Comparison between artificial neural networks and Hermia´s models to assess ultrafiltration performance

María-José Corbatón-Báguena, María-Cinta Vincent-Vela, José-Marcial Gozálvez-Zafrilla, Silvia Álvarez-Blanco*, Jaime Lora-García, David Catalán-Martínez

Institute for Industrial, Radiophysical and Environmental Safety (ISIRYM),

Universitat Politècnica de Valencia, C/Camino de Vera s/n 46022 Valencia, Spain.

Tel: +34 96 3879630 Ext: 79630; Fax: +34 96 3877639

**Corresponding author: sialvare@iqn.upv.es*

Abstract

In this work, flux decline during crossflow ultrafiltration of macromolecules with ceramic membranes has been modeled using artificial neural networks. The artificial neural network tested was the multilayer perceptron. Operating parameters (transmembrane pressure, crossflow velocity and time) and dynamic fouling were used as inputs to predict the permeate flux. Several pretreatments of the experimental data and the optimal selection of the parameters of the neural networks were studied to improve the fitting accuracy.

The fitting accuracy obtained with artificial neural networks was compared with Hermia pore blocking models adapted to crossflow ultrafiltration. The artificial neural networks generate simulations whose performance was comparable to that of Hermia's models adapted to crossflow ultrafiltration. Considering the computational speed, high accuracy and the ease of the artificial neural networks methodology, they are a competitive, powerful and fast alternative for dynamic crossflow ultrafiltration modeling.

Keywords: Crossflow ultrafiltration, artificial neural networks, fouling, modeling.

1. Introduction

In the last decades, the interest in the use of ultrafiltration (UF) technology has focused on wastewater treatment, recovery of high value compounds from wastewater currents, and the production of drinking water and process water [1]. However, membrane fouling is the main obstacle to a wider application of UF processes as it implies great energy consumption and high operation and maintenance costs [2]. Therefore, a better understanding of membrane fouling is the key to solve the problems arising in the application of this technology [1]. The characterization of membrane fouling makes possible to estimate the capacity and efficiency of the membrane under certain conditions.

Artificial neural networks (ANNs) have been used in the last years in a wide range of scientific and business fields [3-6]. One of the main advantages of ANNs is their capability to learn and recognize trends in a series of input and output data without having into consideration prior assumptions or hypothesis about the relationships governing the process parameters [7]. Compared to the conventional mathematical models used to predict the evolution of permeate flux decline with time during membrane filtration processes, it is noteworthy that these models have certain shortcomings: they involve complex mathematical equations, experimental data is sometimes necessary to infer the input parameters, their empirical equations are only valid in the range of experimental conditions tested and should be fitted for each experimental condition at a time [7]. On the contrary, ANNs are able to accurately predict the complex non-linear relationships between input and output variables of a system and to simulate all the experimental conditions tested at once. For these reasons, some authors concluded that ANNs are a competitive, powerful and fast

55 alternative for dynamic crossflow UF modeling [7-17]. One of the latest applications of
56 ANNs corresponds to the dynamic and steady-state modeling [7] for process control
57 purposes [8], especially in the membrane technology field. Some previous works
58 available in the literature have successfully developed and employed ANNs for different
59 applications from microfiltration and UF to nanofiltration and reverse osmosis and
60 different feed solutions [9-17]. For instance, Chakraborty *et al.* [11] studied the UF of
61 aqueous solutions containing chromium (VI) and correlated the permeate flux and the
62 membrane performance index to different operating conditions (feed flow rate,
63 transmembrane pressure, polymer to metal ratio and pH) using an ANN model. They
64 developed a feed-forward ANN consisting of two hidden layers and based on a
65 Bayesian algorithm. These authors found more accurate predictions by means of the
66 ANN model in comparison with those obtained using a conventional multiple regression
67 analysis. Soleimani *et al.* [12] predicted the permeate flux and fouling resistance after
68 the UF of oily wastewaters by applying ANN models. They created the feed-forward
69 ANN with the Levenberg-Marquadt back-propagation algorithm and they used the
70 transmembrane pressure, the crossflow velocity, the feed temperature and the pH as
71 input variables. They obtained an excellent agreement (values of coefficient of
72 determination greater than 0.99) between the predicted values and the experimental
73 data. Purkait *et al.* [13] investigated the prediction of permeate flux obtained in
74 nanofiltration and reverse osmosis treatments of leather plant effluents. They applied a
75 multi-layered feed-forward ANN with back-propagation algorithm for both batch and
76 crossflow experiments. The optimal ANN consisted of two hidden layers and provided
77 mean absolute error values lower than 1 %. Finally, Rahmanian *et al.* [17] designed an
78 ANN to predict the experimental data obtained from a wastewater micellar-enhanced
79 UF process. These authors tested a three-layer feed-forward ANN using the
80 Levenberg-Marquadt algorithm for training and seven variables as input
81 (transmembrane pressure, pH, electrolyte concentration, feed SDS concentration, etc.).

They observed that there was a good agreement between the ANN model results and the experimental data, being the ANN developed an effective tool to predict complex non-linear relationships.

In this paper, feed-forward ANNs with one intermediate layer and based on a Levenberg-Marquadt training algorithm were created to predict the permeate flux decline with time during the crossflow UF of polyethylene glycol (PEG). In addition, the influence of two pretreatment methods (the normalization of the output variable and the introduction of a fouling indicator as an additional input) of the experimental data on the fitting accuracy of the ANNs models was evaluated. Since only few papers available in the literature deal with the comparison between the goodness of fit provided by the ANN models and the classical ones [18, 19], in this paper ANN predictions were compared with those of Hermia´s classical fouling models, once the optimum ANN parameters were determined and the training of the network with a set of UF experimental data was performed.

2. Theory

2.1. Hermia's models adapted to crossflow ultrafiltration

Hermia's models adapted to crossflow UF are four semi-empirical models based on constant pressure filtration laws [20], whose general equation is as follows (Eq. 1):

$$\frac{d^2t}{dV^2} = K_{DF} \cdot \left(\frac{dt}{dV}\right)^n \quad (1)$$

Where t is the filtration time, V is the permeate volume, K_{DF} is a phenomenological coefficient for dead-end filtration and n is the characteristic model constant.

The classical dead-end filtration models were modified by Field *et al.* [21] to account for the back-transport mass transfer occurring in crossflow filtration by including the permeate flux obtained at the steady-state [22-24]. This modification results in the following general differential equation Eq. (2).

$$-\frac{dJ_P}{dt} = K_{CF} \cdot \left(J_P - J_{PSS}\right) \cdot J_P^{2-n} \qquad (2)$$

Where J_P is the permeate flux at a given time, J_{PSS} is the permeate flux when steady-state was achieved and K_{CF} is a phenomenological coefficient for crossflow filtration. The value of the characteristic model constant (n) depends on the type of fouling mechanism and thus, Hermia distinguished four different types of fouling named as complete blocking (n = 2), intermediate blocking (n = 1), standard blocking (n = 3/2) and gel layer formation (n = 0).

One of the main advantages of the models developed by Hermia is the physical meaning of their phenomenological coefficients, as they allow a deeper comprehension of the fouling mechanisms taking place onto the membrane surface and/or inside its pores. The main hypotheses of each fouling mechanism are well described in the literature [22, 25] and can be resumed as follows: if the solute molecules have a much smaller size than the membrane pores, they can enter in the pores, attach to their walls and diminish the internal diameter of such pores (standard blocking); when solute molecules are approximately of the same size as membrane pores, these molecules are able to seal the pore and accumulate one on each other (intermediate blocking) or

they form a monolayer (complete blocking); if the solute molecules cannot pass through the membrane pores as the former ones are much bigger than the latter, solute molecules can form a cake on the membrane surface (cake/gel layer formation). General equations for each fouling mechanism and their phenomenological coefficients are represented in Eqs. (3) to (9):

- Complete blocking: $$J_P = J_{PSS} + \left(J_0 - J_{PSS}\right) \cdot e^{-K_c \cdot J_0 \cdot t} \quad (3)$$

- Intermediate blocking: $$J_P = \frac{J_0 \cdot J_{PSS} \cdot \left(e^{K_i \cdot J_{PSS} \cdot t}\right)}{J_{PSS} + J_0 \cdot \left(e^{K_i \cdot J_{PSS} \cdot t} - 1\right)} \quad (4)$$

$$K_C = K_i = \frac{3}{4} \cdot \frac{\rho_m \cdot X_m}{\rho_S \cdot a_p \cdot \psi} \quad (5)$$

- Standard blocking: $$J_P = \frac{J_0}{\left(J_0 + J_0^{1/2} \cdot K_S \cdot t\right)^2} \quad (6)$$

$$K_S = 2 \cdot \frac{K_B}{A_0} \cdot A \cdot J_0^{1/2} \quad (7)$$

- Gel layer formation: $$t = \frac{1}{K_{gl} \cdot J_{PSS}^2} \cdot ln\left[\left(\frac{J_P}{J_0} \cdot \frac{J_0 - J_{PSS}}{J_P - J_{PSS}}\right) - J_{PSS} \cdot \left(\frac{1}{J_P} - \frac{1}{J_0}\right)\right] \quad (8)$$

$$K_{gl} = \frac{a \cdot K_G}{J_0 \cdot R_m} \quad (9)$$

Where K_c, K_i, K_s and K_{gl} are the phenomenological coefficients for complete blocking, intermediate blocking, standard blocking and gel layer formation mechanisms, respectively; ρ_m and ρ_s are the feed solution and the solute densities, respectively; X_m, a_p and ψ are characteristics of the solute (mass fraction at the membrane surface, molecule radius and solute form factor, respectively); A_0 is the membrane porous surface; A is the membrane area; K_B and K_G represent the decline in the cross-sectional area of membrane pores and the gel layer mass, respectively, per unit of total

