Membrane and Water Treatment, *Vol. 17, No. 1 (2026) 21-43*
https://doi.org/10.12989/mwt.2026.17.1.021

Techno-Press

A QSPR–ANN-driven predictive framework for assessing organic micropollutant rejection in forward osmosis systems

Fouad Kratbi*, Yamina Ammi, Salah Hanini

Laboratory of Biomaterials and Transport Phenomena (LBMPT), University of Medea, Algeria

(Received December 20, 2022, Revised September 7, 2025, Accepted December 30, 2025)

Abstract. Forward Osmosis (FO) is the subject of many current studies, given existing and future conditions around the world. This work is the continuation of the series of research that implicates Artificial Neural Networks in the processes of membrane separation. Three databases (with the same size of 193 points), two learning algorithms, two function transfers, five subdivisions of the database, and eleven (11) inputs were used with the aim to extract the optimal QSPR-NN model which is chosen based on the best values of coefficient of correlation (R) and the Root Mean Squared Error (RMSE). QSPR-NN (Quantitative Structure-Property Relationships - Neural Networks) model obtained was characterized by eleven (11) neurons on the input layer, fourteen (14) neurons in the hidden layer, and one (1) neuron in the output layer, the Bayesian regularization (Trainbr) was the learning algorithm, tangent sigmoid (Tansig), and purelin were the transfers functions for the hidden and output layers respectively. The performance of the QSPR-NN optimal model obtained was demonstrated with a higher value of (R = 0.9895) and low Root Mean Squared Error (RMSE = 4.3683%), and other errors as RER and RPD more than 2.5 and equal to 21.4356 and 3.4290 respectively, the (NSE) more than 0.9. Furthermore, the comparison with other work in the same orientation demonstrated the excellence of our model developed in this work compared to the others.

Keywords: artificial neural networks; forward osmosis; organic molecules; prediction; rejection

1. Introduction

The world knows recently a rapid development at all levels, human, economic, and industrial, resulting in a severe increase in the demand for water and energy, imposing us to preserve the existing conventional resources or find alternatives to them [1]. Pollution, current drought, desertification, exploitation of aquifers, and climatic fluctuations are the main causes of the scarcity of one of the most important elements of life and in sync reduced the level of the existing water resources [2]. A possible alternative source of water is wastewater via processes of membrane separation such as microfiltration (MF), ultrafiltration (UF), nanofiltration (NF), and reverse osmosis (RO) [2-4].

In the last decade, a new approach based on the osmotic pressure difference between two solutions separated by a membrane, called, Forward Osmosis (FO), this technique has been investigated in both academic research and industrial development as an alternative process for the conventional existing especially reverse osmosis. Forward osmosis is one of the technologies

*Corresponding author, Ph.D Student, E-mail: kratbi.fouad@univ-medea.dz

ISSN: 1225-4568 (Print), 1598-6217 (Online)

applied in a wide range of areas; water, energy, and life science with different applications seawater/brackish, water desalination, wastewater treatment, power generation, and food processing, FO can also be used for controlling drugs release and dehydration of pharmaceuticals [1, 5-6]. FO is characterized by many potential advantages such as less energy input, lower fouling tendency, easier fouling removal, and higher water recovery in comparison with other processes of the same nature (membranes separation) as reverse osmosis (RO), nanofiltration (NF), and ultrafiltration (UF) [7-11]. Membrane separation processes are known by the complexity of their transfer models due to the combinations between the proprieties of solute, solvent, and membrane characteristics, for this, many studies used other methods to evaluate the effect of inputs of the system on his target or output variables [12-13]. Machines learning have already been successful for modeling processes such as fermentation, drying, and other chemical processes, especially the membranes separations processes [14].

An Artificial Neural Network is one of these machines developed originally by Rosenblatt [14], this technique has a large successful use in modeling the membranes separation processes, many studies applied ANN with the aim to predict the impact of inputs on the outputs, optimization, and maintenance cost of membrane process such as microfiltration (MF), ultrafiltration (UF), nanofiltration (NF), and reverse osmosis (RO) [14]. Yangali-Quintanilla et al. (2009) have been using the artificial neurons network to predict the rejection of organic composites by nanofiltration and Reverse Osmosis membranes, in the same target of work [15]. Ammi et al. (2015) have created a QSPR-NN model whose objective is the prediction of the rejection of organic composites by nanofiltration and reverse osmosis membranes [16]. Abbassi et al. (2018) have been studied the oily wastewater using the ANN to predict the permeate flux [17].

To the best of our knowledge, a few studies investigated the application of ANN on the forward osmosis membranes. Pardeshi et al. (2016), used the ANN to determine the optimum conditions for the forward osmosis (FO) groundwater desalination, their work gave an honorable result and the ANN used can predict the optimum conditions for the FO system study [18]. In 2020 a study by Jasir Jawad et al was published, it was about the modeling of forward osmosis process using an artificial neurons network to predict the permeate flux, they studied the effect of nine inputs (membrane type, membrane orientation, feed solution concentration, draw solution concentration, draw solution molecule weight, feed solution velocity, draw solution velocity, feed and draw solutions temperature) on the permeate flux with different parameters of the ANN used (number of neurons, number of the hidden layers), their results obtained were very satisfying and demonstrated its ability to predict the relationships between inputs and outputs in a way better than another simple learning machine such as Multiple Linear Regressions (MLR) [19].

For this purpose, our work is the first study that consists to apply the ANN to develop a QSPR-NN model for predicting the rejection of organic molecules (OM) by the forward osmosis membranes using the properties of organic molecules, the membrane characteristics, and conditions operations as the inputs of the system. The database was collected from the literature; many models of predicting were developed with different inputs of the database, different training algorithms, different transfer functions, and different database subdivisions. To compare between these models developed, a method based on the best error for the testing phase and the best coefficient of correlation is applied. As soon as the optimum QSPR-NN model was obtained, with the aim of identifying the input that's the most impact on the output, sensitivity analysis using the weight method was studied to detect the most important input that affected the rejection of the organic molecules by the forward osmosis membranes, which motivated us to study this influence separately. For all the models of predicting, a range of applicability is mandatory, it was

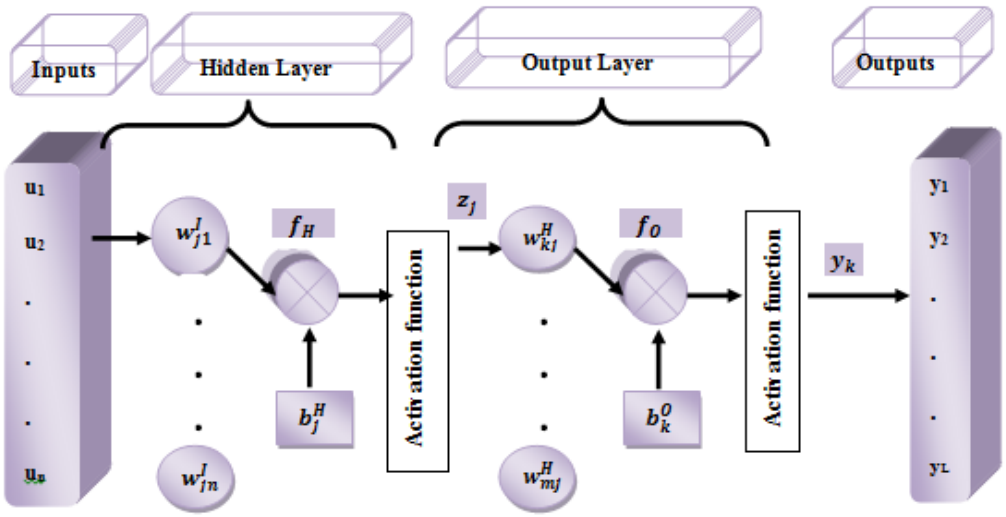


Figure 1. Architecture of the ANN [20]

determined by the method of William's plots during this work. Additional to the Root Mean Squared Error (RMSE) and coefficient correlation, twelve (12) other errors were used to evaluate the performance of the QSPR-NN model obtained with evaluation criteria. As mentioned previously, another work applied ANN for modeling of the forward osmosis processes. Therefore, a comparative study between our work and others mentioned above was studied in this work.

