Comparison of response surface method and artificial neural network in predicting fluoride removal by nanofiltration

F.Z. Addar, M. Farah, M. Belfaquir, M. Tahaikt, M. Taky*, A. Elmidaoui

Laboratory of Advanced Materials and Process Engineering, Faculty of Sciences, Ibn Tofail University, P.O. Box: 1246, Kenitra, Morocco, emails: mohamed.taky@uit.ac.ma/takymogamed@gmail.com (M. Taky), Fatima.zahra@uit.ac.ma (F.Z. Addar), mohamed.farah@uit.ac.ma (M. Farah), Belfaquir.mustapha@uit.ac.ma (M. Belfaquir), mustapha.tahaikt@uit.ac.ma (M. Tahaikt), elmidaoui@uit.ac.ma (A. Elmidaoui)

ABSTRACT

During the last two decades, the theoretical calculation has emerged as one of the most effective approaches to model, predict and optimize the behavior of chemical processes, which could save time and money. Some of the models used are the response surface methodology (RSM) and the artificial neural network (ANN). The present study aims to compare the predictive efficiencies of the RSM and the ANN models applied to the removal of fluoride ions from NaF-doped groundwater by nanofiltration (NF) using three membranes (TR60, NF270 and NF90). An RSM-based central composite model (CCD) and ANN-based on feed-forward, back propagation network (FFBBN) are used in which the effects of input variables are initial fluoride concentration (IC) and transmembrane pressure (TMP) on the fluoride rejection that is considered as a response. The two methodologies are compared for their predictive abilities in terms of root mean square error (RMSE), coefficient of determination (R^2) and average absolute deviation (AAD). For RSM model, a regression coefficient $R^2 > 0.83$ is obtained for fluoride rejection efficiency for all three membranes and both parameters (IC and TMP) have a significant effect on fluoride rejection for both membranes (TR60 and NF270), whereas for the NF90 membrane they have a slight effect. The ANN model shows excellent prediction of fluoride rejection with correlation coefficient values close to unity ($R^2 > 0.998$) for the three membranes. In terms of comparison and based on the estimation parameters (RMSE, R^2 and AAD), both models show good predictions for fluoride rejection. while, the ANN model proves to be more accurate compared to the RSM model. Furthermore, RSM has the advantage of providing a regression equation for prediction and shows the effect of experimental factors and their interactions on the response compared to ANN.

Keywords: Response surface methodology (RSM); Artificial neural networks (ANN); Nanofiltration; Fluoride removal; Modelling; Prediction

1. Introduction

The third pillar of science and engineering is mathematical modeling. It aims to describe the different aspects of the real world, their interaction and their dynamics through mathematics, accomplishing the two more traditional disciplines of theoretical analysis and experimentation [1]. It is a representation or an abstract interpretation of physical reality, accessible to analysis and calculation, which uses simulation. Indeed, there is a narrow relationship between modeling and simulation, in particular to simulate a system, a model and/or several models are needed [2,3]. These models play an important role in the estimation and optimization of any system, leading to efficient and economical designs of whatever field [4,5]. Among its models, statistical models can be defined as a field that studies phenomena

* Corresponding author.

through the collection, processing, analysis, interpretation and presentation of data in order to make them understandable to all [6]. Applied statistics is used in almost all fields of human activity: engineering, management, economy, biology, computer science, physics and chemistry [7].

In the field of membrane science and technology, modeling has become a valuable tool for the prediction of membrane separation processes. Some of the modeling tools that are capable of solving linear and nonlinear multivariate regression problems are the RSM and ANN. Both approaches do not require an expressive physical meaning of the system or process under consideration. They determine a relationship between the system variables and the response [8–11]. In addition, these techniques are very useful tools for reducing the time and cost of studies [12].

RSM is considered a fast and useful procedure for the study, optimization and modeling of complex processes [13]. It is based on experimental data using the basic principles of statistical design, namely design of experiments (DoE), as well as regression modeling and process optimization techniques [14,15]. There are several types of response surface designs, including composite central designs (CCD), Box–Behnken designs (BBD), and Doehlert designs (DD) [16]. CCD is an extremely efficient and practical method to design and study an experimental space. It is used to design experiments with minimal calculation effort. In addition, the CCD defines center points and star points that increase the ability to estimate, optimize and predict responses with high accuracy and to observe details [17].

ANN is a statistical modeling tool that was developed to understand nonlinear multi-variable systems [18]. It has been used in many areas of science and engineering [19]. A typical ANN consists of an input layer, one or more hidden layers with many hidden neurons and an output layer. These layers are mathematically linked by weights and biases. There are several neural network models, such as feedforward model (FF), multilayer perceptron (MLP), and radial basis function (RBF) [20]. The most commonly used neural network algorithm is the back-propagation neural network (FFBPN), including Levenberg–Marquardt (LM) [20]. The main use of FFBPN is to learn and map the relationships between inputs and outputs to achieve minimum error [21].

Due to their effectiveness in many applications, these two approaches (RSM and ANN) are used in membrane processes to predict and model the rejection of fluoride ions by the NF process.

Fluoride is an essential constituent for humans and animals. However, the total amount ingested or its concentration in drinking water must be within certain limits. When the fluoride concentration is present in a narrow concentration range (0.7–1.5 mg/L), it plays an imperative role in bone mineralization and also acts as an antibacterial agent in the mouth. On the contrary, an excessive concentration of fluoride can have a negative effect on bones and teeth (dental and skeletal fluorosis) [5]. For this reason, the WHO has set a guideline value of 1.5 mg/L for fluoride in drinking water [22]. In many parts of the world, groundwater has very high concentrations of fluoride, so that fluoridation of drinking water and the resulting fluorosis in humans is a global problem. The latest estimates suggest that approximately 200 million people suffer the terrible fate of fluorosis

[5,18,22]. Therefore, many technologies have been developed to remove excess fluoride from drinking water, including adsorption [22], ion exchange [23], electrodialysis (ED) [24], precipitation/coagulation [25] and pressurized membrane technologies such as nanofiltration (NF) and reverse osmosis (RO) [5,18,26].