153 permeated volume; R_m is the hydraulic resistance of the original membrane and *a* is
154 the gel layer specific resistance [22].
155
156 *2.2. Artificial neural networks*
157
158 ANNs are computational models able to simulate the processing and learning functions
159 of a human brain [6, 26]. The general ANN architecture is depicted in Fig. 1 and, as it
160 can be observed, an ANN is formed by a group of parallel, processing elements named
161 neurons, units of knots [6, 19]. Neurons in a certain layer of the ANN are connected to
162 those from the previous layer by a number of weighted connections. In addition, there
163 is an extra weight, named bias, which is summed to the rest of input weights [18]. As
164 usual, neurons are distributed in different layers, according to Fig. 2: input, intermediate
165 (or hidden) and output layers [10, 15]. Thus, according to Fig. 1, the output of a neuron
166 in a certain layer acts as input signal for the neurons in the following layer. In order to
167 calculate an output of a neuron, a transfer function is required for its net input to be
168 transformed. As a consequence of all these connections, the learning process can be
169 fitted by selecting the optimal combination of neurons and weights for each studied
170 system [6, 19].
171
172 ## 3. Materials and Methods
173
174 *3.1. Experimental procedure*
175
176 A model solution consisting of polyethylene glycol (PEG) was used as feed during the
177 UF process in a conventional pilot plant. PEG used had an average molecular weight
178 of 35 kDa according to the manufacturer (Merck, Germany) and its concentration in the
179 feed solution was set at 5 g/L. UF experiments were carried out with monotubular

ceramic membranes from Orelis, France (Carbosep M2 of zirconium dioxide with a porous carbon support). Their molecular weight cut-off (MWCO) was 15 kDa and their useful area was 35.5 cm^2. The experimental procedure consisted on a first step in which membrane water permeability was determined, followed by fouling tests using PEG solutions at different experimental conditions, according to Vincent-Vela *et al.* [25]. These experimental tests were developed in total recirculation mode during 7 hours at a temperature of 25 ºC and different values of transmembrane pressure (TMP of 0.1, 0.2, 0.3 and 0.4 MPa) and crossflow velocity (CFV of 1, 2 and 3 m/s).

3.2. ANN modeling

In this work, a MATLAB® software was used to construct and run the feed-forward artificial neural networks (FF ANNs) tested. According to Fig. 2, three operating parameters were considered as input variables: the transmembrane pressure (TMP), the crossflow velocity (CFV) and the operating time.

As it was abovementioned, a transfer function is required to obtain the output values from the neurons. Table 1 shows the types of transfer functions employed in this work: firstly, the hyperbolic tangent sigmoid ('tansig') function was selected to connect the input layer to the intermediate one; and then, the linear transfer function ('purelin') linked the intermediate and the output layers. These transfer functions were selected according to the information provided in [27].

The procedure followed to complete the study of the ANNs construction and performance consisted of several steps:

1. Experimental data was divided in three independent groups for training (50 %), validation (25 %) and test (25%). This division was randomly performed. The total number of experimental data and its division is shown in Table 2. Some other authors [28, 29] have also used this division (50% - 25% - 25%) in order to present more new data to the ANN once trained than that established by default and thus, to improve the generalization process of the developed ANN. These authors achieved high regression coefficient values for both training and test processes with this division.

2. The influence of different pretreatments of the experimental data on the ANN fitting accuracy was studied. For this purpose, as summarized in Table 3, the accuracy of the ANN model predictions without pretreating the experimental data was compared to that achieved after three different situations: when the permeate flux was normalized as in Eq. (10) [9, 13, 14, 16, 30], after adding a new input consisting of a fouling indicator (Eq. (11)), and when both pretreatments (flux normalization and an additional input) were used. It is important to highlight that the use of a fouling indicator allows taking into account the dynamic performance of the UF process as a function of some experimental parameters, such as TMP.

$$J_{normalized} = \left(1 - \Delta^{L} - \Delta^{U}\right) \cdot \frac{J_P - J_{min}}{J_{max} - J_{min}} + \Delta^{L} \quad (10)$$

$$R(t) = \frac{TMP}{\mu \cdot J_P(t)} - R_m \quad (11)$$

Where J_{min} and J_{max} are the minimum and maximum permeate flux measured, respectively (with values of 25 and 175 L/m^2·h, respectively); Δ^L and Δ^U are the lower and upper limits for the extrapolation ability of the ANN (with values of 0.01 for each limit); and μ is the feed solution viscosity.

3. The training step was carried out using the Levenberg-Marquardt algorithm with early stopping. As other authors reported [31, 32], this algorithm has the fastest convergence ability among the available training algorithms. In addition, the mathematical algorithm used during the learning step was the gradient descent with momentum weight and bias learning function.

4. Two different types of weights initialization were tested: null initialization and random initialization (see Table 3).

5. A FF ANN was trained taking into consideration all these different alternatives for data pretreatment and weights initialization and their simulation results were compared in terms of fitting accuracy to the experimental permeate flux measured.

6. An analysis of the variance (ANOVA) was performed as a final step to check if the main effects studied (pretreatment, weights initialization and number of neurons in the intermediate layer) were statistically significant for the ANN model fitting accuracy. This accuracy was expressed in terms of the regression coefficient, R^2, and the normalized mean square error, NMSE, according to Eqs. (12) and (13). In addition, NMSE values during the training, validation and test processes was plotted against the number of iterations in order to check if any overfitting effect occurs.

$$R^2 = \left(\frac{cov\left(y_{calc}, y_{exp}\right)}{\sigma\left(y_{calc}\right)\sigma\left(y_{exp}\right)} \right)^2 \quad (12)$$

$$NMSE = \frac{\sum_{i=1}^{N}\left(y_{exp} - y_{calc}\right)^2}{\sigma\left(y_{exp}\right)^2 \cdot N} = \frac{\sum_{i=1}^{N}\left(y_{exp} - y_{calc}\right)^2}{\sum_{i=1}^{N}\left(y_{exp} - \overline{y_{exp}}\right)^2} / \; 0 < NMSE < 1 \quad (13)$$

Where y_{calc} and y_{exp} are the predicted and the experimental values, respectively; σ is the standard deviation; $\bar{y}$ is the mean value of y; N is the number of processed data; and the covariance cov is defined as in Eq. (14):

$$cov\left(y_{calc}, y_{exp}\right) = \frac{\sum_{i=1}^{N}\left(y_{calc,i} - \overline{y_{calc}}\right)\left(y_{exp,i} - \overline{y_{exp}}\right)}{N-1} \quad (14)$$

4. Results and discussion

The data set was used to train the ANN and the fitting accuracy of the ANN model was compared to that obtained with Hermia's models adapted to crossflow UF described in a previous work [22].

4.1. Network architecture

For the identification of the best modeling methodology with the ANNs, a statistical analysis of variance (ANOVA) on the fitting accuracy results was performed. It is important to highlight that, as it is well known, the regression coefficient R^2 could surpass its maximum value ($R^2 > 1$) or has a negative value in some cases. Therefore, a normal distribution of R^2 was used to avoid possible inconsistencies and thus, the response variable used in the ANOVA was $[-\log_{10}(1-R^2)]$. Regarding the ANOVA test, Table 4 summarises the results when a 95 % confidence level was used in the analysis. Statistics evaluated in the ANOVA test were sum of squares, degrees of freedom (Df), mean square, F-ratio and p-value for the main effects (pretreatment, A, weights initialization, B, and neurons in the intermediate layer, C), and their double and triple interactions. F-ratio is an indicator of the variance of the data about the mean

282 value. When the F-value departs from the unity, the design variables are adequate in
283 providing a suitable explanation for the variation in the mean of the data [33]. Based on
284 this statistical, the p-value is calculated with the F-value and the degrees of freedom
285 [11]. Using a confidence interval of 95 %, p-values lower than 0.05 indicate statistically
286 significance of the design variables on the response one. According to the results
287 shown in Table 4, it is remarkable that only the single factors A, B and C have
288 statistically significant effects based on their p-values (0.0024, 0.0082 and 0.0026,
289 respectively) and F-ratios (5.15, 7.29 and 3.96, respectively) on the response variable
290 $[-\log_{10}(1-R^2)]$.
291
292 The influence of the abovementioned factors on the response variable can be
293 determined by using the Least Significant Difference (LSD) intervals analysis. This
294 statistical analysis allows the calculation of the smallest significant difference between
295 two means. This means that, if the absolute value of the difference between two means
296 is greater than the LSD interval (i.e. the LSD intervals do not overlap), the comparison
297 is significant at the selected confidence level [34]. LSD intervals for the main factors A
298 (pretreatments), B (weights initialization) and C (neurons in the intermediate layer) are
299 depicted in Figs. 3-5 respectively. Fig. 3 shows the LSD intervals for the response
300 variable $[-\log_{10}(1-R^2)]$ for the different pretreatments considered. It can be observed
301 that the use of pretreatments improves the accuracy obtained. The interval of the
302 pretreatment 2 (Pret 2) and 3 (Pret 3) does not overlap with the interval of the
303 pretreatment 0 (Pret 0). This means that pretreatments 2 and 3 significantly improve
304 ANN accuracy, while pretreatment 1 does not, as its LSD interval overlaps with
305 pretreatment 0. The pretreatment that offered the best accuracy was the double
306 pretreatment (Pret 3). However, comparing both intervals for pretreatments 2 and 3, it
307 can be concluded that the difference between these two different pretreatments was
308 not statistically significant. Regarding the effect of the weight initializations (null and