2. Experimental

2.1 Artificial neural networks

Artificial Neural Networks (ANN) is a machine learning technique used to model the input and output variables based on a learning mechanism effected by a set of data instead of using mathematical equations. ANN using generally for predicting the nonlinear system, the membrane separation is one of these complex systems, when the results obtained from the ANN model were close to the results of the conventional models [14].

The Artificial Neural Networks (ANN) is made with three stages of layers, called, Input, Hidden, and output layer respectively. Each stage or layer forms of neurons (nodes) where the number of these nodes is equal to the number of the inputs and outputs for the input and output layers respectively, for the Hidden layer the number of the optimum neurons was determined by the method of trial and error.

The real values of inputs present the food of the input layer and each succeeding layer receives weighted outputs from the previous layers as its inputs resulting, ANN assigns weights and biases to each neuron in the input and hidden layer as shown in the architecture of the ANN illustrated in Fig. 1 [16], [20-24].

In the learning phase, the values of the output of the neuron in the hidden layer and the output layer are given by the Eqs. (1, 2) respectively [16], [24].

$$z_j = f_h\left(\sum_{i=1}^{i=n} w_{JI}^{I} U_I + b_J^H\right);\ j=1, 2, \ldots m \quad (1)$$

z_j is the instance output of the Hidden layer, w_{JI}^{I} ; the weight of the Input- Hidden layer connection,U_I; is the input, b_J^H is the Bias of the Hidden layer, I present the number of inputs (varied entre 1 et n), J is the number of the neurons in the Hidden layer (varied from 1 to m (the optimum), and $f_h$is the transfer function of the hidden layer.

$$Y_k = f_o\left(\sum_{j=1}^{j=m} w_{kj}^{h} z_j + b_k^O\right) \quad k = 1, 2, \ldots L \quad (2)$$

Y_k: present the output, w_{kj}^{h}, are the weight of the Hidden-Output connections, z_j, is the instance output of hidden layer, b_k^O, is the Bias of output layer, k the number of the output layer, and $f_o$is the output layer's transfer function.

The most frequent activation (transfer) functions are the hyperbolic tangent sigmoid transfer function (tansig) and Logarithmic sigmoid transfer function (logsig) for the hidden layer (are given by Eqs. (3, 4), and Pure linear (purelin) for the output layer (Eq. (5)) [14], [24-25].

$$f(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}}, \quad (3)$$

$$f(a) = \frac{1}{1 + e^{-a}}, \quad (4)$$

$$f(a) = a. \quad (5)$$

2.1 Modeling procedure

The steps used in this work, for design and optimization of the architecture of the neural networks are illustrated in Fig. 2.

2.2.1 Database, Collection, Division, and Reduction

We have collected a database from the available literature [26-42] intending to group all the characteristics of the studied system, the size of this database was 193 points of 53 organic molecules (forty-four (44) pharmaceuticals, seven (7) perfluorochemicals, and two (2) organic acids). Fig. 3 shows the rejection of these organic molecules as a function of their molecular weights when the values of the molecular weights varied between 100 and 700 (g/mol) for the X-axis (Molecular weight) and from 5% to 100% for the experimental rejections.

The selection of the input and output variables was based on interactions between organic molecule properties, membrane characteristics, and filtration operating conditions for the rejection of organic molecules by forward osmosis "FO" membranes. The inputs considered in this work are molecule descriptors (the effective diameter of an organic molecule in water " d_c", molecular length "Length", molecular width "Width", molecular depth "Depth", molecular equivalent width "eqwidth", minimal surface "S_{min}", maximal surface "S_{max}", molecular volume, Dipole moment, The logarithm of the octanol-water partition coefficient "Log P", membrane characteristics (Surface membrane charge as "Zeta potential", and the Hydrophobicity "as Contact angle"), and operating conditions (pH, Cross flow velocity (CFV), and the water flux).

The software used to compute the inputs values were mentioned in Table 1.

