Bentonite in a Model Wine
692 Solution, *Am. J. Enol. Vitic.*, 39, no 3, Art. no 3, (1988).
- 693 [46] F.-X. Sauvage, B. Bach, M. Moutounet, et A. Vernhet, Proteins in white wines:
694 Thermo-sensitivity and differential adsorption by bentonite, *Food Chem.*, 118, no 1, Art. no 1,
695 (2010), doi: 10.1016/j.foodchem.2009.02.080.
- 696 [47] X. Sun, C. Li, Z. Wu, X. Xu, L. Ren, et H. Zhao, Adsorption of Protein from Model
697 Wine Solution by Different Bentonites* *Supported by the National Natural Science
698 Foundation of China (No.20466002), the Program of Ministry of Education for New Century
699 Excellent Talents (NCET-04-0989), the Doctor Funds of Xinjiang Bingtuan (04BSZJ04) and
700 the Shihezi University’s Key Scientific and Technological Project (ZDGG2004-01)., *Chin. J.*
701 *Chem. Eng.*, 15, no 5, Art. no 5, (2007), doi: 10.1016/S1004-9541(07)60137-2.



702

703 *Figure 1: TMP build up during ultrafiltration of synthetic solutions containing (a) alginate, (b) BSA, (c) alginate*
 704 *and BSA, (d) alginate and bentonite, (e) BSA and bentonite and (f) alginate, BSA and bentonite*

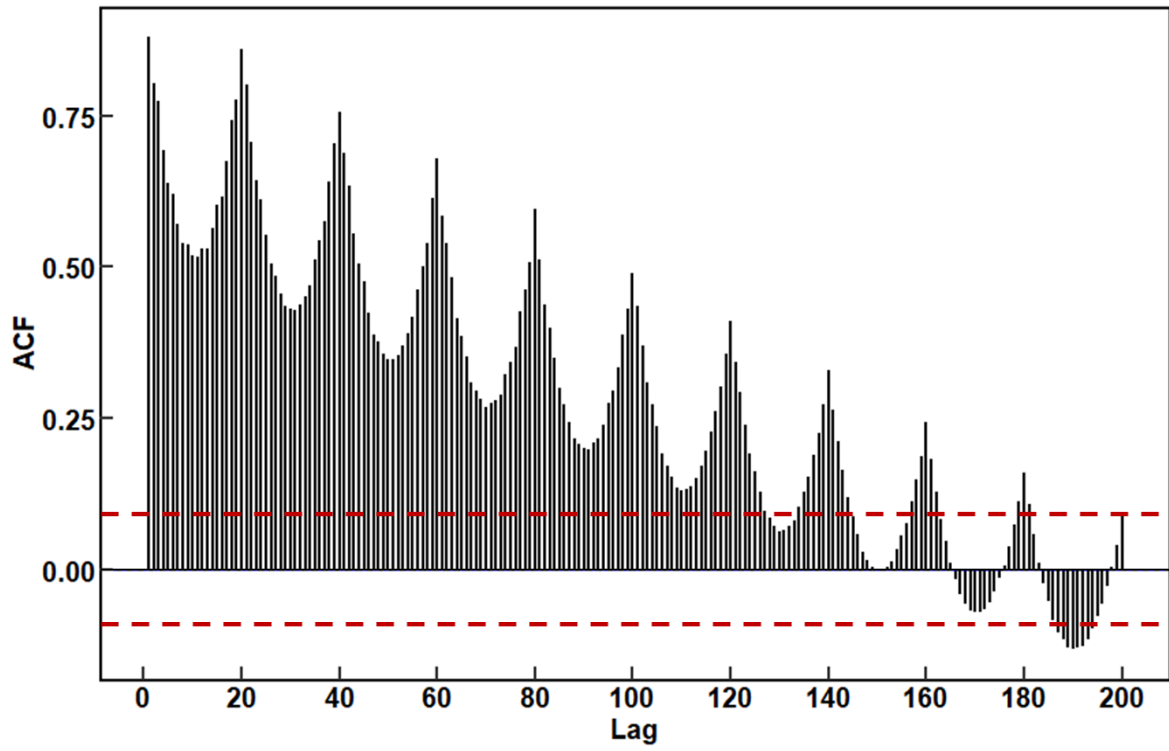


705

706

707

Figure 2: (a) Average TFI and HIFI values for tested conditions, (b) Boxplot showing the TFI distribution for tested conditions with orange dot points representing TFI values for every filtration cycle.