ANN and RSM are modeling methods that have been progressively applied during the last years for the simulation and optimization of separation processes, among them membrane processes, especially NF. Addar et al. [5], used the RSM based on the CCD to optimize and model the fluoride removal process by three NF membranes by employing as input variable (TMP and initial fluoride concentration) and as output variable the permeate fluoride concentration, fluoride rejections and permeate flux. The analysis of variance shows a high value of the coefficient of determination ($R^2 > 0.83$), for the three membranes and for the three responses thus ensuring a satisfactory fit of the second-order regression model with the experimental data. In another study, Addar et al. [18], investigated groundwater defluoridation by NF by testing three membranes (NF90, NF270 and TR60) where they compared one statistical method (ANN) and the other mathematical, coupled Film Theory model with Nernst–Planck equation (NP-FT) to predict and explain the variation of fluoride rejections as a function of permeate flux. For ANN, the results obtained showed a perfect correlation (output exactly equal to the target) with R^2 values > 0.9483, in terms of comparison between the two models used, the ANN model presents a superiority in the prediction of fluoride rejection. Jadhav et al. [27], Used RSM-based CCD to modeling the removal of multiple contaminants such as fluoride, arsenic, sulfate and nitrate by two membranes (NF90 and NF270). The significance of the quadratic model is determined by the F-value of the model, a large F-value (85189.92 for NF90 and 6352140.52 for NF270) is obtained indicates that the model is significant for both membranes. Bowen et al. [28], applied ANN to provide a means of modeling the performance of the NF process, thus used ANN to predict the rejection of single salts (NaCl, Na_2SO_4, $MgCl_2$ and $MgSO_4$) and mixtures of these salts at a spiral rolled NF membrane. They found good agreement between ANNs predictions and experimental data single salts and mixtures. Srivastava et al. [29], used predictive models based on machine learning techniques such as RSM and ANN to predict the permeate flow rate, water recovery, salt rejection, and specific energy consumption (SEC) of the RO and NF pilot plants, in order to optimize and compare the RO and NF for better performance. They found that the difference between the RSM and ANN predictions is small for the two pilot plants. Emami et al. [30] studied the prediction and optimization of removal efficiency and permeate flux behavior of aqueous Pb^{2+} solution in NF process using RSM and different multilayer perceptron neural network (MLP-ANN) structures. Regression coefficients of $R^2 = 0.99$ and $R^2 = 0.9986$ were obtained for the RSM model and for the best (MLP-ANN) structure, respectively. Moreover, the comparison between the models showed that the MLP-ANN model is more precise than the RSM model in predicting the empirical data.

The objective of our work is to compare two statistical mathematical methods (RSM and ANN) in describing

and modeling the phenomenon of fluoride ion removal by NF using three membranes (TR60, NF270 and NF90), for groundwater doped with different initial fluoride concentrations (0.5, 10, 15 and 20 mg/L). The RSM and ANN are used by employing as input variable (TMP and IC) and as output variable the fluoride rejections. The results obtained by these two methods are discussed and analyzed. The comparison of the efficiency of these two methods is based on the calculation of the different estimation parameters (RMSE, ADD, and R^2), this is the first report comparing RSM and ANN in the fluoride removal by NF.

2. Experimental set-up

2.1. Characteristics of the feed water

The experiments are carried out on natural groundwater from the Benguerir region doped with NaF at different concentrations. A very slight variation in pH and conductivity were detected. The results of the feed water analysis are provided in Table 1.

2.2. NF unit pilot testing

The experiments are performed on a pilot plant NF/RO (E 3039) supplied by TIA (France), equipped with two pressure vessels in series (Fig. 1). The applied TMP can be varied in the range of 5 to 70 bar using manually operated valves. Each pressure vessel contains one element. The pressure drop is about 2 bar corresponding to 1 bar of each pressure vessel. The two spiral modules are equipped with two identical commercial membranes. The water to be treated is taken from the tank by a pump and admitted to the first vessel, the retentate is admitted to the second vessel and the two permeates are recovered and mixed.

The washing is carried out by a basic solution of sodium hydroxide NaOH at pH between 9 and 10 for 10 min, followed by a rinse with water, then a washing with a solution of sulfuric acid H_2SO_4 at pH between 3 and 4 for 15 min.

The temperature is maintained at 29°C using the heat exchanger. Permeate samples are collected and the water parameters are determined analytically following the standard methods previously described [31,32]. The other parameters followed are the ones listed below:

The permeate flux is given by Eq. (1) [5,18]:

$$J_v = \frac{Q_P}{S}\left(\frac{L}{m^2 \cdot s} \text{ or } \frac{m^3}{m^2 \cdot s}\right) \tag{1}$$

where S is the membrane area (m²) and Q_P the flow rate of the permeate (L/h or m³/s).

The recovery rate (Y) is defined as:

$$Y(\%) = \frac{Q_p}{Q_0} \times 100 \tag{2}$$

where Q_P is the permeate flow (L/h) and Q_0 the feed flow (L/h).

Salt rejection (R) is defined as:

$$R = \left(1 - \frac{C_p}{C_0}\right) \times 100 \tag{3}$$

where C_p is the solute concentration in permeate (g/L) and C_0 the solute concentration in the feed water (g/L).
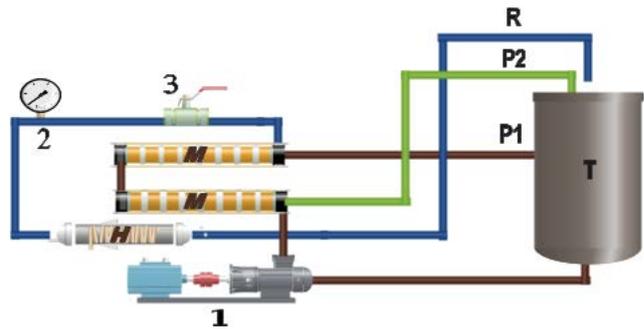
Fig. 1. Schematic diagram of the NF/RO pilot plant [5,18]. T: Tank; M: NF module; P: Permeate recirculation; R: Retentate recirculation; H: Heat exchanger; 1: High pressure pump; 2: Pressure sensor; 3: Pressure regulation valves.

Table 1
Characteristics of the feed water

Parameters	Feed water	Moroccan guidelines [5]	WHO standards [18]
Temperature (°C)	29	–	–
Turbidity, NTU	<2	–	–
pH	7.41	6–9.2	6.5–8.5
pHs	7.80	–	–
Electric conductivity, µS/cm	1,492	2,700	–
Hardness, mg/L $CaCO_3$	440	500	500
Alkalinity, mg/L $CaCO_3$	320	200	–
Fluoride, mg/L	0.5, 5, 10, 15, 20	1.5	1.5
Sulphate, mg/L	116	200	200
Nitrate, mg/L	20	50	50
Chloride, mg/L	560	750	250
Sodium, mg/L	246	–	–

2.3. Characteristics of the membranes

The two modules are equipped with two identical spiral wound NF membranes. Table 2 gives the characteristics of the three NF membranes used. After the run, membranes are cleaned with alkaline and acidic cleaning solutions according to the manufacturer recommendations.

2.4. Modelling by RSM

The experimental design of fluoride removal is performed using RSM. In this study, CCD, which is a well exploited model of RSM is employed to modeling fluoride removal by NF using three membranes. The input parameters used: TMP (X_1) and IC (X_2) which are varied at five different levels (–1.14, –1, 0, +1, + 1.14), as shown in Table 3.

The process response studied for the model is fluoride rejection. The sum of the series of experimental designs N, can be evaluated using Eq. (4):

$$N = 2^k + 2k + \text{Nc} \tag{4}$$

where k is the number of input factors. The terms 2^k, $2k$ and Nc represent the factorial points, the axial points, and the center points, respectively.

The experimental data are collected after 13 experimental runs have been completed. The experimental results for all coded factors and the actual NF response values for the three membranes are presented in Table 4. The performance of each NF membrane is evaluated in terms of rejection of fluoride (Y_1) which is considered a response. The following polynomial equation describes the predicted values of the Y_1 response as:

$$Y_1 = \beta_0 + \sum_{i=1}^{2} \beta_i X_i + \sum_{i=1}^{2} \sum_{j=1}^{2} \beta_{ij} X_i X_j + \sum_{i=1}^{2} \beta_{ii} X_i^2 + \xi \tag{5}$$

where Y_1 is the predicted response, β_0 is the constant coefficient, β_i is the linear coefficients, β_{ij} is the interaction coefficients, β_{ii} is the quadratic coefficients, and X_i and X_j are the coded values of the variables TMP and IC, ξ is the residual term (followed by the equation). The experimental design and the analysis of the experimental data are carried out by

software (Design–Expert). Model adjustment and significance are determined by analysis of variance (ANOVA).