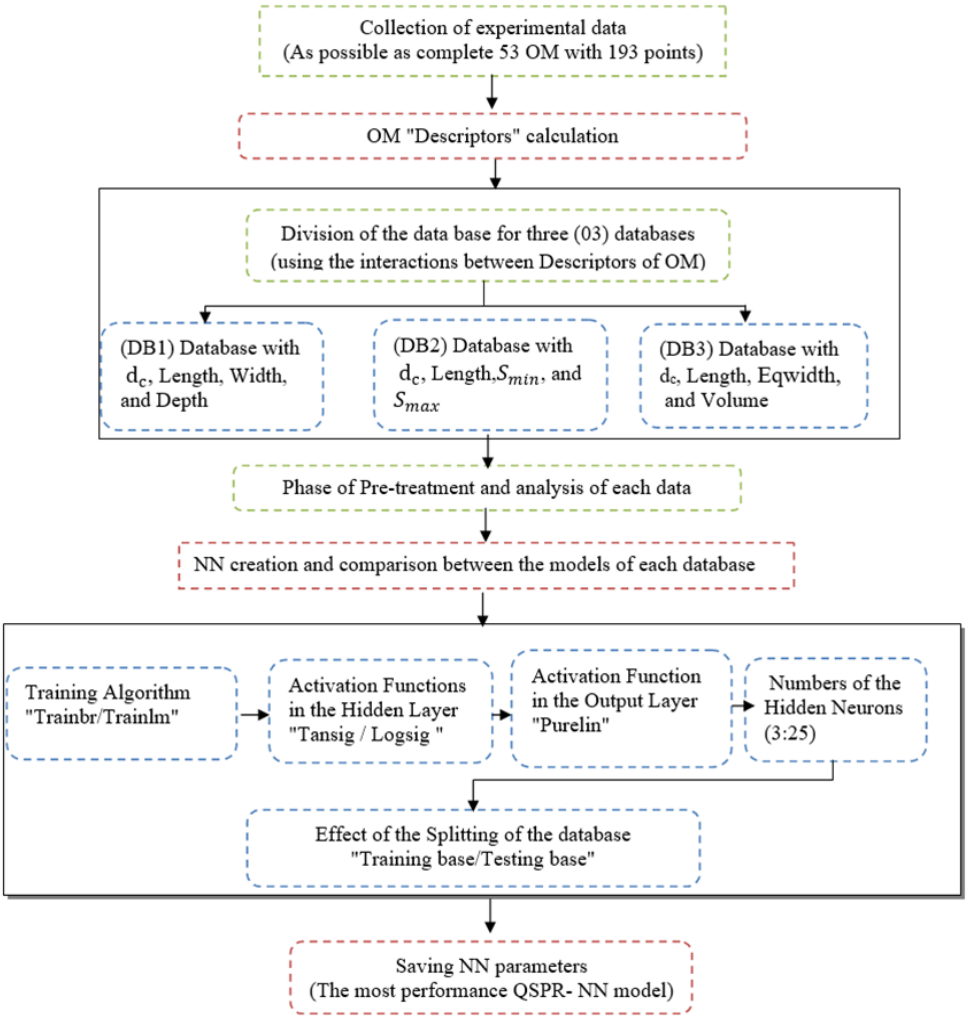


Figure 2. Steps used for design and optimization of neural networks

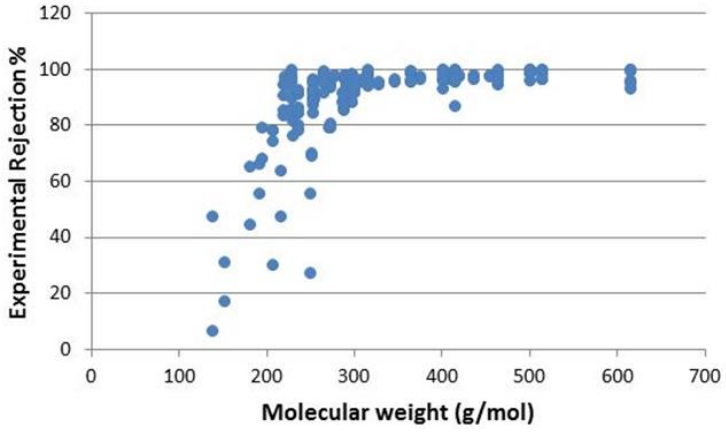


Figure 3. Rejection as a function of the molecular weight

Table 1. The software used to compute the input values

Input values	Software
Log P, Dipole moment, S_{min} , $S_{max,}$ and molecular volume	Calculated with Hyperchem 2008
Molecular length	Calculated with Chembio 2014

Table 2. Inputs and output for each database

	DB1	DB2	DB3
Inputs	d_c	d_c	d_c
	Dipole moment	Dipole moment	Dipole moment
	Log P	Log P	Log P
	Length	Length	Length
	Width	Eq Width	S_{min}
	Depth	Volume	S_{max}
	Contact angle	Contact angle	Contact angle
	Zeta potential	Zeta potential	Zeta potential
	pH	pH	pH
	Crossflow Velocity	Crossflow Velocity	Crossflow Velocity
	Water flux	Water flux	Water flux
Output	Rejection	Rejection	Rejection

The values of the effective diameter of an organic molecule in water " d_c", molecular width "Width", molecular depth "Depth", and Equivalent molecular width "Eqwidth" are calculated with the Eqs. (6- 9), respectively, [43- 45].

$$d_c = 0.065 * M_w^{0.0438} \tag{6}$$

Where: M_w = Molecule Weight.

$$Width = \frac{1}{2}\sqrt{S_{min}}, \tag{7}$$

$$Depth = \frac{1}{2}\sqrt{S_{max}}, \tag{8}$$

$$Eq\ width = \sqrt{molecular\ width * molecular\ depth}. \tag{9}$$

To achieve the objective of finding the parameters (which have a relationship between them as seen in the preceding Eqs. (7-9), that give the highest performance of the "QSPR-NN" obtained. The approach of this work is to divide the original database into three databases according to the relations between the inputs and to edit all the properties of the organic molecules (the inputs and outputs for each database are illustrated in Table 2 while respecting the process cited in Fig. 2. Statistical analysis is preliminary (standard deviations (STD), minimum, maximum, and mean) are shown in Table 3.

For each database, we used a matrix correlation to edit the interactions between the variables (organic molecule properties, membrane characteristics, and operating conditions) and to reduce its size which can be accessed as Supplementary data (Table 1-3).

Table 3. Statistical analyses of the Input and Output variables of all databases

	Min	Max	STD	Mean
d_c	0.0914	1.1000	0.1291	0.7917
Log P	0.1612	6.0000	1.2105	2.4347
Dipole moment (Debey)	-2.2200	9.3000	2.9023	2.3499
Length (nm)	0.1468	0.2000	0.0119	0.1579
Width (nm)	0.5000	1.3000	0.1116	0.9969
Depth (nm)	0.3795	1.4000	0.1126	1.0509
EqWidth(nm)	0.6529	1.3493	0.1019	1.0216
Volume (nm^3)	0.4093	1.3472	0.1583	0.7409
S_{min} (nm^2)	2.3329	7.1000	0.8288	4.0522
S_{max} (nm^2)	0.5761	7.4000	0.8704	4.4680
Contact angle (°)	42.7	90.3000	6.2294	64.6435
Zeta potential (mv)	-39.91	19.7000	11.9568	-4.3641
pH	3	9	1.5434	6.057
Crossflow Velocity(m/h)	288	1094.4000	388.0198	634.5699
Water flux (L/m^2h)	5.2579	17.6000	4.3444	10.0116
Rejection (%)	6.3636	100	14.9789	90.0579

2.2.2 Development of the QSPR-NN model

For each database, several networks are developed, at first, the training algorithm was fixed and the activations functions in the hidden layer were variants, for this purpose, the Bayesian Regularization (trainbr) and Levenbarg Marquardt (trainlm) were used as training algorithms, Hyperbolic tangent sigmoid (tansig) and Logarithmic sigmoid (logsig) as activation functions for the hidden layer, the Pure linear (purelin) transfer function was used in the output layer. The initial division was 80%, 20% for the training phase and test phase respectively with trainbr as the training algorithm and 80%, 10%, and 10% for training, validation, and testing phases respectively, using the training algorithm trainlm. The number of neurons in the hidden layer was between 3 and 25 neurons for each neural network model. At the end of this loop, the effect of the database splitting on the best configuration (database, training algorithm, activation function) was studied.

Software MATLAB 2020 was used for the ANN modeling of the rejection of the organic molecules by the forward osmosis "FO" membranes by trial-and-error method.

For this and for each database, many subdivisions have been used with two (02) algorithms (Bayesian Regularization (Trainbr) and Levenbarg-Marquardt (Trainlm)), two (02) transfer functions for the hidden layer (tangent sigmoid and log sigmoid) and purelin as the transfer function for the output layer have been formed to create the structure of each ANN model.

2.3 Sensitivity analysis

To edit the influence of each input variable (organic molecule properties, membrane characteristics, and filtration operating conditions) on the output variable (Rejection of the organic molecules by FO membranes), a sensitivity analysis is recommended, for this, a method was developed by Garson [46-47] called "weight method", this method is based on the calculation the

relative importance (RI) of each input on the output variable. The process of calculating the relative importance by weight method is summarized in the flow steps [48-49].

$$\text{Step1: } P_{i\,j} = \left|w^i\, w^j\right|, \tag{10}$$

$$\text{Step 2: } Q_{i\,j} = \frac{P_{i\,j}}{\sum_{i=1}^{n_i} P_{i\,j}}, \tag{11}$$

$$\text{Step 3: } S_i = \sum_{i=1}^{n} Q_{ij}, \tag{12}$$

$$\text{Step 4: } RI_I\,(\%) = 100 * \frac{S_i}{\sum_{i=1}^{k_i} S_i}. \tag{13}$$

w^i, w^j are the connections weights for the input-hidden layer and hidden-output layers respectively.

2.4 Applicability domain

A precision range is recommended for any prediction model satisfying its application, this one is defined by the applicability domain, or external of this domain, the use of this model can offer an erroneous result, many approaches can be used to determine this domain, in this work, we based on the leverage approach (William's plot) in order to determinate this important domain. The impact of a section on the model is presented by the leverage (h_i). The compound's leverage in the original variable space is given by the following Eq. (14):

$$\mathrm{H} = \mathrm{X}(\mathrm{X^tX})^{-1}\mathrm{X^t} \tag{14}$$

Where X refers to the m × n matrix, (m and n represent the number of samples and the parameters (input variables) of the model, respectively). The leverage values are obtained from the main diagonal of the H matrix and are always between 0 and 1. The critical leverage(h^*) is calculated generally with the formula presented underside:

$$\mathrm{h^*} = \frac{3(\mathrm{n}+1)}{\mathrm{m}} = \frac{3(11+1)}{193} = 0.1865 \tag{15}$$

The normalized residuals are calculated from the experimental data of the rejection and that calculated by the model.

$$(\mathrm{R\_Norm})_\mathrm{i} = \frac{\left(\mathrm{Rejection_i^{exp}} - \mathrm{Rejection_i^{cal}}\right)}{\sqrt{\mathrm{Var(Rejection^{exp} - Rejection^{cal})}}}\ \mathrm{i} = 1, \ldots \mathrm{m} \tag{16}$$

Normalized residuals equal to ±3 is considered margin limits for validated data, whereas residuals outside these ranges define suspected data. In addition, all data points with Hat values > H* are also considered suspected data [49-50].