309 random) on the response variable [$-\log_{10}(1-R^2)$], the corresponding LSD intervals are
310 shown in Fig. 4. In this case, the intervals for both initializations are clearly separated
311 one from each other and thus, the random initialization significantly achieves a more
312 accurate prediction. In the same way, Fig. 5 shows the LSD intervals for the response
313 variable [$-\log_{10}(1-R^2)$] for different number of neurons in the intermediate layer. The
314 best accuracies were obtained for the highest number of neurons of the intermediate
315 layer for the range tested. For 8 neurons and above, there is no significant difference in
316 the accuracy because the intervals overlap. This can be due to overfitting when
317 introducing excessive nodes in the intermediate layer. These results are similar to
318 those obtained by other authors [9, 14, 15] in the application of ANNs to dynamic
319 permeate flux in MF and UF. Other studies on NF showed that the best fitting was
320 obtained for 6 to 8 neurons in the intermediate layer [32].
321
322 Taking into account all the information provided from Figs. 3-5, it can be concluded that
323 the best methodology for the developed ANN model consists of double pretreatment
324 (normalization of the permeate flux values and the use of an additional input, which
325 was a fouling indicator), random initialization of the weights and 8 neurons in the
326 intermediate layer.
327
328 Based on these optimal results, a simulation of the ANN performance was carried out
329 and shown in Figs. 6 and 7. On the one hand, Fig. 6 represents the fitting accuracy
330 obtained with the ANN model for the complete experimental data ('Results') and the
331 different datasets ('Training', 'Validation' and 'Test'). In this figure, the experimental
332 permeate flux data (as target values) was compared to the predicted permeate flux
333 values (or output values). The linear regression determined for each dataset is shown
334 in its corresponding graph and, according to the value of the regression coefficients R^2,
335 highly accurate fitting results were obtained using these equations. Moreover, all

simulation results were below 5 % of deviation as it could be observed in Fig. 6. On the other hand, the evolution of the NMSE with the number of iterations during the training, validation and test processes was used to evaluate if any overfitting effect occurs. When overfitting takes place, the validation error decreases up to a minimum value and then it starts to increase. After this iteration in which the validation error increases, overfitting occurs if the training process does not stop. The main effect of overfitting is that the developed ANN is unable to generalise from the trained values to new ones [27]. As it can be observed in Fig. 7, the pattern for both the validation and test errors were almost the same, without no overfitting detected by the iteration 78 where the training stopped. In addition, as it was abovementioned for Fig. 6, the good agreement between the experimental and the predicted results for the 'Test' dataset leads to the conclusion that no significant overfitting occurs during the ANN performance [35] and thus, training algorithm used with the optimal ANN (Levenberg-Marquadt with early stopping) was appropriate to avoid overfitting when the 'Test' data was provided to the ANN.

A confirmation of the high accuracy obtained with the ANN model was corroborated by comparing the ANN predictions to the experimental data at different transmembrane pressures and crossflow velocities. Figs. 8 a, b and c show the results of the experimental permeate flux (represented in dots and previously reported in [22]) and the predictions of the neural network model (represented as solid lines) for crossflow velocity values of 1, 2 and 3 m/s, respectively. Regarding the experimental variation of permeate flux with time, it can be observed that an increase in transmembrane pressure (Fig. 8a) resulted in a sharp decline of permeate flux during the first minutes of operation. In the same way, when comparing Figs. 8a and c for the same transmembrane pressure (for instance, the highest one, 0.4 MPa) and different crossflow velocities (1 and 3 m/s), it is remarkable that that the sharp decline of

363 permeate flux that took place at the lowest crossflow velocity was significantly reduced
364 at 3 m/s. This is due to the fact that less pore blockage phenomena occurred when
365 high crossflow velocity was applied. In addition, the steady-state permeate flux
366 obtained is greater at 3 m/s than that achieved at 1 m/s. This pattern can be explained
367 by the greater the shear stress that high crossflow velocity causes on the proximity of
368 the membrane surface and thus, the solute molecule deposited as a cake layer on the
369 membrane surface diminishes [25, 36]. In addition, concentration polarization has been
370 reported to be a significant foulant phenomenon to take into account [37-40]. At this
371 regard, and according to the mathematical description provided by Jonsson [37], the
372 concentration polarization layer thickness can be calculated from the general film
373 model equation considering the relationship between the permeate flux, the osmotic
374 pressure and the solute concentration at the membrane surface. By this mathematical
375 development, the concentration polarization layer thickness was determined for the
376 different transmembrane pressures and crossflow velocities tested in this work.
377 Regarding the steady-state values obtained at the lowest crossflow velocity used (1
378 m/s), this layer increases from $2.678 \cdot 10^{-4}$ m at 0.1 MPa to $4.732 \cdot 10^{-4}$ m at 0.4 MPa. As
379 it is well-known, concentration polarization increases when the transmembrane
380 pressure increases and thus, the boundary layer near the membrane surface where the
381 concentration polarization phenomenon takes place is thicker [39]. This may be
382 explained by the fact that at a high transmembrane pressure, solute molecules are
383 forced towards the membrane surface and thus, they can accumulate on its
384 proximities. On the contrary, regarding the values of the concentration polarization
385 layer thickness obtained at the highest transmembrane pressure used (0.4 MPa), the
386 effect of crossflow velocity was less significant, achieving values of δ ranging from
387 $4.732 \cdot 10^{-4}$ m at 1 m/s to $4.498 \cdot 10^{-4}$ m at 3 m/s. This demonstrated that the higher the
388 crossflow velocity is applied, the lower the concentration polarization phenomenon is
389 observed. This is due to the high shear stress generated when using high crossflow

velocities, which prevents solute molecules from accumulating on the membrane surface [40].

As it can be also observed in Figs. 8 a, b and c, the ANN model predictions fitted with high accuracy the permeate flux decline along the ultrafiltration process, especially at the steady-state values, for all the transmembrane pressures and crossflow velocities tested. This fact confirms that the optimal methodology selected to create and train the ANN proposed in this work results in an adequate model to predict the permeate flux decline with time. The high fitting accuracy obtained is comparable to that of the ANNs predictions found in the literature for different feed solutions and transmembrane pressures [10-16]. The experimental conditions, type of membrane process, configuration of the ANN and main results of these previous studies are summarized in Table 5. According to the provided information, some authors used ANN models with two or more intermediate layers [10, 13, 16], while other authors have chosen the data entered in the training step manually [10, 14-16]. Regarding the former ones, an increase in the number of intermediate layers results in an increase in the complexity of the developed model. In addition, the training time, the risk of overfitting and the network error may decrease by reducing the number of intermediate layers [10, 14]. As the number of these layers depends on the complexity of the input data, in this work one intermediate layer was selected as the optimal ANN methodology, due to the high accuracy obtained when predicting the permeate flux decline with time, its high computational speed and low complexity. On the other hand, regarding the training step, a random selection of the data used in this step is the most often used [9, 13, 17, 30, 41] to guarantee that the statistical differences obtained with the ANN model in the output parameters are not due to a manual selection of the data. Therefore, in this work a random distribution of the data in the training process was performed, achieving high fitting accuracies with this ANN methodology.

417

418 *4.2. Comparison between the ANN selected and Hermia pore blocking models adapted*
419 *to crossflow ultrafiltration*

420

421 Hermia's models were used to fit experimental data in a previous study [22]. As
422 experimental conditions can highly influence the prediction accuracy, the effect of such
423 experimental conditions (TMP and CFV) was evaluated for both ANNs and Hermia's
424 models (Table 7). Firstly, the square regression coefficient values achieved for each
425 combination of transmembrane pressure and crossflow velocity in the case of ANN are
426 shown in Table 6.

427

428 Regarding the ANOVA test shown in Table 7, three different main effects and their
429 double interactions on the response variable [$-\log_{10}(1-R^2)$] were considered: the
430 transmembrane pressure (A), the crossflow velocity (B) and the type of model used
431 (C). As the type of model used is a character variable, the following codification was
432 employed to convert this variable into a numeric one: 0 for complete blocking, 1 for
433 intermediate blocking, 2 for gel layer and 3 for ANN model. Statistics evaluated in this
434 ANOVA test were, as in the ANOVA test shown in Table 4, sum of squares, degrees of
435 freedom (Df), mean square, F-ratio and p-value. Based on the latter statistical and
436 using a confidence interval of 95 %, p-values indicated that factors A, B, C and the
437 interactions AB and AC have statistically significant effects on the fitting accuracy (p-
438 values of 0.0000, 0.0001, 0.0015, 0.0000 and 0.0406, respectively). Taking into
439 account these results of significance, a comparison of the means obtained for the main
440 factors A, B and C was displayed in a LSD intervals test, considering [$-\log_{10}(1-R^2)$] as a
441 response variable.