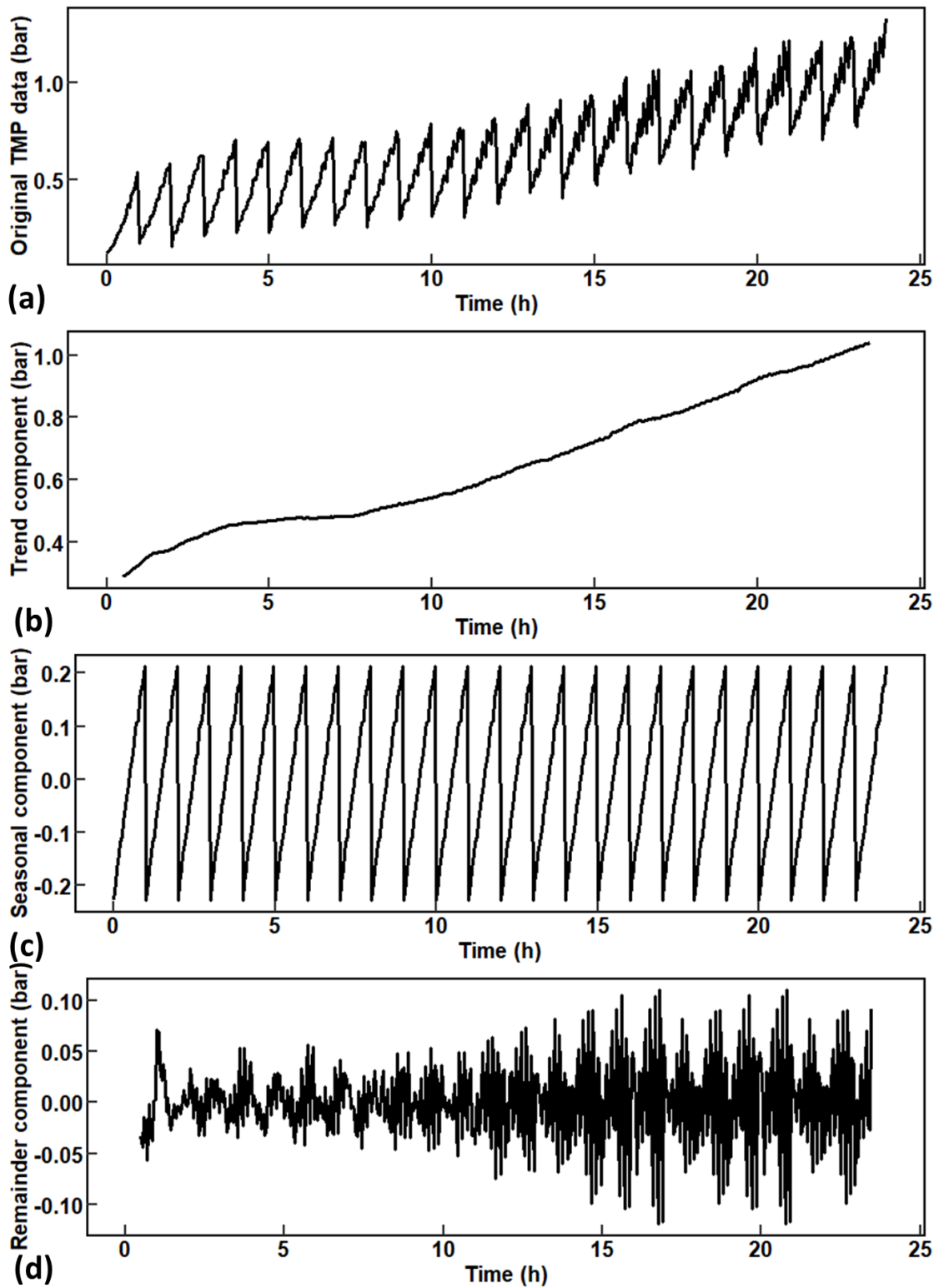


708

709

710

Figure 3: ACF plot for the filtration curve obtained from the filtration of 2 mg/L of Alginate and 100 mg/L of Bentonite