3. Results and discussions

3.1 Creation and optimization of the QSPR-NN Model

In order to search for the ideal architecture (as developed in Fig. 2), for each database, a neural network was created with different training algorithms, different activation functions, and different

Table 4. Results of neural network models for database 1

QSPR-NN number	Training Algorithms	Activate Functions	Phase	RMSE	R
$QSPR\text{-}NN_1$	Bayesian Regularization (Trainbr)	Tansig	Test	8.5517	0.8848
			Train	2.3224	0.9895
			All	4.3683	0.9601
$QSPR\text{-}NN_2$		Logsig	Test	7.5823	0.6926
			Train	2.3756	0.9892
			All	4.0151	0.9639
$QSPR\text{-}NN_3$	Levenbarg-Marquardt (Trainlm)	Tansig	Test	9.7899	-0.1144
			Train	3.2961	0.9821
			Val	3.3324	0.8455
			All	4.3879	0.9574
$QSPR\text{-}NN_4$		Logsig	Test	9.0571	0.7887
			Train	3.2641	0.9789
			Val	5.1623	0.3913
			All	4.3882	0.9561

splitting of the database, the number of the neurons in the hidden layer varied between 3 and 25 neurons. to evaluate the performance of each neural network, two parameters were used the Root Mean Squared Error (RMSE) is represented the error between the calculated and the experimental values of the rejection of organic molecules which is given by Eq. (17) [43], and the correlation coefficient (R) (was obtained with the plots and parameters of the linear regression by the function Matlab Postreg).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(\boldsymbol{Y}_{i,exp} - \boldsymbol{Y}_{i,cal}\right)^2}{\boldsymbol{n}}} \tag{17}$$

Where n is the number of data points, $Y_{i,cal}$ represents the calculated values from the neural network, and $Y_{i,exp}$ is the experimental values.

The next Tables 4-6 show the results of neural network based on quantitative structure proprieties relationships models "QSPR-NN" obtained for database 1, database 2, and database 3 respectively.

Table 4 shows that the result of the first neural network model "$QSPR\text{-}NN_1$" gives the higher correlation coefficients (R = 0.8848 for the testing phase and R = 0.9601 for all phases) and the Root Mean Squared Error (RMSE equal to 8.5517% and 4.3683% for the testing and all phase respectively) with training algorithm "trainbr" and activation function "tansig" than the others neural network models ($QSPR\text{-}NN_2$, $QSPR\text{-}NN_3$, and $QSPR\text{-}NN_4$). The results demonstrate that the training algorithm "trainbr" and activation function "tansig" are more suitable to describe the rejection of organic molecules by forward osmosis for the first database. Table 5 represents that the result of the sixth neural network model ($QSPR\text{-}NN_6$) offers the higher correlation coefficients (R = 0.8778 for the testing phase and R = 0.9552 for all phases) and the Root Mean Squared Error (RMSE equal to 8.0873% and 4.4929% for the testing and all phases respectively) with training

Table 5. Results of neural network models for database 2

QSPR-NN number	Training Algorithms	Activate Functions	Phase	RMSE	R
$QSPR\text{-}NN_5$	Bayesian Regularization (Trainbr)	Tansig	Test	8.9324	0.7838
			Train	1.0866	0.9581
			All	4.3238	0.9981
$QSPR\text{-}NN_6$		Logsig	Test	8.0873	0.8778
			Train	2.9556	0.9804
			All	4.4929	0.9552
$QSPR\text{-}NN_7$	Levenbarg-Marquardt (Trainlm)	Tansig	Test	4.7210	0.3978
			Train	2.9109	0.9845
			Val	7.8253	0.4405
			All	3.8765	0.9664
$QSPR\text{-}NN_8$		Logsig	Test	8.1503	0.7401
			Train	5.1784	0.9463
			Val	6.3081	0.3581
			All	5.6562	0.9260

Table 6. Results of neural network models for database 3

QSPR-NN number	Training Algorithms	Activate Functions	Phase	RMSE	R
$QSPR\text{-}NN_9$	Bayesian Regularization (Trainbr)	Tansig	Test	9.0857	0.8056
			Train	2.7247	0.9828
			All	4.9881	0.9435
$QSPR\text{-}NN_{10}$		Logsig	Test	8.5970	0.7979
			Train	1.8572	0.9934
			All	4.2056	0.9623
$QSPR\text{-}NN_{11}$	Levenbarg-Marquardt (Trainlm)	Tansig	Test	10.2653	0.8053
			Train	3.5504	0.9751
			Val	8.7500	0.3047
			All	5.2947	0.9387
$QSPR\text{-}NN_{12}$		Logsig	Test	10.3427	0.8169
			Train	4.5325	0.9559
			Val	6.4999	0.7971
			All	5.5847	0.9282

algorithm “trainbr” and activation function “logsig” than the others neural network models ($QSPR\text{-}NN_5$, $QSPR\text{-}NN_7$, and $QSPR\text{-}NN_8$). The results prove that the training algorithm “trainbr” and activation function “logsig” are more appropriate to describe the rejection of organic molecules by forward osmosis for the second database. In Table 6, the result of the higher

Table 7. Effect of splitting of database

Splitting of database	Size of the sub division	Phase	RMSE	R
80% Training 20% Testing	38 points	Test	8.5517	0.8848
	155 points	Train	2.3224	0.9895
	193 points	All	4.3683	0.9601
85% Training 15% Testing	29 points	Test	8.9808	0.8507
	164 points	Train	2.2700	0.9883
	193 points	All	4.0617	0.9629
70% Training 30% Testing	58 points	Test	9.8077	0.8384
	135 points	Train	2.5858	0.9834
	193 points	All	5.7952	0.9258
60% Training 40% Testing	77 points	Test	9.7319	0.7402
	116 points	Train	3.132	0.9803
	193 points	All	6.6087	0.9008
55% Training 45% Testing	87 points	Test	10.4090	0.7356
	106 points	Train	3.9617	0.9640
	193 points	All	7.5803	0.8632

Table 8. Structure of the optimized neural network model (QSPR-NNoptimal)

Database	Splitting of database	Training Algorithm	Input layer	Hidden Layer		Output layer	
			Neurons Number	Neurons Number	Activation Function	Neurons Number	Activation Function
DB1	80% training phase 20% testing phase	Bayesian Regularization (Trainbr)	11	14	Tangent Sigmoid (Tansig)	1	Purelin

correlation coefficients (R = 0.8196 for the testing phase and R = 0.9282 for all phases) and the Root Mean Squared Error (RMSE equal to 10.3427% and 5.5847% for the testing and all phases respectively) for the twelfth neural network model (QSPR-NN_{12}) with training algorithm "trainlm" and activation function "logsig" than the others neural network models (QSPR-NN_9, QSPR-NN_{10}, and QSPR-NN_{11}). The results show that the training algorithm "trainlm" and activation function "logsig" are more appropriate to describe the rejection of organic molecules by forward osmosis for the third database.