442

443 Figs. 9-11 show the LSD intervals for the fitting accuracy achieved for the TMPs, CFVs
444 and models tested, respectively. Fig. 9 shows the effect of TMP on the fitting accuracy
445 of the models employed for the different CFV tested. This means that, for each value of
446 TMP (0.1 to 0.4 MPa), the results obtained for 1, 2 and 3 m/s were averaged. It can be
447 observed that the lowest level of TMP (0.1 MPa) corresponds to the worst fitting
448 accuracy regardless of the model used because for this TMP fouling was less severe.
449 As TMP increases, permeate flux decline and the fitting accuracy also significantly
450 increase for the selected confidence level. This pattern is confirmed by the results
451 summarized in Table 6, since the square regression coefficient increased as
452 transmembrane pressure increased for the complete blocking, intermediate blocking,
453 gel layer and ANN model. On the other hand, Fig. 10 shows the effect of CFV on the
454 fitting accuracy of the models employed for the different TMP tested. For each value of
455 CFV (1 to 3 m/s), the results obtained for 0.1 to 0.4 MPa were averaged. It can be
456 observed that increasing the CFV results in a decrease of the fitting accuracy, as
457 fouling is less severe for high CFVs. In this case, the improvement in the fitting
458 accuracy obtained at the lowest CFV (1 m/s) was statistically significant in comparison
459 with that determined at CFV values of 2 and 3 m/s, as their LSD intervals did not
460 overlap. Finally, Fig. 11 shows the accuracy of each model for the different TMP and
461 CFV tested. In this case, the results obtained for each model at all the possible
462 combinations of TMP (0.1 to 0.4 MPa) and CFV (1 to 3 m/s) were averaged. ANN and
463 complete and intermediate blocking models are significantly more accurate than the gel
464 layer model. It can also be observed that, although the ANN has a slightly lower
465 accuracy than the intermediate and complete blocking models, this difference is not
466 significant because the LSD intervals of these models overlap.
467
468 In order to conclude that the ANN models predicted the experimental results with
469 significant higher accuracy than the other models, the interaction between the

transmembrane pressure (factor A) and the type of model (factor C) tested was depicted in Fig. 12 in terms of the response variable [$-\log_{10}(1-R^2)$]. As it was above mentioned regarding the effect of TMP on the fitting accuracy, all the models selected provided more accurate predictions of the experimental results as transmembrane pressure increased from 0.1 to 0.4 MPa. However, the best fitting accuracy at 0.4 MPa was obtained with the ANN model. This indicates that, for the experimental conditions at which the experimental permeate flux showed the most severe decline with the operation time, the model developed by means of the ANN methodology was the most accurate. This better accuracy was compared to that reported in previous studies available in the literature about fitting of semi-empirical classical models and ANN ones. According to Table 8, it can be observed that ANNs have a higher fitting accuracy than classical models. Although in this study both methods, Hermia's models and ANNs, achieved R^2 higher than 0.99, it can be concluded that ANNs are a suitable methodology to predict the permeate flux decline with time that occurs in membrane separation processes.

5. Conclusions

The dynamic performance of the UF process studied was modeled using ANNs.

1. The pretreatment of the data with the two methods proposed improved the fitting accuracy of ANNs. The initialization of the weights with random values gave better results than the null initialization. The optimum number of neurons in the intermediate layer was 8.
2. The ANNs achieved results very accurate with good fitting to experimental data.
3. The fitting accuracy of FF ANNs is comparable to that of Hermia's models adapted to the crossflow UF. The results obtained with ANNs are similar to those obtained for Hermia's intermediate blocking model for high TMPs.

Considering that Hermia's models require to be fitted for each experimental test condition and that ANNs are able to simulate all the experimental conditions tested at once, it can be concluded that ANNs are a competitive, powerful and fast alternative for dynamic crossflow UF modeling.

6. Acknowledgements

The Spanish Ministry for Science and Innovation (Project OPTIMEM CTM2010-20248) is kindly acknowledged.

Nomenclature

A	Membrane area (m^2)
A_0	Membrane porous surface (m^2)
a	Specific resistance of the gel layer (m/kg)
a_p	Radius of the solute molecule (m)
CFV	Crossflow velocity (m/s)
E_m	Average deviation (dimensionless)
E_{max}	Maximum deviation (dimensionless)
E_{min}	Minimum deviation (dimensionless)
J_0	Initial permeate flux ($L/m^2 \cdot h$)
J_p	Permeate flux ($L/m^2 \cdot h$)
J_{pss}	Steady-state permeate flux ($L/m^2 \cdot h$)
K_c	Constant for complete blocking model for crossflow filtration (m^{-1})
K_{CF}	Phenomenological coefficient—constant
K_{gl}	Constant for gel layer formation model for crossflow filtration (s/m^2)

K_S	Constant for standard blocking model ($m^{-1/2} \cdot s^{-1/2}$)
K_i	Constant for intermediate blocking model for crossflow filtration (m^{-1})
n	Constant for fouling mechanism (dimensionless)
Neur	Number of neurons in the intermediate layer of the ANNs
Norm	Normalization of the permeate flux
Weights	Initialization of the weights in the ANNs
Pret	Data pretreatment
R^2	Square regression coefficient (dimensionless)
$R(t)$	Fouling indicator (m^{-1})
R_m	Membrane resistance (m^{-1})
RE	Relative error (dimensionless)
T	Time (s)
TMP	Transmembrane Pressure (MPa)

Greek letters

μ	Viscosity (kg/m·s))
ρ	Density (kg/m^3)
χ_m	Solute concentration over the membrane surface (dimensionless)
ψ	Solute form factor (dimensionless)
Δ^L and Δ^U	Margins used to give the network limited extrapolation capability in the Eq. (10) (dimensionless).

Abbreviations

ANN	Artificial Neural Network
FF ANN	Feed Forward Artificial Neural Network
LSD	Least Significant Difference
MF	Microfiltration

MP ANN	Multilayer Perceptron Artificial Neural Network
MSE	Mean Square Error
MWCO	Molecular Weight Cut-Off (g/mol)
NF	Nanofiltration
NMSE	Normalized Mean Square Error
PEG	Polyethylene glycol
RMSE	Root Mean Square Error
UF	Ultrafiltration

References

[1] W. Guo, H.-H. Ngo, J. Li, A mini-review on membrane fouling, Biores. Technol. 122 (2012) 27-34.

[2] W. Gao, H. Liang, J. Ma, M. Han, Z.-L. Chen, Z.-S. Han, G.-B. Li, Membrane fouling control in ultrafiltration technology for drinking water production: A review, Desalination 272 (2011) 1–8.

[3] I.A. Basheer, M. Hajmeer, Artificial neural networks: fundamentals, computing, design, and application, J. Microbiol. Method. 43 (2000) 3–31.

[4] M. Paliwal, U.A. Kumar, Neural networks and statistical techniques: A review of applications, Expert Syst. Appl. 36 (2009) 2–17.

[5] J. Misra, I. Saha, Artificial neural networks in hardware: A survey of two decades of progress, Neurocomputing 74 (2010) 239–255.

[6] S. Walczak, N. Cerpa, Artificial Neural Networks, in: R.A. Meyers (Ed.), Encyclopedia of Physical Science and Technology (Third Edition), Academic Press, San Diego (CA), 2003, pp. 631-645.

576 [7] H. Nourbakhsh, Z. Emam-Djomeh, M. Omid, H. Mirsaeedghazi, S. Moini, Prediction
577 of red plum juice permeate flux during membrane processing with ANN optimized using
578 RSM, Comput. Electron. Agric. 102 (2014) 1-9.
579 [8] S. Curcio, V. Calabrò, G. Iorio, Reduction and control of flux decline in cross-flow
580 membrane processes modeled by artificial neural networks and hybrid systems,
581 Desalination 236 (2009) 234-243.
582 [9] S. Chellam, Artificial neural network model for transient crossflow microfiltration of
583 polydispersed suspensions, J. Membr. Sci. 258 (2005) 35-42.
584 [10] H. Chen, A.S. Kim, Prediction of permeate flux decline in crossflow membrane
585 filtration of et al colloidal suspension: a radial basis function neural network approach,
586 Desalination 192 (2006) 415-428.
587 [11] S. Chakraborty, J. Dasgupta, U. Farooq, J. Sikder, E. Drioli, S. Curcio,
588 Experimental analysis, modeling and optimization of chromium (VI) removal from
589 aqueous solutions by polymer-enhanced ultrafiltration, J. Membr. Sci. 456 (2014) 139-
590 154.
591 [12] R. Soleimani, N.A. Shoushtari, B. Mirza, A. Salahi, Experimental investigation,
592 modeling and optimization of membrane separation using artificial neural network and
593 multi-objective optimization using genetic algorithm, Chem. Eng. Res. Des. 91 (2013)
594 883-903.
595 [13] M.K. Purkait, V. Dinesh Kumar, D. Maity, Treatment of leather plant effluent using
596 NF followed by RO and permeate flux prediction using artificial neural network, Chem.
597 Eng. J. 151 (2009) 275-285.
598 [14] M.A. Razavi, A. Mortazavi, M. Mousavi, Dynamic modelling of milk ultrafiltration by
599 artificial neural network, J. Membr. Sci. 220 (2003) 47-58.
600 [15] M.A. Razavi, A. Mortazavi, M. Mousavi, Application of neural networks for
601 crossflow milk ultrafiltration simulation, Int. Dairy J. 14 (2004) 69–80.