2.5. Modelling by ANN

The FFBPN is used for its capacity to model any function. Fig. 2 shows the general framework of the FFBPN model. In addition, the FFBPN learning rule is used to adjust the weights and threshold values of a system to obtain the minimum possible error [33]. The input neurons have received the experimental data and the network has given its outputs (ANN simulation data). If the ANN output is not equal to the measured experimental outputs, then our procedure calculates the mean square error between both values and modifies the ANN weights to minimise it. The learning data is normalised to the range {0–1}. Sixty percent of the Collected data is used for model training, while 30% of the data is equally divided for testing and validation, respectively. The model is then trained in accordance with Eq. (6) until the mean square error is (or becomes) minimal [33].

$$x_J = \sum_{i}^{n} W_{ji} x_i \tag{6}$$

where x_J is the variable's new value, x_i is the variable's initial value and W_{ji} is the neuron/variable's weight link value. The activation function between the input and the hidden layer is (tansig) and log-sigmoid (logsig), as indicated by Eq. (7).

Table 3
Independent input variables range in terms of coded levels

Factors	Coded level				
	–1.14	–1	0	+1	+1.14
IC (mg/L): TR60-NF270-NF90	0.51	3	9	15	17.48
TMP (bar) TR60	3.96	5	7.5	10	11.03
TMP (bar) NF270	2.92	5	10	15	17.07
TMP (bar) NF90	4.82	10	22.5	35	40.177

Table 2
Characteristics of the membranes used

	NF270*4040	NF90*4040	TR60*4040
Area (m²)	7.6	7.6	6.8
Salt rejection (%)	>97.0%	97%	–
P_{max} (bar)	41	41	10
Material	Polyamide	Polyamide	Polyamide
Contact angle (°) [18]	27	54	–
Zeta potential (mV) [18]			
pH = 3	4.9	3.7	–
pH = 12	–25.6	–19.4	–

Salt rejection based on the following test conditions 2.000 mg/L MgSO₄, 77°F (25°C), and 15% recovery rate at TMP 4.8 bar.
Salt rejection based on the following test conditions 2.000 mg/L NaCl, 77°F (25°C), and 15% recovery rate at TMP 10 bar.
Salt rejection based on the following test conditions 2.000 mg/L NaCl, 77°F (25°C), and 15% recovery rate at TMP 15.5 bar.

Table 4
CCD design matrix of two variables and the experimentally determined, RSM predicted and ANN predicted values for three membranes

Run	Coded variables values		Responses values								
			TR60			NF270			NF90		
	X_1	X_2	Y_{Exp}	Y_{RSM}	Y_{ANN}	Y_{Exp}	Y_{RSM}	Y_{ANN}	Y_{Exp}	Y_{RSM}	Y_{ANN}
1	0	0	0.765	0.765	0.765	0.805	0.805	0.804	0.99	0.986	0.990
2	0	0	0.710	0.725	0.710	0.805	0.805	0.804	0.995	0.989	0.994
3	–1	1	0.827	0.860	0.826	0.806	0.765	0.805	0.993	0.991	0.993
4	1	–1	0.765	0.765	0.765	0.805	0.805	0.810	0.994	0.995	0.994
5	0	0	0.765	0.765	0.765	0.742	0.797	0.741	0.993	0.991	0.993
6	–1.14	0	0.765	0.765	0.765	0.63	0.601	0.630	0.99	0.992	0.99
7	0	0	0.718	0.716	0.718	0.805	0.805	0.806	0.977	0.985	0.976
8	0	1.14	0.765	0.765	0.765	0.805	0.805	0.806	0.995	0.997	0.995
9	1.14	0	0.766	0.730	0.766	0.42	0.434	0.422	0.982	0.983	0.981
10	1	1	0.723	0.709	0.723	0.742	0.753	0.740	0.993	0.991	0.993
11	0	0	0.824	0.805	0.824	0.734	0.689	0.734	0.993	0.991	0.993
12	0	–1.14	0.55	0.572	0.549	0.466	0.481	0.466	0.995	0.998	0.995
13	–1	–1	0.6	0.597	0.6	0.726	0.743	0.726	0.993	0.991	0.993
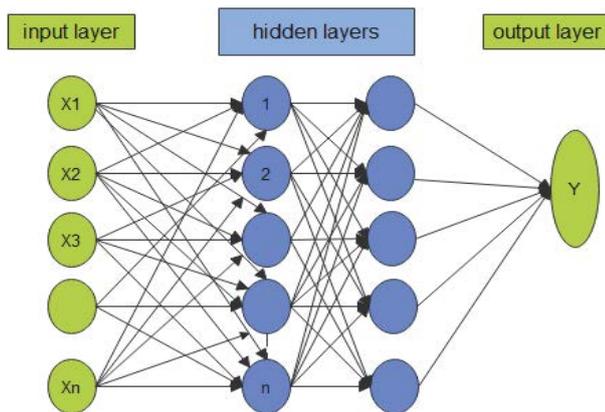
Fig. 2. Proposed neural network architecture [18].

Fig. 3. Effect of TMP on permeate flux for the three membranes [5,18].

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \text{ and } f(x) = \frac{1}{1 + e^{-x}} \tag{7}$$

The purelin function is used as Eq. (8) between the hidden layer and the output layer:

$$f(x) = x \tag{8}$$

Finally, the reliability of the model is verified with a new (unknown) dataset and the results are found to be satisfactory. The justification for these steps is discussed in the results section. The FFBPN used as an ANN network to predict fluoride rejection by the three NF membranes tested, uses the same input and output as the RSM-CCD. The results obtained are mentioned in Table 4.

3. Results and discussion

3.1. Effect of TMP on the flux of permeate

The relationship between permeate flux and TMP, illustrated in Fig. 3, was investigated using batch mode for groundwater samples from the Benguerir region, which were doped with various fluorides. This dataset has already been published in our previous papers [5,18].

As shown in Fig. 3, the permeate flux increases almost linearly with TMP according to Darcy's law [18]. This flow behavior is well established in the literature [5,18]. In this illustration, the permeate flux follows the following order: NF270 > TR60 > NF90.

3.2. Effect of TMP and IC on the fluoride rejection

3.2.1. Modeling by RSM

3.2.1.1. Validation by ANOVA

The significance and performance of the regression model are examined by ANOVA analysis of variance for

fluoride removal by NF, the results are presented in Table 5. The P values for fluoride rejection for all three membranes are strictly below the significance level (α = 0.05), indicating that the mathematical models or regression for the response fit the experimental data extremely closely. The reliability of the model and the quality of the adjustment of the model values to the experimental data is demonstrated by the R-squared correlation coefficient which is greater than 0.83 for all three membranes. These results indicate that more than 83% of the sample variation for fluoride rejection is attributed to the two factors.