According to the previous results, the first database, training algorithm "trainbr", and activation function "tasing" were the optimal structure for the modeling of the rejection of the organic molecules by the forward osmosis "FO".

This optimal combination (database1, trainbr, and tansig) was used to study the effect of the splitting of the database, as illustrated in Table 7. This table shows that the division1 (80% for the training phase and 20% for the testing phase) for the neural network model (QSPR-NN_1) gives the higher correlation coefficients (R = 0.8848 for the testing phase and R = 0.9601 for all phases) and the Root Mean Squared Error (RMSE equal to 8.5517% and 4.3683% for the testing and all phase respectively) than the other divisions for the same QSPR-NN_1.

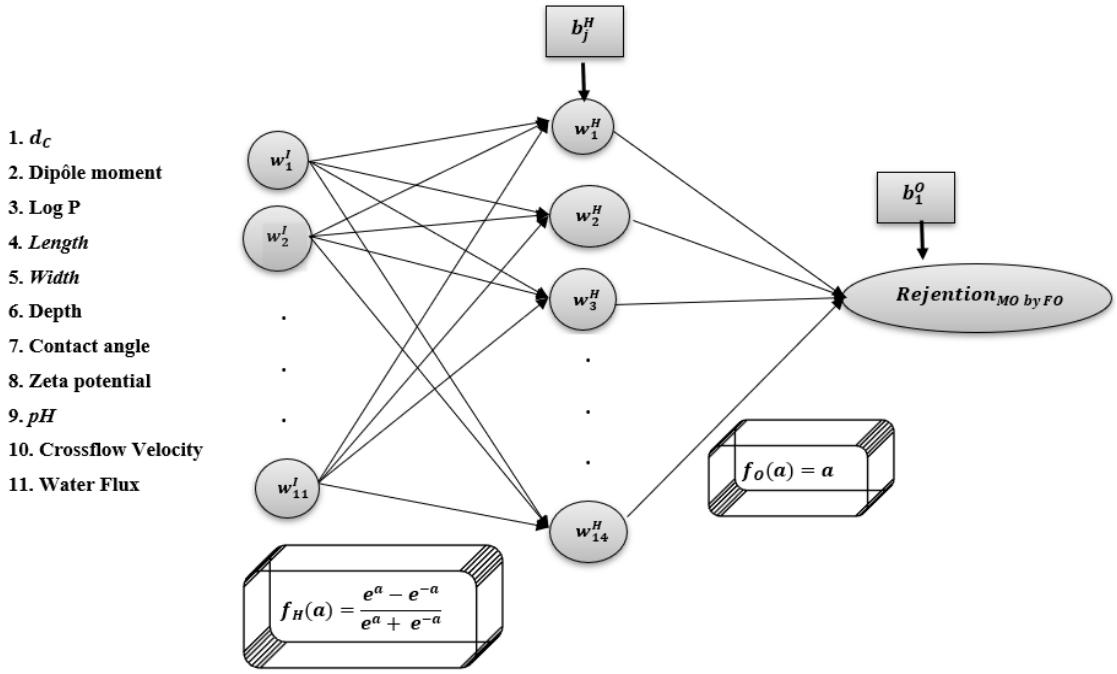


Figure 4. The architecture of the QSPR-NN$_{optimal}$ obtained

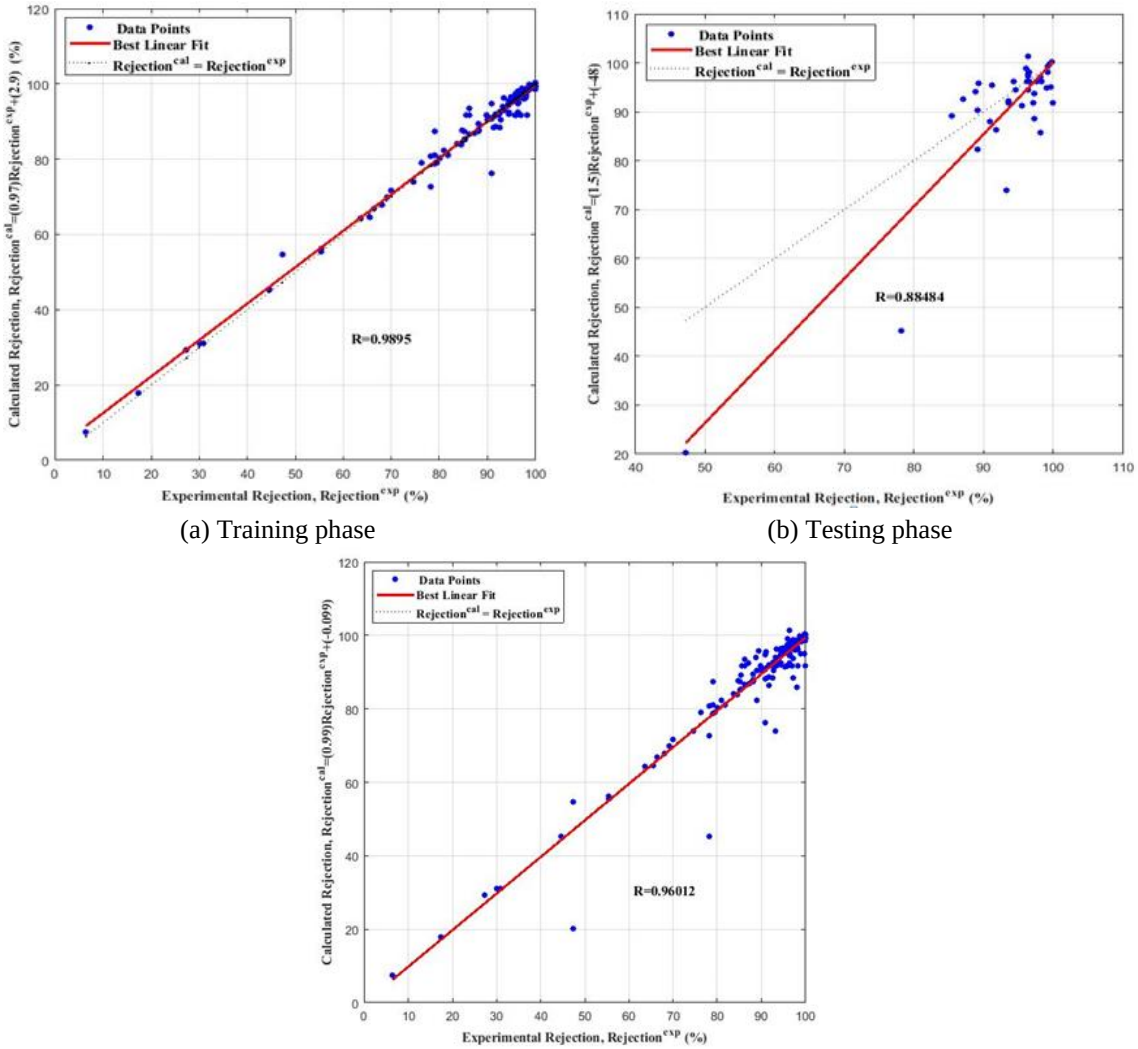
Table 9. Linear regression vectors $y^{cal} = \alpha * y^{exp} + \beta$ with α slope, β intercept, R coefficient correlation, and RMSE (Root Means Squared Error)

	Phase	α	β	R	RMSE
QSPR-NN$_{optimal}$	Training	0.9674	2.9117	0.9895	2.3224
	Testing	1.4775	-47.5579	0.8848	8.5517
	Total	0.9945	-0.0995	0.9601	4.3683

Following the above results, the parameter of the optimized QSPR-NN model with the database and its best division, the training algorithm, the neuron number for all layers (input, hidden, and output) of the QSPR-NN model, and finally the activation functions were represented in Table 8.