602 [16] Sahoo Goloka Behari, Ray Chittaranjan, Predicting flux decline in crossflow
603 membranes using artificial neural networks and genetic algorithms, Journal of
604 Membrane Science 283 (2006) 147-157.
605 [17] B. Rahmanian, M. Pakizeh, S.A.A. Mansoori, R. Abedini, Application of
606 experimental design approach and artificial neural network (ANN) for the determination
607 of potential micellar-enhanced ultrafiltration process, J. Hazard. Mat. 187 (2011) 67-74.
608 [18] H. Al-Zoubi, N. Hilal, N.A. Darwish, A.W. Mohammad, Rejection and modeling of
609 sulphate and potassium salts by nanofiltration membranes: neural network and
610 Spiegler-Kedem model, Desalination 206 (2007) 42-60.
611 [19] B.K. Nandi, A. Moparthi, R. Uppaluri, M.K. Purkait, Treatment of oily wastewater
612 using low cost ceramic membrane: Comparative assessment of pore blocking and
613 artificial neural network models, Chem. Eng. Res. Design 88 (2010) 881-892.
614 [20] J. Hermia, Constant pressure blocking filtration laws – application to power-law
615 non-newtonian fluids, Trans. Inst. Chem. Ing. 60 (1982) 183-187.
616 [21] R.W. Field, D.Wu, J.A.Howell, B.B. Gupta, Critical flux concept for microfiltration
617 fouling, J. Membr. Sci. 100 (1995) 259–272.
618 [22] M.C. Vincent-Vela, S. Álvarez-Blanco, J. Lora-García, E. Bergantiños-Rodríguez,
619 Analysis of membrane pore blocking models adapted to crossflow ultrafiltration in the
620 ultrafiltration of PEG, Chem. Eng. J. 149 (2009) 232–241.
621 [23] C. Jarusutthirak, S. Mattaraj, R. Jiraratananon, Influence of inorganic scalants and
622 natural organic matter on nanofiltration membrane fouling, J. Membr. Sci. 287 (2007)
623 138–145.
624 [24] S.T.D. de Barros, C.M.G. Andrade, E.S. Mendes, L. Peres, Study of fouling
625 mechanism in pineapple juice clarification by ultrafiltration, J. Membr. Sci. 215 (2003)
626 213–224.
627 [25] M. C. Vincent-Vela, Beatriz Cuartas-Uribe, Silvia Álvarez-Blanco, Jaime Lora-
628 García, Enrique Bergantiños-Rodríguez, Analysis of ultrafiltration processes with

629 dilatant macromolecular solutions by means of dimensionless numbers and
630 hydrodynamic parameters, Sep. Purif. Technol. 75 (2010) 332–339.
631 [26] A. Ferrari, in: F. Angeli (Ed.), Aspetti applicativi delle reti neurali artificiali, Milano,
632 1996.
633 [27] M.H. Beale, M.T. Hagan, H.B. Demuth, Neural Network Toolbox. User Guide, The
634 MathWorks Inc. (2011).
635 [28] M. Kabsch-Korbutowicz, M. Kutylowska, Use of artificial intelligence in predicting
636 the turbidity retention coefficient during ultrafiltration of water, Env. Protect. Eng. 37
637 (2011) 75-84.
638 [29] S. Chatterjee, A. Bhattacherjee, Genetic algorithms for feature selection of image
639 analysis based quality monitoring model: An application to an iron mine, Eng. Appl.
640 Artif. Intell. 24 (2011) 786-795.
641 [30] B. Sarkar, A. Sengupta, S. De, S. DasGupta, Prediction of permeate flux during
642 electric field enhanced cross-flow ultrafiltration—A neural network approach, Sep. Purif.
643 Technol. 65 (2009) 260-268.
644 [31] MathWorks Inc., MATLAB, 3 Apple Hill Drive, Natick, MA, USA, 2002.
645 [32] G.R. Shetty, S. Chellam, Predicting membrane fouling during municipal drinking
646 water nanofiltration using artificial neural networks, J. Membr. Sci. 217 (2003) 69–86.
647 [33] R.H. Myers, D.C. Montgomery, C.M. Anderson-Cook, Response Surface
648 Methodology: Process and Product Optimization Using Designed Experiments (Third
649 Edition), John Wiley & Sons, Inc., Hoboken, New Jersey, 2009.
650 [34] L.J. Williams, H. Abdi, Fisher's Least Significant Difference (LSD) Test in N.
651 Salkind (Ed.), Encyclopedia of Research Design, Thousand Oaks, CA, Sage, 2010.
652 [35] Q.-F. Liu, S.-H. Kim, S. Lee, Prediction of microfiltration membrane fouling using
653 artificial neural network models, Sep. Purif. Technol. 70 (2009) 96-102.

[36] E. Alventosa-deLara, S. Barredo-Damas, M.I. Alcaina-Miranda, M.I. Iborra-Clar, Ultrafiltration technology with a ceramic membrane for reactive dye removal: optimization of membrane performance, J. Hazard. Mater. 209–210 (2012) 492–500.

[37] G. Jonsson, Transport phenomena in ultrafiltration: membrane selectivity and boundary layer phenomena, Pure & Appl. Chem. 58 (1986) 1647-1656.

[38] A.L. Zydney, Stagnant film model for concentration polarization in membrane systems, J. Membr. Sci. 130 (1997) 275-281.

[39] P. Bacchin, D. Si-Hassen, V. Starov, M.J. Clifton, P. Aimar, A unifying model for concentration polarization, gel-layer formation and particle deposition incross-flow membrane filtration of colloidal suspensions, Chem. Eng. Sci. 57 (2002) 77-91.

[40] P. Srijaroonrat, E. Julien, Y. Aurelle, Unstable secondary oil/water emulsion treatment using ultrafiltration: fouling control by backflushing, J. Membr. Sci. 159 (1999) 11-20.

[41] M. Jamshidi, M. Ghaedi, K. Dashtian, A.M. Ghaedi, S. Hajati, A. Goudarzi, E. Alipanahpour, Highly efficient simultaneous ultrasonic assisted adsorption of brilliant green and eosin B onto ZnS nanoparticles loaded activated carbon: Artificial neural network modeling and central composite design optimization, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 153 (2016) 257-167.

[42] S. Ghandehari, Mohammad Mehdi Montazer-Rahmati, Morteza Asghari, A comparison between semi-theoretical and empirical modeling of cross-flow microfiltration using ANN, Desalination 277 (2011) 348-355.

[43] C. Aydiner, I. Demir, E. Yildiz, Modeling of flux decline in crossflow microfiltration using neural networks: the case of phosphate removal, J. Membr. Sci. 248 (2005) 53-62.

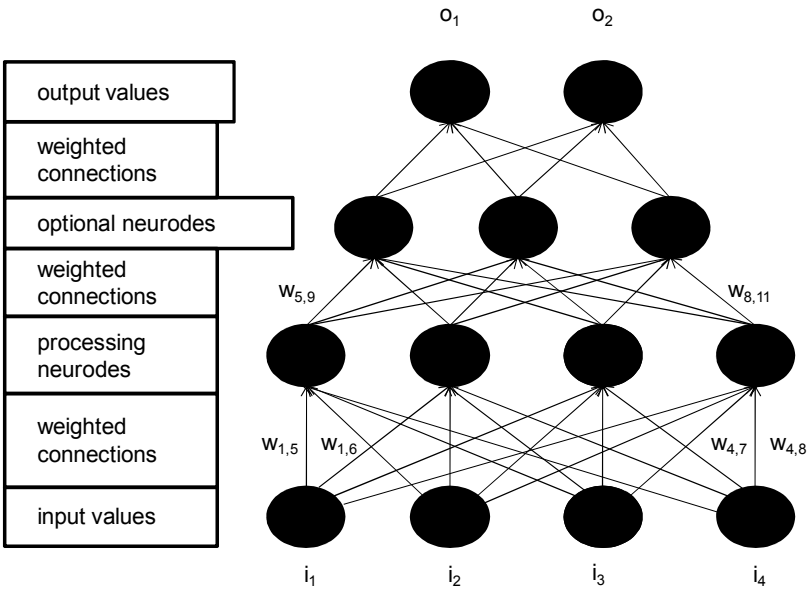


Fig. 1. Sample of an artificial neural network architecture (not all weights are shown) [6].

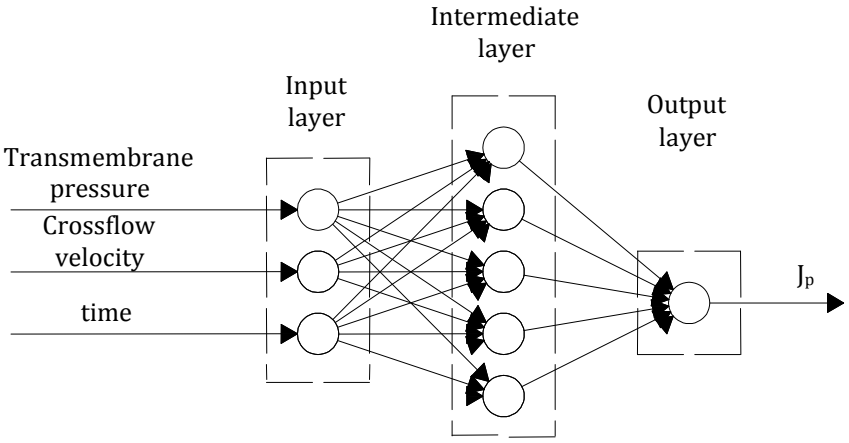


Fig. 2. Feed-forward artificial neural network with five neurons in the intermediate layer.