3.2.1.2. Regression equation

A quadratic polynomial is established to identify the relationship between the Y_i response and the various factors. The fluoride rejection regression equations generated for TR60, NF270 and NF90 are Eqs. (9)–(11), respectively.

$$Y_1 = +42.50729 + 0.485391X_1 + 4.52497X_2$$
$$-0.088667X_1X_2 + 0.163200X_1^2 - 0.167500X_2^2 \qquad (9)$$

$$Y_1 = +6.22187 + 5.13522X_1 + 8.43217X_2$$
$$-0.136167X_1X_2 - 0.176475X_1^2 - 0.284705X_2^2 \qquad (10)$$

$$Y_1 = +96.88400 + 0.048940X_1 + 0.283158X_2$$
$$-0.002533X_1X_2 - 0.000040X_1^2 - 0.008646X_2^2 \qquad (11)$$

According to regression Eqs. (9)–(11), for NF270 and TR60 membranes, the variables X_1 and X_2 have positive effects on Y_1, since their associated regression coefficients are positive. The variables X_1X_2 and X_2^2 have negative effects on Y_1, however for X_1^2 has an opposite effect on Y_1. In addition, the variables (X_1X_2, X_1^2 and X_1^2) have insignificant effects on Y_1 since their coefficients are close to zero. A positive value represents an effect that favors optimization, while a negative value indicates an inverse relationship between the factor and the response [16,34]. This is confirmed by the overall analysis of the ANOVA results. The variables found by the NF270 membrane are higher than those found by the TR60 membrane. For the NF90 membrane variables are all almost zero.

Fig. 4 compares the experimental rejection values of fluoride ions to the predicted data for three membranes.

Fig. 4 shows the plot of predicted vs. actual response values. The predicted values were uniformly and closely distributed to the actual responses and exhibited reasonable agreement (R^2 > 0.83) for all three membranes. This shows that the regression models generated can effectively describe the relationship between the factors and the responses in the range studied. The distribution of the data points were almost uniformly close to a straight line.

3.2.1.3. Fluoride rejection

Figs. 5–7 show the 3D response surfaces and 2D contour plot for the interaction effect of two parameters (TMP and IC) on the response for the three membranes.

From Figs. 5–7 we can notice:

- For TR60 and NF270, have a significant effect of IC on fluoride rejection. Whereas the effect of TMP less important. In addition, TMP seems to affect fluoride rejection more when IC is high than when it is low.
- For NF90, both parameters slightly affect the fluoride rejection,
- For the range of IC and TMP studied, the fluoride rejection follows the sequence: $R_{NF90} > R_{TR60} > R_{NF270}$.

All results of the 3D response surface and 2D contour plot for the interaction effect of two parameters for the three membranes are already confirmed by the regression equations.

3.2.2. Modeling by ANN

3.2.2.1. Fluoride rejection

Using ANN with 10 hidden neurons, as input IC and TMP are taken, and as output the fluoride rejections and the results obtained are mentioned in Table 4. Fig. 8 shows the fluoride rejection predicted by the ANN model compared to the experimental values for the three membrane, the operation of prediction is performed on the MATLAB software.

According to Fig. 8, for the three membranes, all the points are located very close to the straight line, which

Table 5
ANOVA of fluoride rejection response for the three membranes

Response	Membranes	Variation source	Sum of squares	Degree of freedom	Mean square	F-value	P-value	R^2
Fluoride rejection	TR60	Regression	716.74	5	143.35	27.27	0.0002	0.9512
		Residual	85.76	7	12.25	–	–	–
		Total	1,056.57	12	–	–	–	–
	NF270	Regression	1,983.79	5	396.76	33.12	<0.0001	0.9594
		Residual	121.37	7	17.34	–	–	–
		Total	343.05	12	–	–	–	–
	NF90	Regression	3.00	5	0.6000	6.93	0.0122	0.8320
		Residual	0.5620	7	0.0803	–	–	–
		Total	3.73	12	–	–	–	–

Fig. 4. Predicted vs. actual values for fluoride rejection.
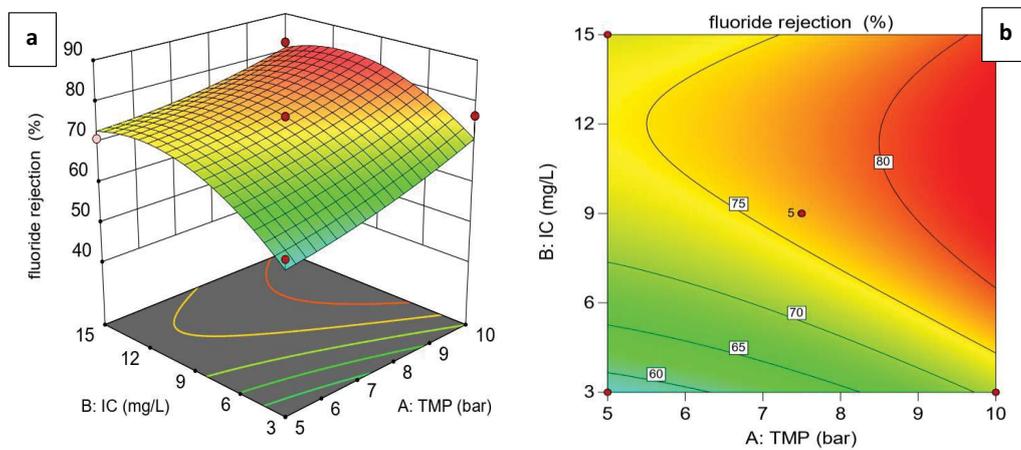
Fig. 5. 3D response surface (a) and 2D contour plot (b) for the interaction effect of two parameters on the response for TR60 membrane.
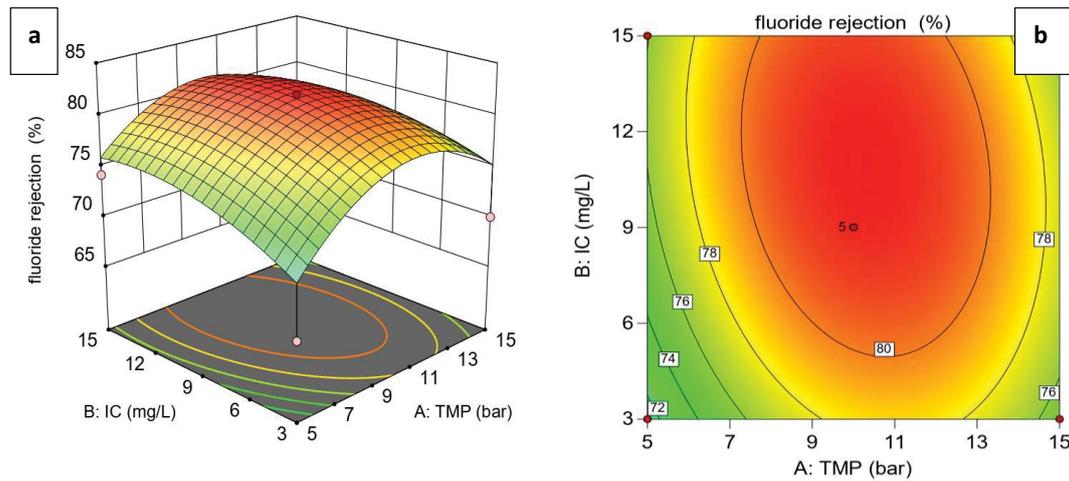
Fig. 6. 3D response surface (a) and 2D contour plot (b) for the interaction effect of two parameters on the response for NF270 membrane.
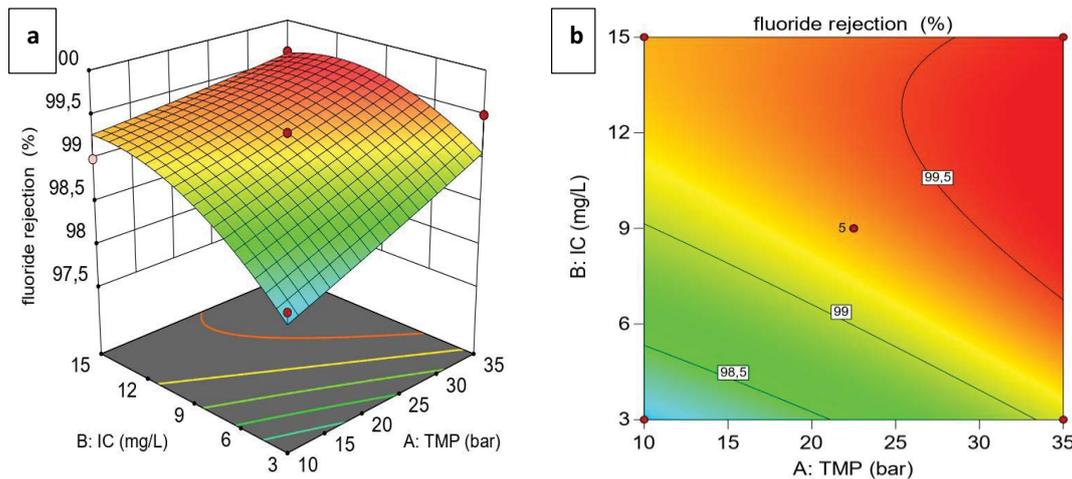
Fig. 7. 3D response surface (a) and 2D contour plot (b) for the interaction effect of two parameters on the response for NF90 membrane.