The architecture of the QSPR-NN$_{optimal}$ obtained is illustrated in Fig. 4. The Fig. 5 shows the agreement plots for the rejection of organic molecules with agreement vectors of the neural network based on quantitative structure proprieties relationships "QSPR-NN$_{optimal}$" model for the training, testing, and total phase, the values of the parameters of the linear regression vectors for all phases with the QSPR-NN$_{optimal}$ model obtained [α, β, R] were mentioned in Table 9 with α represent the slope of the vector, β is the intercept, and R correlation coefficient gives the difference by the contribution of the ideal.

Fig. 5 and Table 9 show a comparison between experimental and calculated rejections of organic molecules by forward osmosis (training, testing, and total phases) for the QSPR-NN$_{optimal}$ model with agreement vectors approaching the ideal in the adjustment of the profiles of the neural networks based on quantitative structure proprieties relationships, [α, β, R] = [0.9674, 2.9117, 0.9895] for training phase; [α, β, R] = [1.4775, -47.5579, 0.8848] for the testing phase, and [α, β,



(a) Training phase (b) Testing phase

(c) Total phase

Figure 5. Comparison of the experimental and calculated values with the QSPR-$NN_{optimal}$ obtained

R] = [0.9945, -0.0995, 0.9601] for total phase. Correlation coefficients are generally considered to be excellent ($0.90 \leq R \leq 1.00$) for this QSPR-$NN_{optimal}$, which shows good robustness of this model and the possibility of predicting the different parameters that characterize the rejection of organic molecules during forward osmosis "FO".

For the QSPR-$NN_{optimal}$ obtained**,** a mathematical formula for predicting the rejection of the organic molecules "OM" was developed using the equation of the training algorithm Trainbr and the equation of the activation function Tansig for the hidden layer and Purelin as a transfer function for the output layer. The instance output h_j (output of the Hidden layer) is presented by:

$$h_j = f_h\left[\sum_{i=1}^{11} w_{ji}^{I} E_i + b_j^{H}\right] = \frac{\exp\left(\sum_{i=1}^{11} w_{ji}^{H} E_i + b_j^{H}\right) - \exp\left(\sum_{i=1}^{11} w_{ji}^{H} E_i + b_j^{H}\right)}{\exp\left(\sum_{i=1}^{11} w_{ji}^{H} E_i + b_j^{H}\right) + \exp\left(\sum_{i=1}^{11} w_{ji}^{H} E_i + b_j^{H}\right)} \quad (18)$$

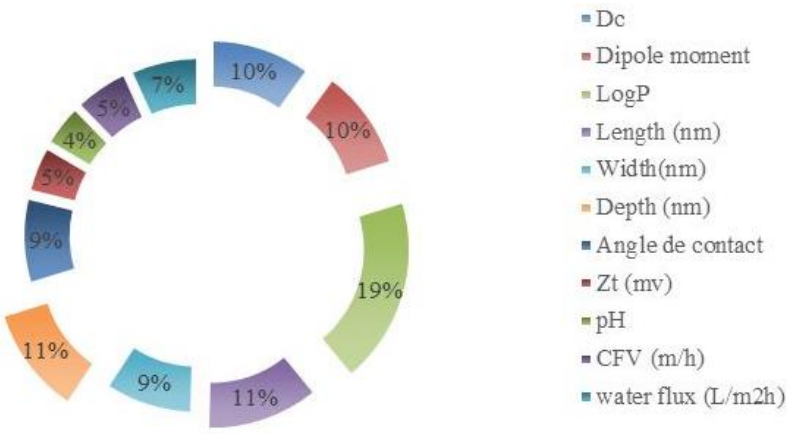


Figure. 6 The values of Relative Importance (RI) for each input

The output (Rejection) is given by the follow Eqs. (19, 20).

$$Rejection_{OM\ by\ FO} = f_o\left[\sum_{j=1}^{14} w_{1j}^{o} h_j + b_1^{o}\right] \tag{19}$$

$$Rejection_{OM\ by\ FO} = \sum_{j=1}^{14} w_{1j}^{H} \frac{\exp\left(\sum_{i=1}^{11} w_{ji}^{H} E_i + b_j^{H}\right) - \exp\left(\sum_{i=1}^{11} w_{ji}^{H} E_i + b_j^{H}\right)}{\exp\left(\sum_{i=1}^{11} w_{ji}^{H} E_i + b_j^{H}\right) + \exp\left(\sum_{i=1}^{11} w_{ji}^{H} E_i + b_j^{H}\right)} + b_1^{o} \tag{20}$$

Where E_i represented the input values, w_{ji}^{H}, w_{1j}^{o} are the connections weights for the input-hidden layer and hidden-output layer respectively, b_j^{H}, b_1^{o} are the bias of the hidden and output layer respectively.

3.2 Sensitivity analysis

For the QSPR-$NN_{optimal}$ model obtained, the table of the weight for the connections between Input-Hidden Layers and Hidden-output layers was available and can be accessed in the Supplementary data (Table 4). Fig. 6 shows the contribution of the input variables to the rejection of the organic molecules by the forward osmosis membranes.

This Fig. 6 proves that log p, d_c, dipole moment, length, and Depth are the most influenced on the rejection of the organic molecules by FO membranes with values of the relative importance of 19%, 11%, 11%, 10%; and 10% respectively. The method of the weights demonstrates the right choice of the input variables for modeling where all the values of relative importance are more than 4%.

3.3 Applicability domain

The results of the outlier analysis are mentioned in Fig. 7 based on the above-mentioned figure, the critical lever H* value is 0.1865, and only the data shown by red circles is the valid data and for which the model's accuracy is confirmed. The analysis of the outliers of the QSPR-$NN_{optimal}$

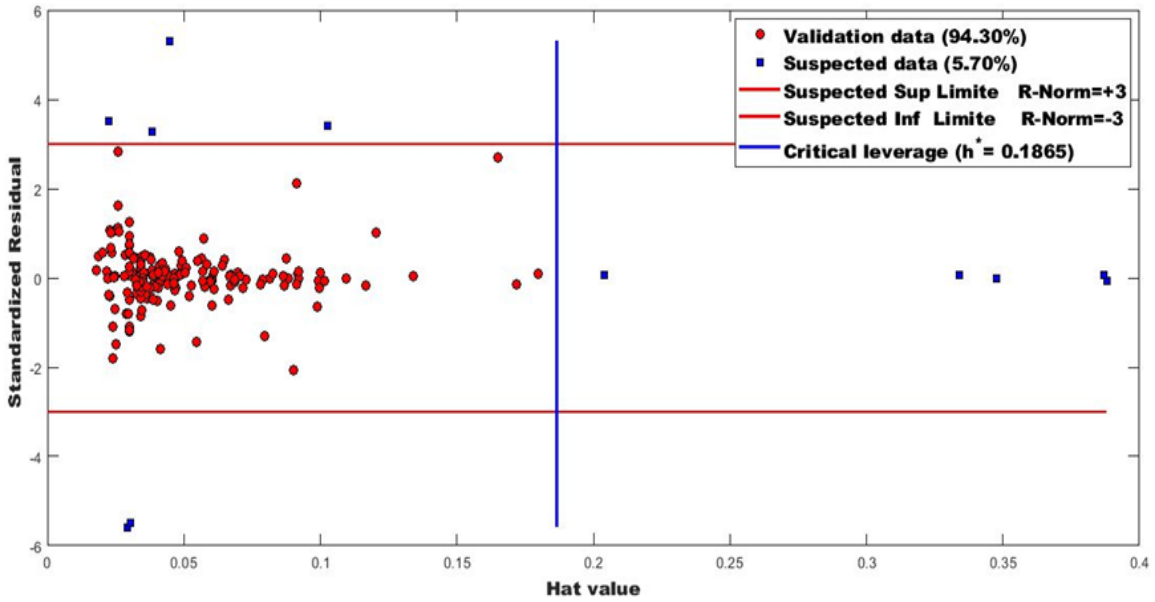


Figure 7. William's Diagram for detection of outliers for the QSPR-$NN_{optimal}$ model for total phase

model "trainbr, tansig" revealed the results represented on the William diagram where it is observed that 182/193 (94.30%) points belong to the domain of validity and that only 11/193 (5.70%) points are outside the domain of applicability of the optimized QSPR-NN model.