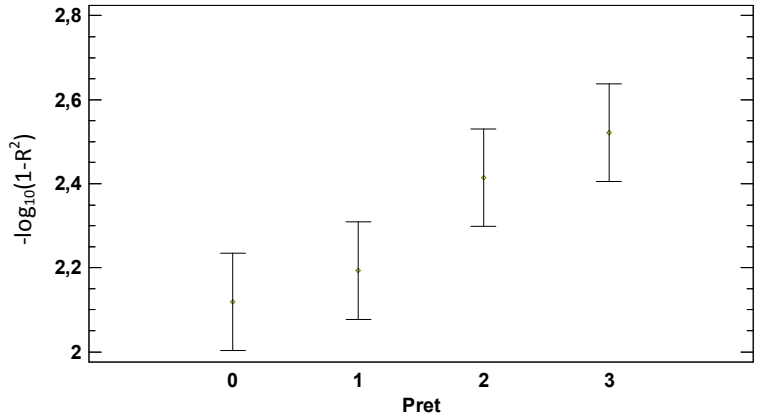


Fig. 3. Means and LSD intervals for $-\log_{10}(1-R^2)$ as a function of the type of pretreatment (0: null pretreatment; 1: normalization of the permeate flux; 2: fouling indicator as an additional input; 3: double pretreatment: normalization of the permeate flux and use of the fouling indicator as an additional input).

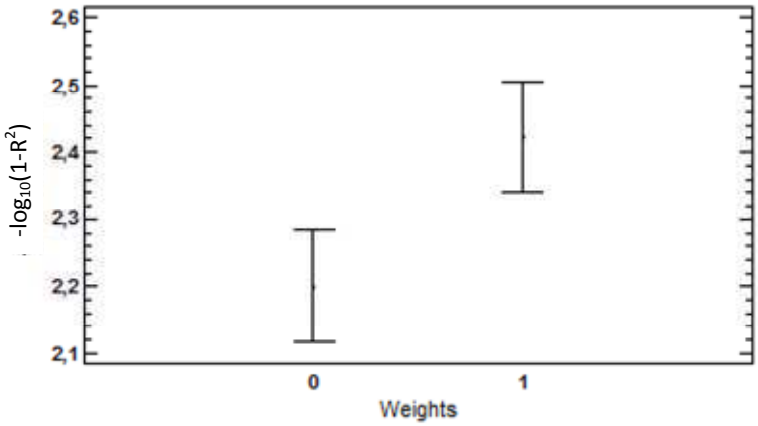


Fig. 4. Means and LSD intervals for $-\log_{10}(1-R^2)$ as a function of the type of initialization of the weights of the artificial neural networks (0: null initialization; 1: random initialization).

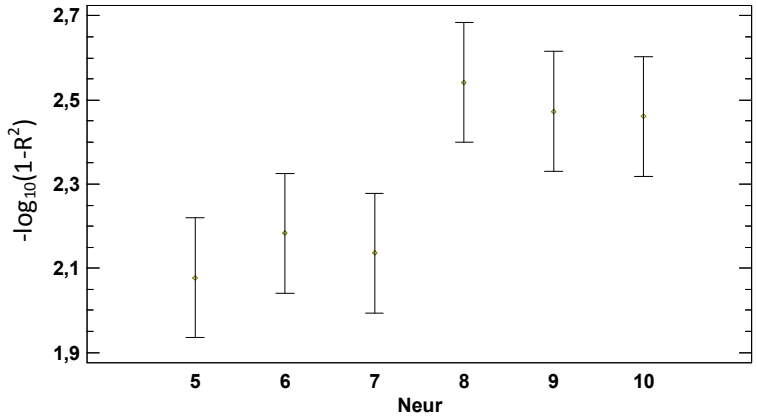


Fig. 5. Means and LSD intervals for $-\log_{10}(1-R^2)$ as a function of the number of neurons in the intermediate layer of the artificial neural network.

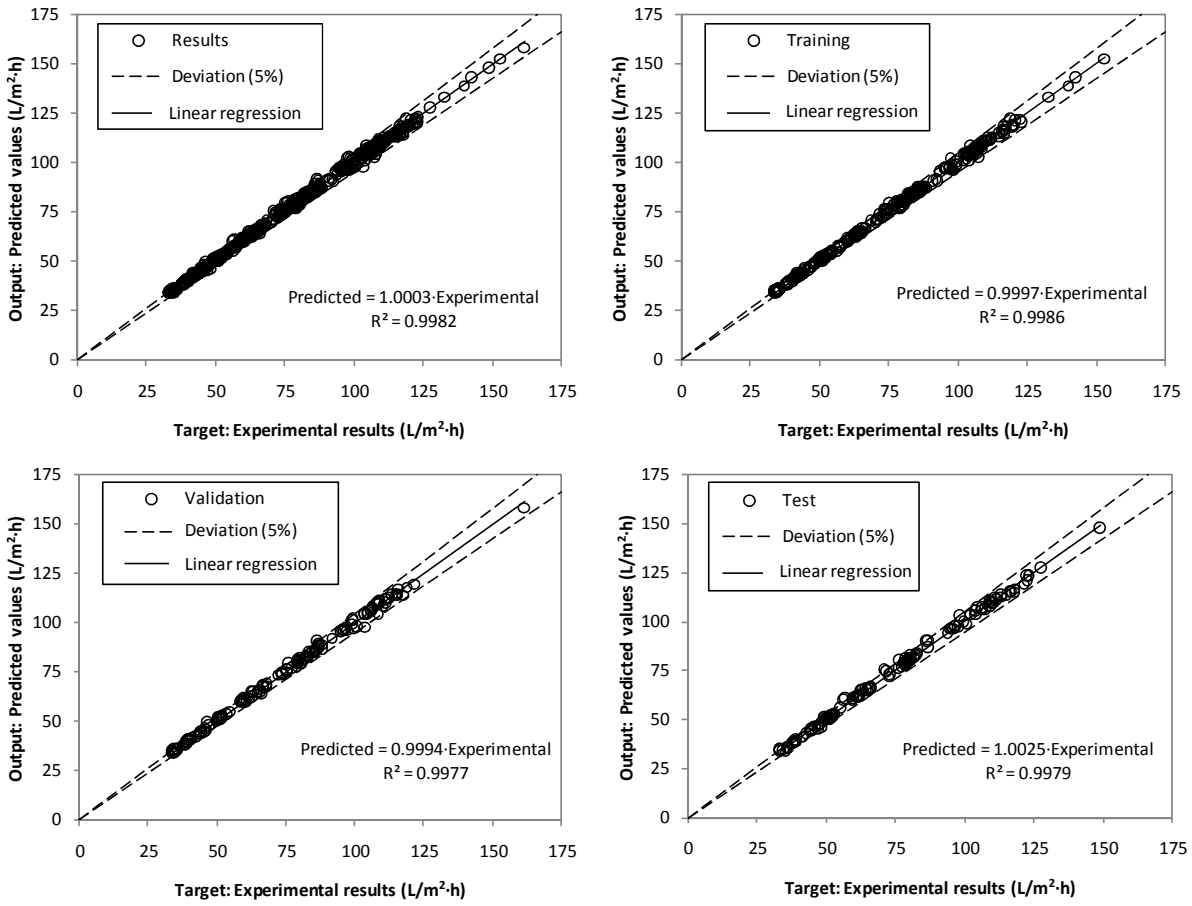


Fig. 6. Fitting accuracy for the artificial neural network model and the training, validation and test datasets