indicates that the prediction of the ANN model is excellent in the valid region and the value of the correlation coefficient is closed to unity ($R^2 > 0.998$), showing the linear relationship between the experimental and predicted fluoride rejection.

3.2.2.2. Validation of the model ANN

The results obtained for the training, the test, the validation and the global R^2 for the training data set, for the three membranes are depicted in Fig. 9.

In order to predict fluoride rejection values using the ANN model, 75% of the data are used randomly for training. The remaining data are classified as test and validation data [18]. Fig. 9 shows the summaries of the R^2 plots during the training, test and validation phases for the three membranes. The best linear fit equations for the training, validation, test, and global subsets mostly had a slope between 0.99 and 1.

The R^2 values are all greater than 0.99 for all three membranes, indicating close agreement between the experimental and modeling results. On the other hand, for all curves, the points are located very close to the straight line. Therefore, the trained ANN model shows an accurate simulation of fluoride rejection for the NF fluoride removal process.

3.2.3. Comparison of RSM and ANN

RSM and ANN are modeling models used to solve linear and nonlinear multivariate regression problems [4]. To evaluate the performance of ANN and RSM for the fluoride removal process using the three NF membranes. The predicted experimental results were evaluated in terms of estimation parameters, namely the root mean square error (RMSE), coefficient of determination (R^2) and average absolute deviation (AAD) which are defined by Eqs. (12)–(14), respectively:
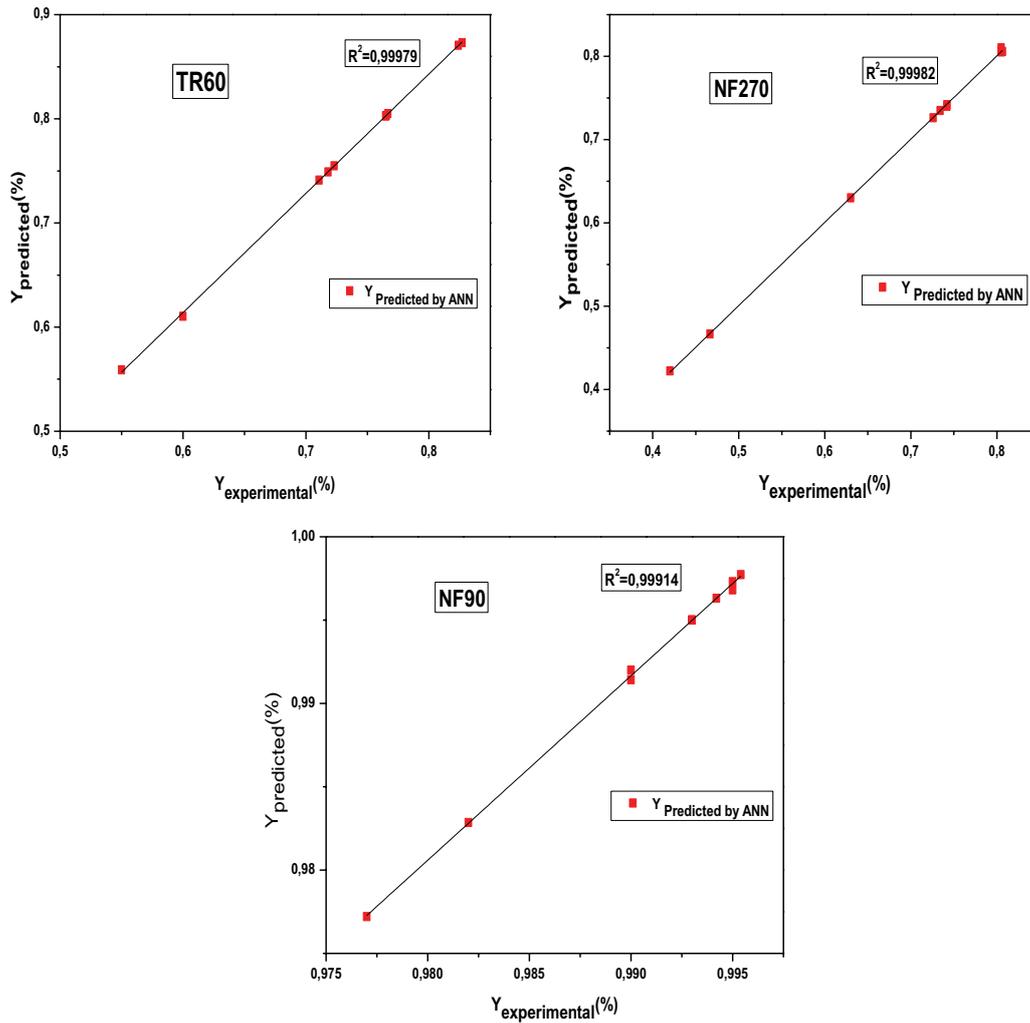
Fig. 8. Fluoride rejection predicted by the ANN model compared to experimental values.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{\text{ex}} - y_p\right)^2} \qquad (12)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_p - y_{\text{ex}}\right)^2}{\sum_{i=1}^{n}\left(y_p - y_m\right)^2} \qquad (13)$$

$$\text{ADD} = \left\{\left[\sum_{i=1}^{n}\left(\frac{\left|y_p - y_{\text{ex}}\right|}{y_{\text{ex}}}\right)\right]^{1/n}\right\} \times 100 \qquad (14)$$

where y_p is the value predicted by the ANN model, y_{ex} is the experimental value, n is the number of data, and y_m is the mean of the experimental value, the results obtained are presented in Table 6.