3.4 Evaluation criteria

To evaluate the quality of the prediction by the QSPR-$NN_{optimal}$ model obtained, twelve (12) other than the errors mentioned above "the Correlation Coefficient (R) and the Root Mean Squared Error (RMSE)", these errors are summarized at the Mean Absolute Error (MAE), the Model Predictive Error (MPE), the Standard Error of Prediction (SEP), Range Error Ratio (RER), Residual Predictive Deviation (RPD), the Mean Square Error (MSE), the Mean Relative Squared Error (MRSE), the Relative Absolute Error (RAE), the Accuracy factor (A_f), Bias factor (B_f), Relative Absolute Error (RAE), and Nash-Sutcliffe Efficiency (NSE), which are given by the next Eqs. (21-33), [48, 51, 52].

$$\mathrm{MAE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}\left|(\mathrm{y_{i,exp}} - \mathrm{y_{i,cal}})\right|, \tag{21}$$

$$\mathrm{MPE(\%)} = \frac{100}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{n}}\left|\frac{(\mathrm{y_{i,exp}} - \mathrm{y_{i,cal}}}{\mathrm{y_{i,exp}}}\right|, \tag{22}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{\mathrm{i=1}}^{\mathrm{N}}(\mathrm{Y_{i,cal}} - \mathrm{Y_{i,exp}})^2}{\mathrm{N}}}, \tag{23}$$

$$\mathrm{SEP}(\%) = \frac{\mathrm{RMSE}}{\mathrm{Y_e}} \times 100, \tag{24}$$

$$\mathrm{RER} = \frac{\mathrm{Max} - \mathrm{Min}}{\mathrm{RMSE}}, \tag{25}$$

$$\mathrm{RPD} = \frac{\mathrm{SD}}{\mathrm{RMSE}}, \tag{26}$$

$$\mathrm{MSE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}\left(\mathrm{y_{i,exp}} - \mathrm{y_{i,cal}}\right)^2, \tag{27}$$

$$\mathrm{MRSE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=1}}^{\mathrm{N}}\left(\frac{\mathrm{y_{i,exp}} - \mathrm{y_{i,cal}}}{\mathrm{y_{i,exp}}}\right)^2, \tag{28}$$

$$\mathrm{REA} = \sum_{\mathrm{i=1}}^{\mathrm{N}}\left|\frac{\mathrm{y_{i,exp}} - \mathrm{y_{i,cal}}}{\mathrm{y_{i,exp}}}\right|, \tag{29}$$

$$\mathrm{A_f} = 10^{\left(\sum_{\mathrm{i=1}}^{\mathrm{N}}\left|\log\frac{\mathrm{y_{i,cal}}}{\mathrm{y_{i,exp}}}\right|/\mathrm{N}\right)}, \tag{30}$$

$$\mathrm{B_f} = 10^{\left(\sum_{\mathrm{i=1}}^{\mathrm{N}}\log\frac{\mathrm{y_{i,cal}}}{\mathrm{y_{i,exp}}}/\mathrm{N}\right)}, \tag{31}$$

$$\mathrm{RAE} = \frac{\sum_{\mathrm{i=1}}^{\mathrm{N}}\left|\mathrm{y_{i,cal}} - \mathrm{y_{i,exp}}\right|}{\sum_{\mathrm{i=1}}^{\mathrm{N}}\left|\overline{\mathrm{y_{i,exp}}} - \mathrm{y_{i,exp}}\right|}, \tag{32}$$

$$\mathrm{NSE} = \left|1 - \left[\frac{\sum_{\mathrm{i=1}}^{\mathrm{N}}(\mathrm{Y_{i,exp}} - \mathrm{Y_{i,cal}})^2}{\sum_{\mathrm{i=1}}^{\mathrm{N}}(\mathrm{Y_{i,exp}} - \overline{\mathrm{Y_{i,exp}}})^2}\right]\right|, -\infty \leq \mathrm{NSE} \leq 1. \tag{33}$$

$Y_{i,exp}$, $Y_{i,cal}$ represent the experimental and the calculated values respectively, n is the number of data, Y_e is the mean value of the experimental data, and STD, Max, and Min are the standard deviation, the maximum, and the minimum of the experimental data respectively.

Viscarra Rossel (2006) provided a five-level of values of Residual Predictive Deviation (*RPD*) and Range Error Ratio (RER) to interpret the model of the predicting as described below: RER and RPD higher than 2.5 means excellent predictions; good predictions for values of RER and RPD between 2.0 and 2.5; RER and RPD less than 2.0 and more than 1.8 signifies approximate quantitative predictions; possibility to distinguish high and low values (RER and RPD of 1.4 to 1.8); and unsuccessful for values less than 1.40. Fig. 8 shows the evaluation Criteria of the QSPR-$NN_{optimal}$ model [53].

The error values confirm that the model QSPR-$NN_{optimal}$ gives an excellent prediction of the rejection of organic molecules with a value of RER equal to 21.4356 and RPD more than 2.5 with a value of 3.4290. Nash-Sutcliffe Efficiency (NSE) is used also by D. N. Moriasi [54], Melis Zeybek [55], and Dileep Kumar Gupta [56] to evaluate the performance of the model of predicting, the NSE obtained in this study is more than 0.9 which convince again the excellent

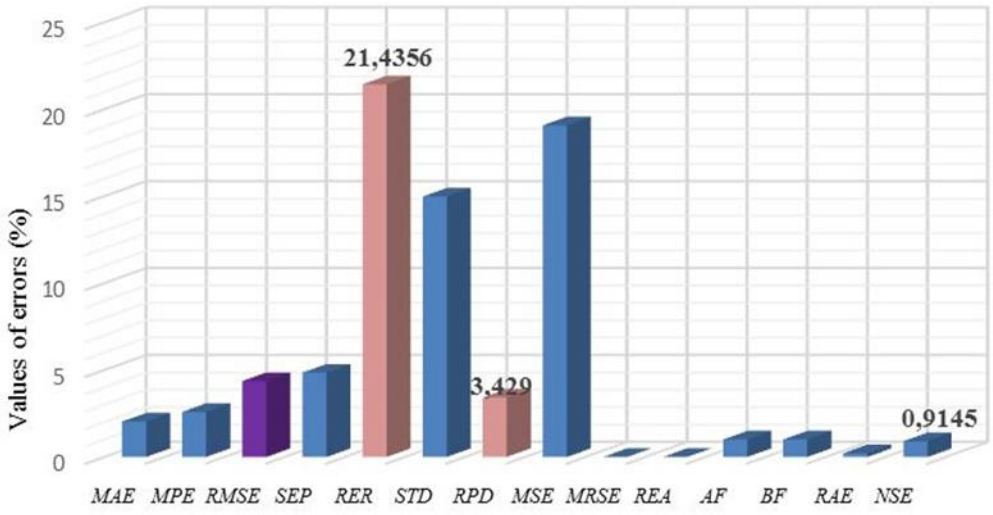


Figure 8. Evaluation Criteria of QSPR-$NN_{optimal}$ model

performance of the QSPR-$NN_{optimal}$ model of the rejection of the organic molecules by the forward osmosis membranes obtained as agreeing with the previous discussion about the levels of interpreting the model of the predicting, the MAE, MPE, SEP, and RMSE have values less than 4%, which mean the good precision of the model obtained, the other errors MRSE, REA, and RAE give poor values of error (close to zero (0)) which explain the accuracy of this model obtained. According to the previous discussion, the highest performance and precision of our QSPR-$NN_{optimal}$ obtained of the rejection of organic molecules by the forward osmosis membranes have demonstrated with the previous values of errors.