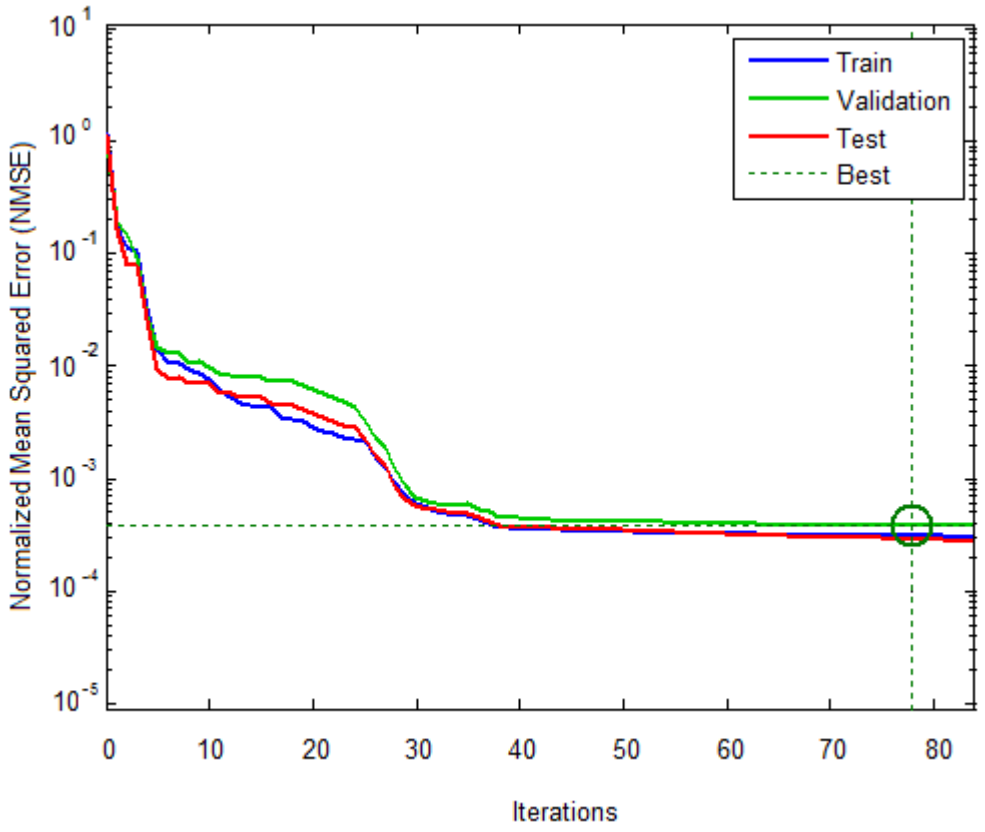
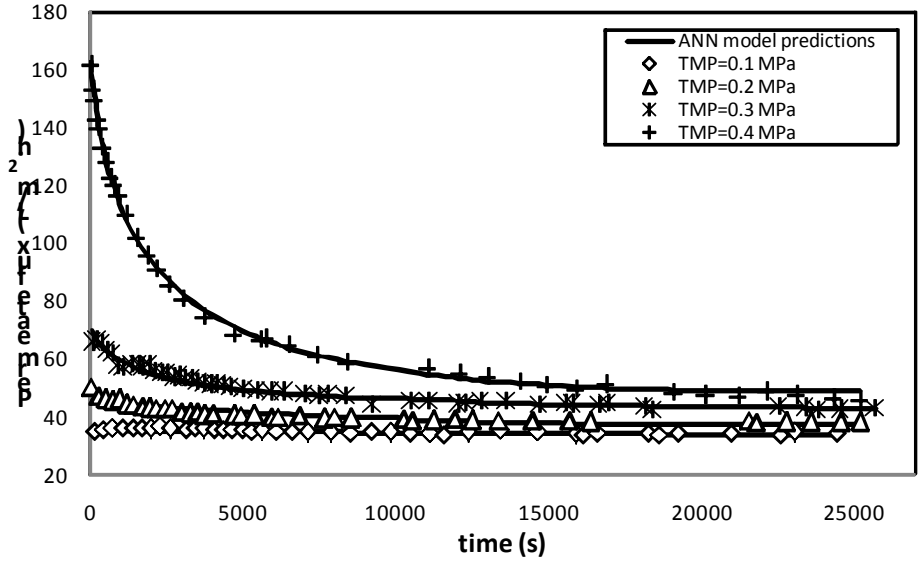
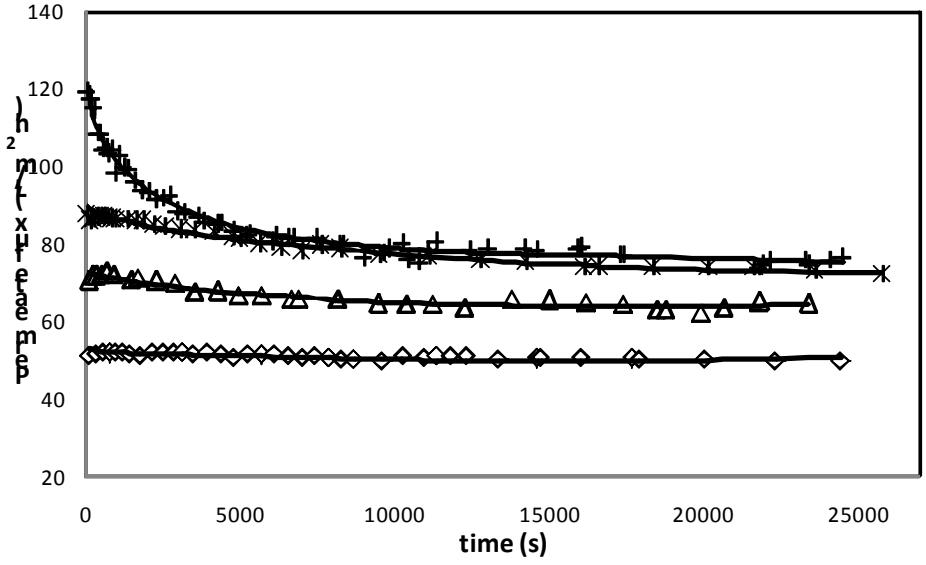


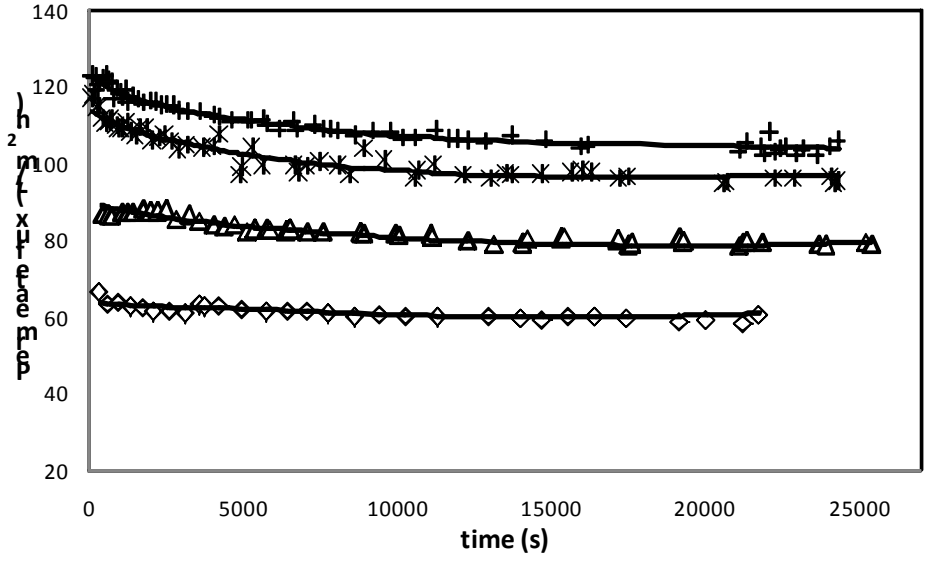
Fig. 7. Evolution of the Normalized Mean Squared Error during the training, validation and test processes.



(a)



(b)



(c)

Fig. 8. ANN simulation results for different transmembrane pressures and (a) CFV of 1 m/s, (b) CFV of 2 m/s and (c) CFV of 3 m/s (dots: experimental data; lines: artificial neural network simulation results).

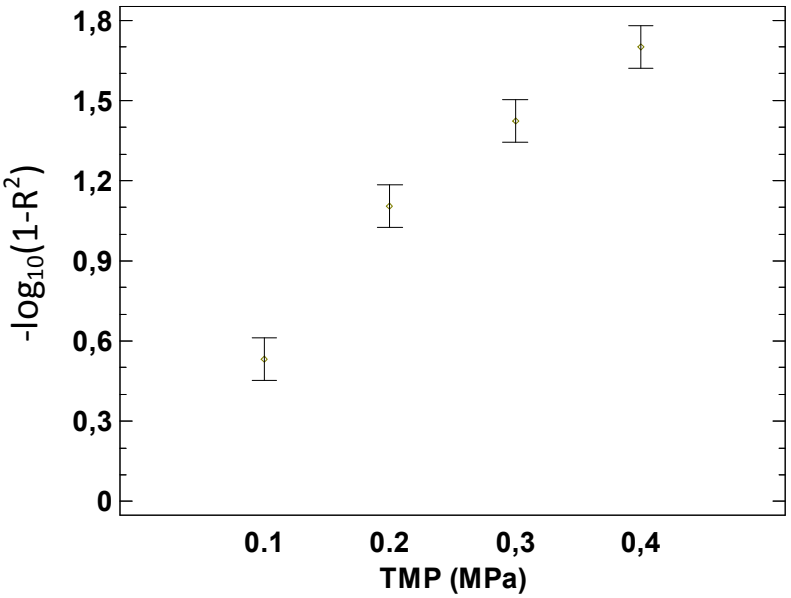


Fig. 9. Means and LSD intervals for $-\log_{10}(1-R^2)$ with TMP (MPa) for the different CFV tested (1, 2 and 3 m/s).