According to the results reported in Table 6 and the deviation of the predicted response values from the experimental data for the two models shown in Fig. 10. The RMSE and AAD values presented by ANN are lower than those of the RSM, while the R^2 values of ANN are greater than

those reported by the RSM. A value of RMSE is always positive and a value of zero would indicate a perfect fit to the data. In most cases, a smaller RMSE value indicates better accuracy than a larger RMSE value. Regarding ADD, the smaller the standard deviation, the more the values are clustered around the mean and the same is true for the inverse [34]. Both models fit the experimental data well. However, the predictive power of the ANN is found to be more strong than that of the RSM. On the other hand, RSM has the advantage of providing a regression equation for prediction and showing the effect of experimental factors and their interactions on the response compared to ANN. Nevertheless, the main limitation of the RSM method involves a quadratic nonlinear correlation only. Since ANN can inherently capture almost any form of nonlinearity, it can easily overcome the limitations of RSM [35]. Another advantage of the ANN model is its flexibility and the possibility to add new experimental data to build a reliable ANN model. On the contrary, the ANN methodology may require a larger number of experiments than RSM [36].

Table 6 shows the results obtained by RSM and ANN for the estimation parameters (RMSE, R^2 and AAD) of the
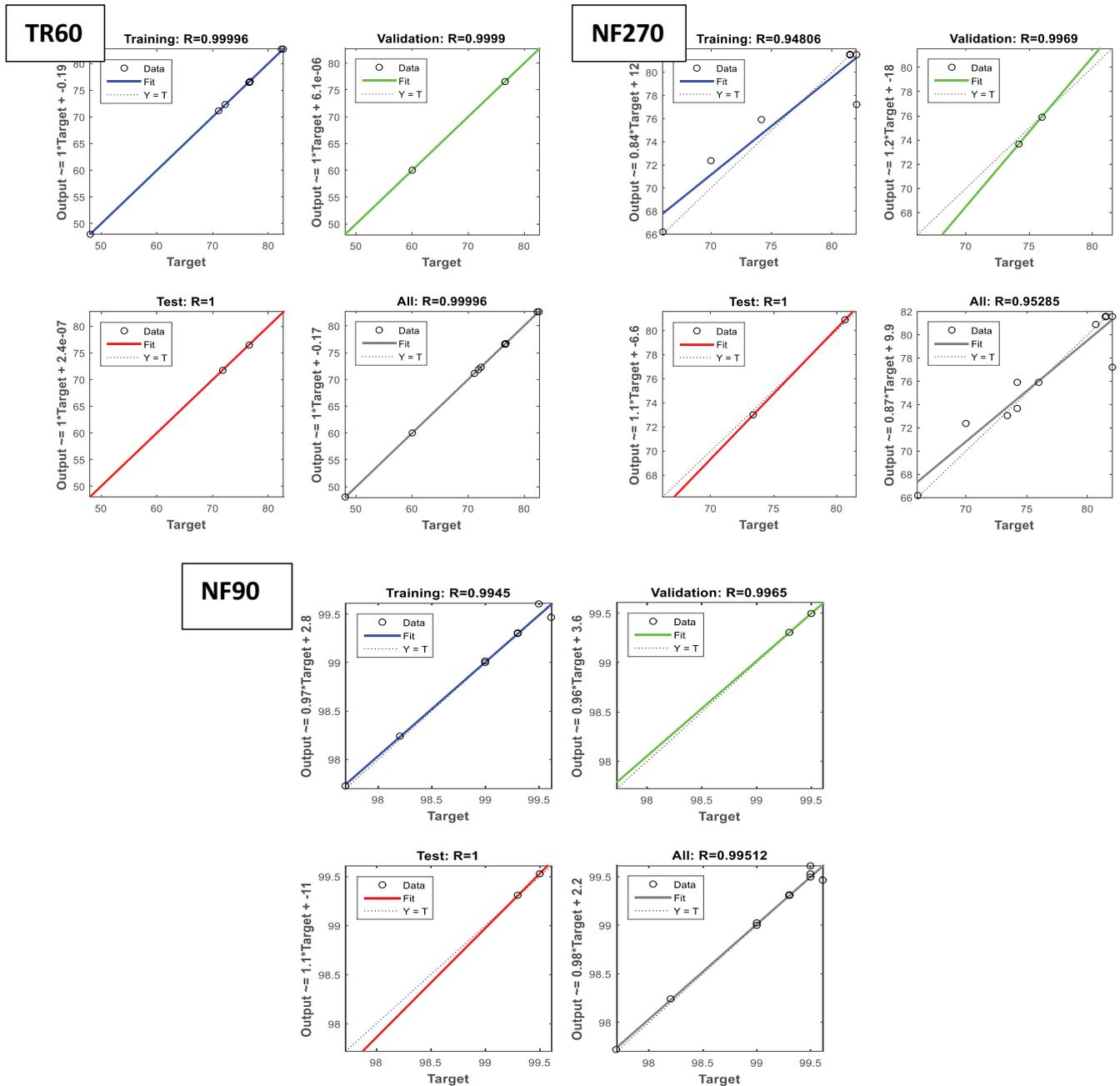
Fig. 9. Network model with training, validation, testing and prediction set for the three membranes.