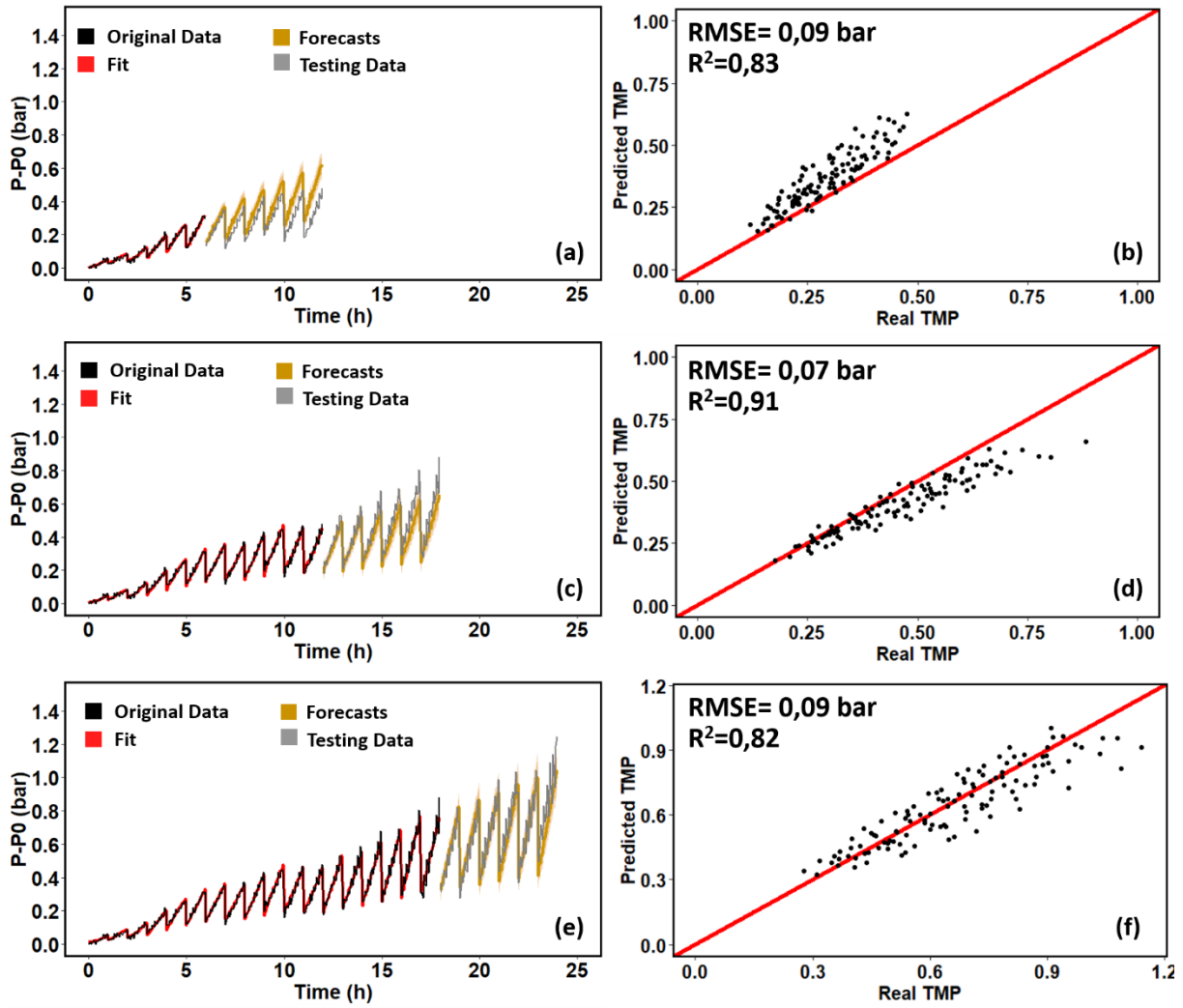


711

712 *Figure 4: A classical multiplicative decomposition of the filtration curve obtained from the filtration of 2 mg/L of*

713 *Alginate and 100 mg/L of Bentonite showing (a) the original curve, (b) the trend, (c) the seasonal component*

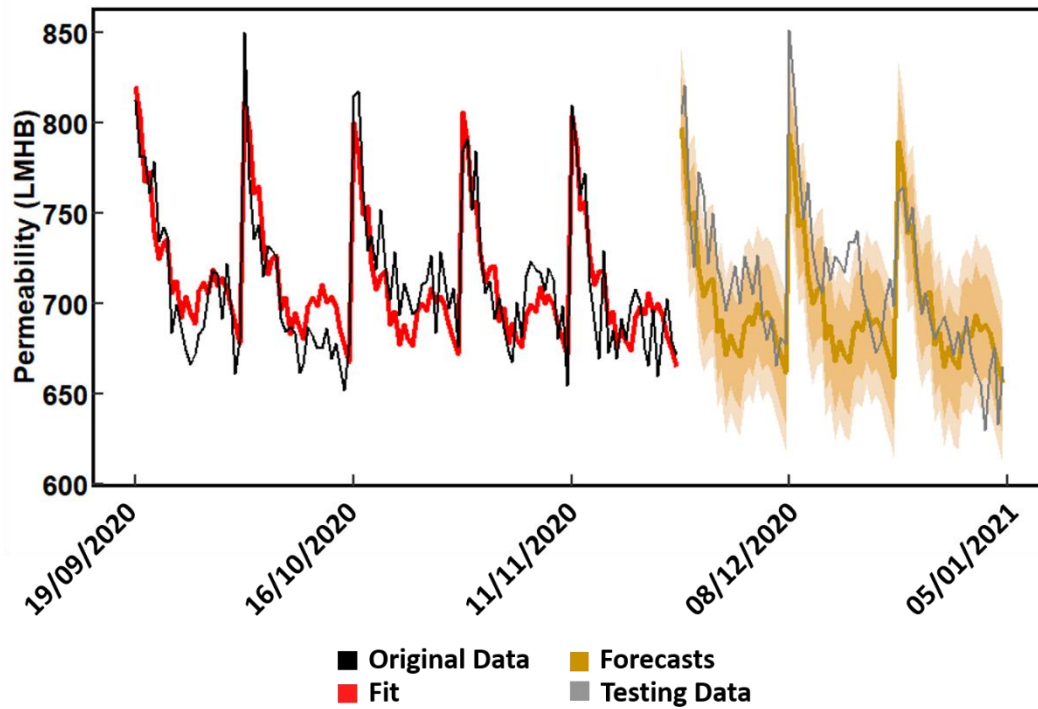
714 *and (d) the remainder component.*



715

716 *Figure 5: (a-c-e) Real-time prediction of the TMP with the actual filtration curve in black (training data), the*
 717 *model fit in red, the forecasts with 95% prediction intervals in yellow and testing data in gray, (b-d-f)*

718 *Comparison between the actual pressure and the predicted pressure*



719

720 *Figure 6: Application of ETS forecasting on ultrafiltration industrial data with the smoothed filtration curve in*
 721 *black (training data), the model fit in red, the forecasts with both 80% and 95% prediction intervals in yellow*
 722 *and the testing data in gray.*