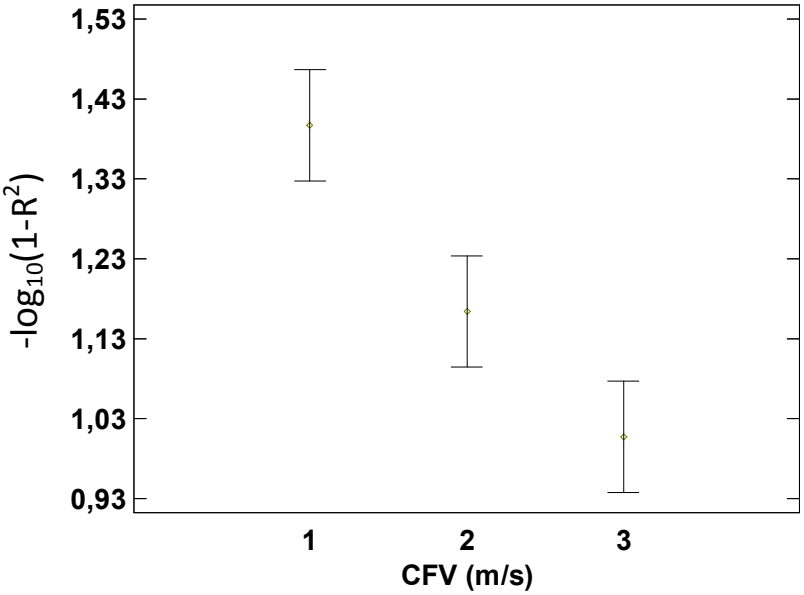


Fig. 10. Means and LSD intervals for $-\log_{10}(1-R^2)$ with CFV (m/s) for the different TMP tested (0.1 to 0.4 MPa).

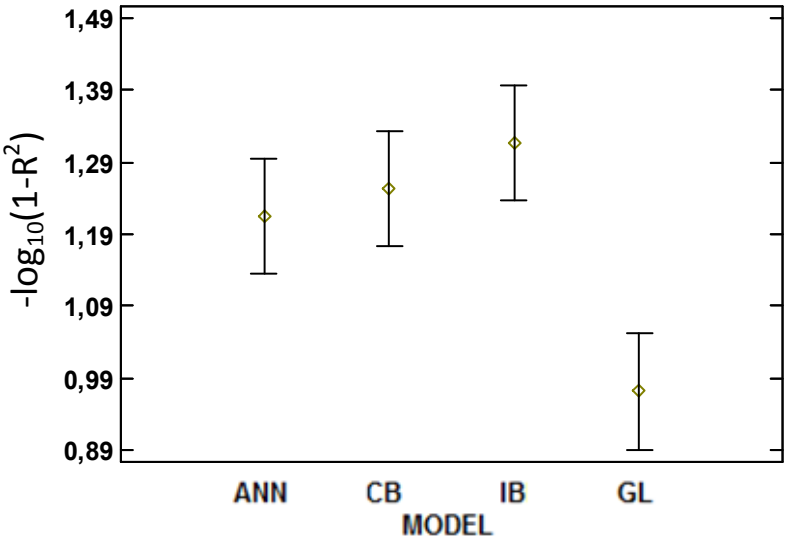


Fig. 11. Means and LSD intervals for the models employed (CB: Complete blocking; IB: Intermediate blocking; GL: Gel layer) for each combination of TMP and CFV tested.

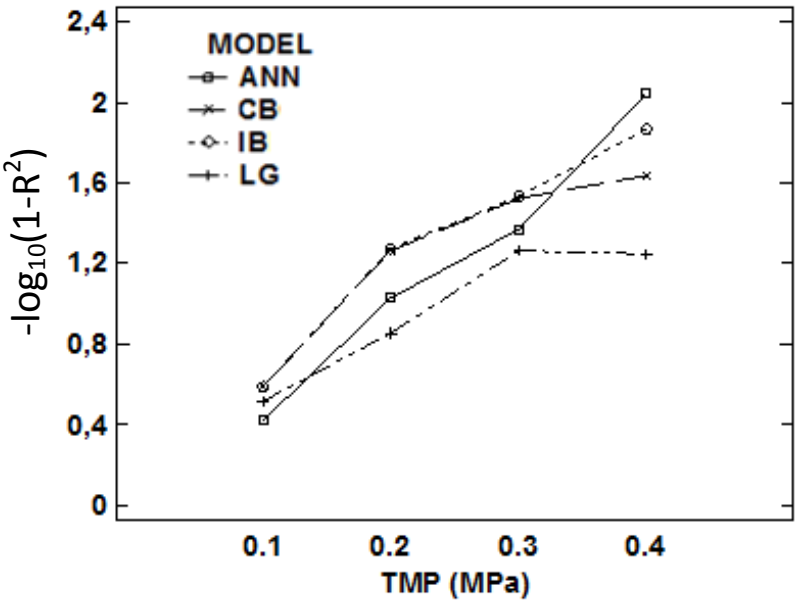


Fig. 12. Interaction between the TMP and the type of model. (CB: Complete blocking; IB: Intermediate blocking; GL: Gel layer).

Table 1. Artificial neural networks transfer functions used.

Transfer function	Layers connected	Equation
Tansig[a]	Input–Intermediate	$f(x) = \frac{2}{1+\exp(-2\cdot x)} - 1$
Purelin[b]	Intermediate–Output	f(x)=x

[a]Tansig: hyperbolic tangent sigmoid

[b]Purelin: linear transfer function

Table 2. Distribution of the data in the artificial neural network groups.

Total Data	Training data	Validation data	Test data
536	268	134	134

Table 3. Code for the methodology employed.

Stage	Abbreviation	Types	Explanation
Pretreatment	Pret	Pret 0	No pretreatment
		Pret 1	Normalization of the permeate flux
		Pret 2	Additional input: a fouling indicator
		Pret 3	Double pretreatment
Weights Initialization	Weights	0	Null initialization
		1	Random initialization
Neurons in the intermediate layer	Neur	5-10	-

Table 4. ANOVA for the response variable [$-\log_{10}(1-R^2)$] (study performed at a 95% confidence level).

Source	*Sum of Squares*	*Df*	*Mean Square*	*F-Ratio*	*p-Value*
MAIN EFFECTS					
A:Pret	3.80337	3	1.267790	5.15	0.0024
B:Weights	1.79461	1	1.794610	7.29	0.0082
C:Neur	4.87450	5	0.974901	3.96	0.0026
INTERACTIONS					
AB	0.77721	3	0.259069	1.05	0.3732
AC	2.37241	15	0.158160	0.64	0.8327
BC	1.01510	5	0.203020	0.82	0.5352
ABC	3.18590	15	0.212393	0.86	0.6068
RESIDUAL	23.6346	96	0.246193		
TOTAL (CORRECTED)	41.4577	143			

Table 5. Literature review of feed-forward artificial neural networks used in membrane processes.

Process	IL	Pret	Data Training (%)	Neur	Accuracy	Ref.
MF	1	Norm	50	3-10	MSE=0.04-0.01	[9]
UF	1	No	10 (manual)	3-11	E_{min}=1.06 E_{max}=3.61	[15]
UF	1	No	10 (manual)	3-15	E_{max}=3.0 E_m=1.0	[14]
UF and MF	2	Norm	28.6 (manual)	(4-9) (2-4)	R^2>0.99	[16]
UF and MF	1	No	16.93 (manual)		R^2=0.988 RMSE=0.082	[10]
UF and MF	2	No	16.93 (manual)	4 2	R^2=0.958 RMSE=0.156	[10]
UF	1	No	1/3	5	R^2 > 92%	[17]
NF and RO	1	Norm	80	2-7	MSE=(0.53-2.03)·10^{-4}	[13]
NF and RO	1	Norm	90	3-4	MSE=(4.33-4.67)·10^{-4}	[13]
	2			(2-6) (1-5)	MSE=(2.10-7.33)·10^{-4}	
UF	1	Norm	50	5-10	R^2 > 95% NMSE < 0.005	This study

IL= Intermediate layers in the ANN. RE =Relative Error. E_{min} =Minimum deviation. E_{max}=Maximum deviation. E_m=Average deviation. Norm=Permeate flux normalization. MSE=Mean square error. RMSE=Root mean square error. NMSE= Normalized mean square error.

Table 6. Square regression coefficient (R^2) for artificial neural networks model.

TMP (MPa)	CFV (m/s)	Square regression coefficient (R^2)
0.1	1	0.507
	2	0.549
	3	0.755
0.2	1	0.875
	2	0.917
	3	0.920
0.3	1	0.972
	2	0.967
	3	0.916
0.4	1	0.999
	2	0.985
	3	0.962

TMP=Transmembrane pressure. CFV=Crossflow velocity.

Table 7. ANOVA for $-\log_{10}(1-R^2)$ of Hermia's models adapted to crossflow ultrafiltration (except standard blocking) and ANN (95% confidence level).

Source	Sum of Squares	Df	Mean Square	F-Ratio	p-Value
MAIN EFFECTS					
A:TMP	9.06001	3	3.020000	86.37	0.0000
B:CFV	1.22697	2	0.613487	17.54	0.0001
C:MODEL	0.82673	3	0.275576	7,88	0.0015
INTERACTIONS					
AB	2.61148	6	0.4352470	12.45	0.0000
AC	0.81707	9	0.0907851	2.60	0.0406
BC	0.05448	6	0.0090800	0.26	0.9486
RESIDUAL	0.62941	18	0.0349676		
TOTAL (CORRECTED)	15.2262	47			

Table 8. Comparison between classical fouling models and artificial neural networks used in membrane technology.

Semi-empirical model			ANN			Ref.
Model	Parameter	R^2	Type	Parameter	R^2	
Hermia's models in MF	MSE_{CB}=0.042849 MSE_{IB}=0.004489 MSE_{SB}=0.003249 MSE_{GL}=0.015376	0.1867 0.9888 0.8661 0.8580	FF	MSE=0.0027	0.9940	[42]
Koltuniewicz's model in MF	-	0.914 - 0.989	FF	-	0.9440-0.9930	[43]
			FF	-	0.9670-0.9990	

FF=Feed-forward.