Table 6
Comparison of RSM and ANN

Parameters	TR60		NF270		NF90	
	RSM	ANN	RSM	ANN	RSM	ANN
RMSE	0.0168	0.0002	0.0253	0.0017	0.0033	0.0002
R^2	0.95	0.99	0.95	0.99	0.83	0.99
AAD	0.073	0.007	0.192	0.088	0.069	0.005
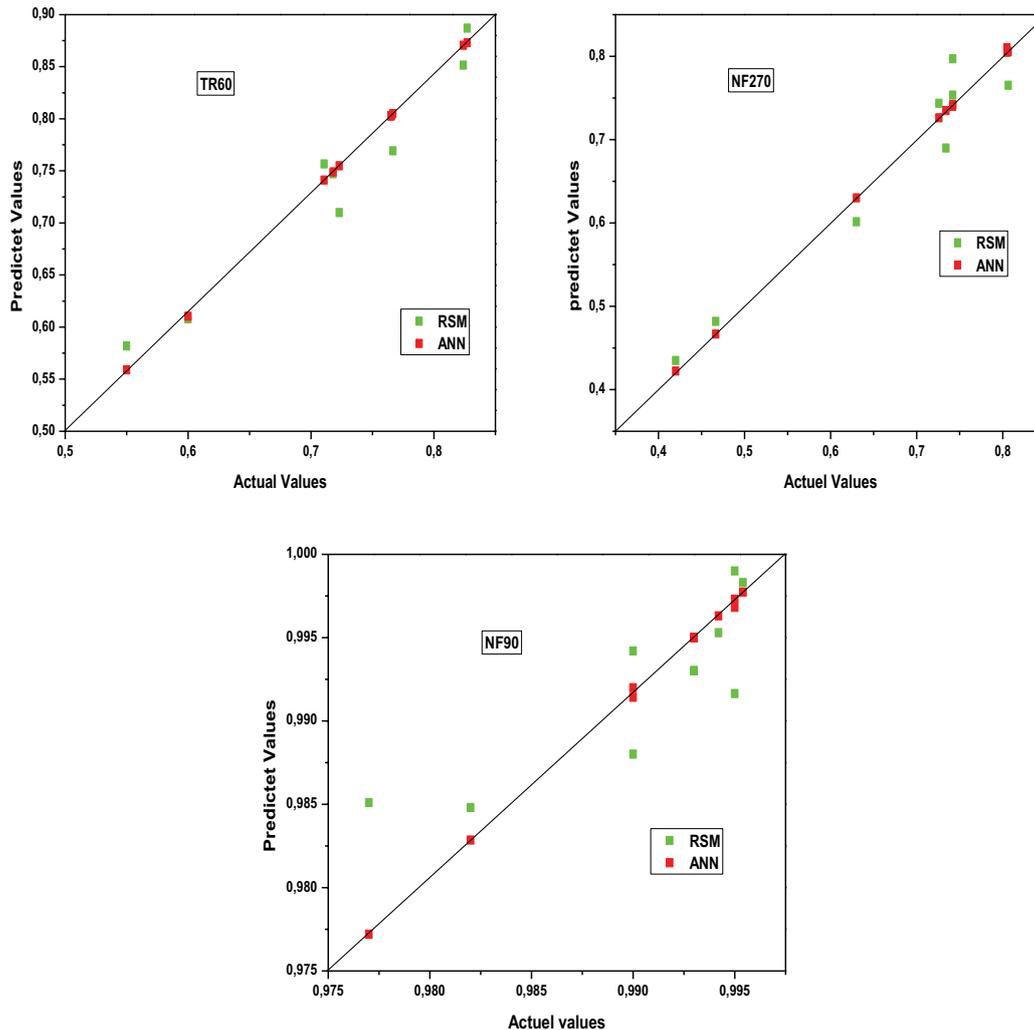
Fig. 10. Predicted values vs. actual values of ANN and RSM models for the three membranes.

three membranes. The scatter plot of the values predicted by the RSM and ANN models vs. the experimental values for the three tested membranes is shown in Fig. 10.

4. Conclusion

The study was led on NaF-doped groundwater using three membranes (TR60, NF270 and NF90), which aim to apply two statistical methods RSM and ANN, in order to predict the fluoride rejection as a function of the input variables (IC and TMP). The following conclusions were reached from this work:

- The RSM model indicated the variables IC and TMP have a significant effect on fluoride rejection for the TR60 and NF270 membranes, while for the NF90 membrane they have a minor effect. In addition, under optimized conditions, fluoride rejection obtained are 79.69%, 72% and 98.75% for TR60, NF270 and NF90, respectively.
- Both models showed good predictions for three membranes with a coefficient of determination that exceeded 0.83, 0.99 for RSM and ANN, respectively.

- In terms of comparison based on estimation parameters (RMSE, R^2 and AAD), the two models showed good predictions for fluoride rejection, while the ANN model proved to be more accurate than the RSM model.
- RSM has the advantage of providing a regression equation for the prediction and showing the effect of experimental factors and their interactions on the response compared to ANN.

References

[1] A. Quarteroni, Mathematical models in science and engineering, Not. AMS, 56 (2009) 10–19.
[2] D. Hocking, J.M. Nougués, J.C. Rodríguez, S. Sama, Chapter 3.3 – Simulation, Design and Analysis, B. Braunschweig, R. Gani, Eds., Computer Aided Chemical Engineering, P.O. Box: 211.1000 AE, Amsterdam, The Netherlands, 2002, pp. 165–191.
[3] S. Pierucci, G. Buzzi Ferraris, 20th European Symposium of Computer Aided Process Engineering, Radarweg 29, P.O. BOX: 211, 1000AE, Amsterdam, The Netherlands, Linacre House, Jordan Hill, Oxford OX28DP, UK, 2010, pp. 1–1342.
[4] A. Fetimi, A. Dâas, Y. Benguerba, S. Merouani, M. Hamachi, O. Kebiche-Senhadji, O. Hamdaoui, Optimization and prediction of safranin-O cationic dye removal from aqueous

solution by emulsion liquid membrane (ELM) using artificial neural network-particle swarm optimization (ANN-PSO) hybrid model and response surface methodology (RSM), J. Environ. Chem. Eng., 9 (2021) 105837, doi: 10.1016/j.jece.2021.105837.

[5] F.Z. Addar, S. El-Ghizel, M. Tahaikt, M. Belfaquir, M. Taky, A. Elmidaoui, Fluoride removal by nanofiltration: experimentation, modelling and prediction based on the surface response method, Desal. Water Treat., 240 (2021) 75–88.

[6] G.M. Henkin, A.A. Shananin, Asymptotic behavior of solutions of the Cauchy problem for Burgers type equations, J. Math. Pures Appl., 83 (2004) 1457–1500.

[7] O. Martin, L'Enquête et ses méthodes: l'analyse de données quantitatives, Paris, Armand Colin, 2005; 2009.

[8] R.J. Fox, D. Elgart, S. Christopher Davis, Bayesian credible intervals for response surface optima, J. Stat. Plann. Inference, 139 (2009) 2498–2501.

[9] S. Ahmadi, M. Mesbah, C.A. Igwegbe, C.D. Ezeliora, C. Osagie, N.A. Khan, G.L. Dotto, M. Salari, M.H. Dehghani, Sono electro-chemical synthesis of LaFeO$_3$ nanoparticles for the removal of fluoride: optimization and modeling using RSM, ANN and GA tools, J. Environ. Chem. Eng., 9 (2021) 105320, doi: 10.1016/j.jece.2021.105320.

[10] S. Chellapan, D. Datta, S. Kumar, H. Uslu, Statistical modeling and optimization of itaconic acid reactive extraction using response surface methodology (RSM) and artificial neural network (ANN), Chem. Data Collect., 37 (2022) 100806, doi: 10.1016/j.cdc.2021.100806.

[11] J. Prakash Maran, B. Priya, Comparison of response surface methodology and artificial neural network approach towards efficient ultrasound-assisted biodiesel production from muskmelon oil, Ultrason. Sonochem., 23 (2015) 192–200.

[12] I. Salehi, M. Shirani, A. Semnani, M. Hassani, S. Habibollahi, Comparative study between response surface methodology and artificial neural network for adsorption of crystal violet on magnetic activated carbon, Arabian J. Sci. Eng., 41 (2016) 2611–2621.

[13] P. Naderi, M. Shirani, A. Semnani, A. Goli, Efficient removal of crystal violet from aqueous solutions with centaurea stem as a novel biodegradable bioadsorbent using response surface methodology and simulated annealing: kinetic, isotherm and thermo-dynamic studies, Ecotoxicol. Environ. Saf., 163 (2018) 372–381.

[14] A. Ghazali, M. Shirani, A. Semnani, V. Zare-Shahabadi, M. Nekoeinia, Optimization of crystal violet adsorption onto date palm leaves as a potent biosorbent from aqueous solutions using response surface methodology and ant colony, J. Environ. Chem. Eng., 6 (2018) 3942–3950.

[15] E. Alian, A. Semnani, A. Firooz, M. Shirani, B. Azmoon, Application of response surface methodology and genetic algo-rithm for optimization and determination of iron in food samples by dispersive liquid–liquid microextraction coupled UV–Visible spectrophotometry, Arabian J. Sci. Eng., 43 (2018) 229–240.

[16] M. Zait, F.Z. Addar, N. Elfilali, M. Tahaikt, A. Elmidaoui, M. Taky, Analysis and optimization of operating conditions on ultrafiltration of landfill leachate using a response surface methodological approach, Desal. Water Treat., 257 (2022) 64–75.

[17] M. Shirani, A. Semnani, S. Habibollahi, H. Haddadi, M. Narimani, Synthesis and application of magnetic NaY zeolite composite immobilized with ionic liquid for adsorption desulfurization of fuel using response surface methodology, J. Porous Mater., 23 (2016) 701–712.

[18] F.Z. Addar, B. Fahid, M. Belfaquir, M. Tahaikt, A. Elmidaoui, M. Taky, Modeling of fluoride removal by nanofiltration: coupled film theory model with Nernst–Planck equation and artificial neural network, Desal. Water Treat., 257 (2022) 76–90.

[19] M. Shirani, A. Akbari, A. Goli, Application of a novel high-performance nano biosorbent for removal of anionic dyes from aqueous solutions using shuffled frog leaping algorithm: isotherm, kinetic and thermodynamic studies, Desal. Water Treat., 203 (2020) 388–402.

[20] E. Kazemi, Y. Yamini, S. Seidi, A. Mohammadi, Determination of myclobutanil in food samples using homogeneous liquid–liquid microextraction via flotation assistance and gas chromatography-mass spectrometry using artificial neural network, Int. J. Environ. Anal. Chem., 98 (2018) 271–285.

[21] M. Shirani, A. Akbari, M. Hassani, A. Goli, S. Habibollahi, P. Akbarian, Homogeneous liquid–liquid microextraction via flotation assistance coupled with gas chromatography-mass spectrometry for determination of myclobutanil in cucumber, tomato, grape, and strawberry using genetic algorithm, Int. J. Environ. Anal. Chem., 98 (2018) 271–285.

[22] M. Mourabet, A. El Rhilassi, H. El Boujaady, M. Bennani-Ziatni, R. El Hamri, A. Taitai, Removal of fluoride from aqueous solution by adsorption on apatitic tricalcium phosphate using Box–Behnken design and desirability function, Appl. Surf. Sci., 258 (2012) 4402–4410.

[23] K. Wan, L. Huang, J. Yan, B. Ma, X. Huang, Z. Luo, H. Zhang, T. Xiao, Removal of fluoride from industrial wastewater by using different adsorbents, Sci. Total Environ., 773 (2021) 145535, doi: 10.1016/j.scitotenv.2021.145535.

[24] M.A. Menkouchi Sahli, S. Annouar, M. Tahaikt, M. Mountadar, A. Soufiane, A. Elmidaoui, Fluoride removal for underground brackish water by adsorption on the natural chitosan and by electrodialysis, Desalination, 212 (2007) 37–45.

[25] Y. Huang, X. Wang, Y. Xu, S. Feng, J. Liu, H. Wang, Amino-functionalized porous PDVB with high adsorption and regeneration performance for fluoride removal from water, Green Chem. Eng., 2 (2021) 224–232.

[26] M. Tahaikt, F. Elazhar, I. Mohamed, H. Zeggar, M. Taky, A. Elmidaoui, Comparison of the performance of three nanofiltration membranes for the reduction of fluoride ions: application of the Spiegler–Kedem and steric hindrance pore models, Desal. Water Treat., 240 (2021) 14–23.

[27] S.V. Jadhav, K.V. Marathe, V.K. Rathod, A pilot scale concurrent removal of fluoride, arsenic, sulfate and nitrate by using nanofiltration: competing ion interaction and modelling approach, J. Water Process Eng., 13 (2016) 153–167.

[28] W. Richard Bowen, M.G. Jones, J.S. Welfoot, H.N.S. Yousef, Predicting salt rejections at nanofiltration membranes using artificial neural networks, Desalination, 129 (2000) 147–162.

[29] A. Srivastava, K. Aghilesh, A. Nair, S. Ram, S. Agarwal, J. Ali, R. Singh, M.C. Garg, Response surface methodology and artificial neural network modelling for the performance evaluation of pilot-scale hybrid nanofiltration (NF) & reverse osmosis (RO) membrane system for the treatment of brackish ground water, J. Environ. Manage., 278 (2021) 111497, doi: 10.1016/j.jenvman.2020.111497.

[30] M.R.S. Emami, M.K. Amiri, S.P.G. Zaferani, Removal efficiency optimization of Pb^{2+} in a nanofiltration process by MLP-ANN and RSM, Korean J. Chem. Eng., 38 (2021) 316–325.

[31] M. Tahaikt, A. Ait Haddou, R. El Habbani, Z. Amor, F. Elhannouni, M. Taky, M. Kharif, A. Boughriba, M. Hafsi, A. Elmidaoui, Comparison of the performances of three com-mercial membranes in fluoride removal by nanofiltration. Continuous operations, Desalination, 225 (2008) 209–219.

[32] M.A. Menkouchi Sahli, S. Annouarb, M. Tahaikt, M. Mountadar, A. Soufiane, A. Elmidaoui, Fluoride removal for underground brackish water by adsorption on the natural chitosan and by electrodialysis, Desalination, 212 (2007) 37–45.

[33] N.B. Shaik, S.R. Pedapati, S.A.A. Taqvi, A.R. Othman, F.A.A. Dzubir, A feed-forward back propagation neural network approach to predict the life condition of crude oil pipeline, Processes, 8 (2020) 661, doi: 10.3390/pr8060661.

[34] M. Taky, F.Z. Addar, S. Qaid, H. Zeggar, H. El Hajji, Ultrafiltration of Moroccan Valencia orange juice: juice quality, optimization by custom designs and membrane fouling, Sustainability Agric. Food Environ. Res., 11 (2023) 0719–3726.

[35] R.G. Pontius Jr., O. Thontteh, H. Chen, Components of information for multiple resolution comparison between maps that share a real variable, Environ. Ecol. Stat., 15 (2008) 111–142.

[36] D. Bingöl, M. Hercan, S. Elevli, E. Kılıç, Comparison of the results of response surface methodology and artificial neural network for the biosorption of lead using black cumin, Bioresour. Technol., 112 (2012) 111–115.

[37] M. Mourabet, A. El Rhilassi, M. Bennani-Ziatni, A. Taitai, Comparative study of artificial neural network and response surface methodology for modelling and optimization the adsorption capacity of fluoride onto apatitic tricalcium phosphate, Univ. J. Appl. Math., 2 (2014) 84–91.