3.5 The rejection as a function of log p

The sensitivity analysis cited above showed that Log p is the property most affecting the rejection of organic molecules by forward osmosis membranes, in this part, a study of the rejection as a function of log p is performed. Fig. 9 demonstrates the relation between the rejection of organic molecules and log p. The trajectories of the experimental (with blue color) and calculated (with red color) rejection are superimposed one on the other which shows the performance of the model obtained, even though the linear of these values are identical, this figure also shows that the evolution of the rejection is proportional with the log p-value [57-59].

The logarithm of repartition (log p) represents the ratio between the solubility of organic molecules in an organic solution and water; it characterizes the hydrophilic and hydrophobic functions of organic molecules. A large value of log p means that this molecule is hydrophobic and insoluble in water and aqueous solutions, and a small value of log p means that the molecule is hydrophilic and soluble in water and aqueous solutions. This characteristic is well demonstrated in this study, and with the increase of the values of log p, the rejection of the organic molecules by the forward osmosis membranes is stable and increases which signifies a poor solubility in water and consequently their rejection by the membranes which means also that the hydrophobicity force is dominant and gives the stability of the rejection of the organic molecules, in the other face, and

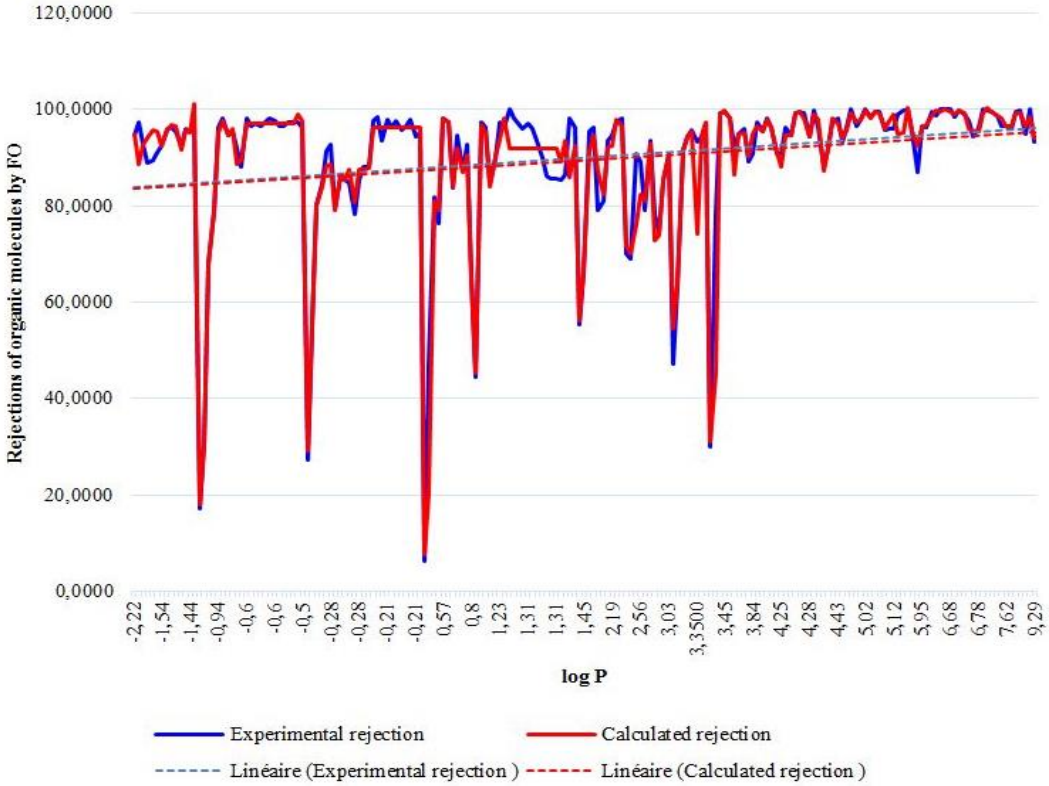


Figure 9. The rejection as a function of log p

Table 10. Overview of various works on model's neural networks for modeling the performance of forward osmosis

Name of model	Target of neural networks	Database size	Number and nature of compounds	Performance of ANN				
				R				RMSE (%)
				Training	Test	Val	all	
Jasir Jawad In 2020 [19]	Permeate flux	709	Salt compounds	0.97358	0.82097	0.85357	0.93097	-
Our work in 2022	Rejection of organic molecules	193	53 Organic molecules	0.9895	0.8848	-	0.9601	4.3683

for small values of log p the rejection is lower than them of the first case and the steric effect of the organic molecules gives the values of the rejection by the membranes. In the range between these two values (small and large) the instability of the rejection is observed and can be interpreted by the range of the existence of other forces as the steric effect [57-60].

3.6 Comparison and analysis

The modeling of the performance of the forward osmosis using the Artificial Neural Networks is reflected in a minimum number of studies (one study in 2020), as cited in the previous

introduction. To the best of our knowledge, this work will be the first attempt to model the organic molecules' rejection mechanisms by forward osmosis membranes using the quantitative structure-property relationships based on neural networks (QSPR-NN).

The comparison between the two studies (as mentioned in Table 10) shows lower values of errors in our study compared to the study of Jasir (2020) [19], which demonstrates the excellence of our QSPR-NN model.

4. Conclusions

This work examines the application of the artificial neurons network in forward osmosis membranes process separation with the ambition of predicting the rejection of organic molecules, however, ANN models can be improved through the practice of QSPR that may summarize the interactions between properties of organic molecules, membranes characteristics, and operations conditions.

- The effect of the database, training algorithms, transfer functions, hidden neurons, and subdivisions of the database is found by the trial and error method in the way to discover the QSPR-$ANN_{optimal}$ model, which is obtained with the eleventh input including the properties of organic molecules (effective diameter of an organic molecule in water " d_c", molecular length "Length", molecular width "Width", molecular depth "Depth", Dipole moment, the logarithm of the octanol-water partition coefficient "Log P"), membrane characteristics (Surface membrane charge as" Zeta potential", and the Hydrophobicity "as Contact angle"), and operating conditions (pH, Crossflow Velocity (CFV), and the water flux). The QSPR-$ANN_{optimal}$ generated can be present with equations using the ideal transfer functions. The logarithm of the octanol-water partition coefficient "Log P" was confirmed as the input the most affecting the target (Rejection of the organic molecules by the forward osmosis membranes), in addition to the effective diameter of an organic molecule in water " d_c", molecular length "Length", molecular depth "Depth", and "Dipole moment" but with less influence than "Log p".
- The QSPR-$ANN_{optimal}$ model was compared with a literature model that applied the forward osmosis which shows satisfactory results with the correlation coefficient close to the ideal value (R = 0.9895) and low value of root mean squared error (RMSE = 4.3683%).
- The performance of the QSPR-$ANN_{optimal}$ is based on error values, this performance is demonstrated by the higher values of RER and RPD more than 2.5 and equal to 21.4356 and 3.4290 respectively, the Nash-Sutcliffe Efficiency (NSE) more than 0.9, and the other errors have fewer values which improved the excellent performance of the QSPR-$ANN_{optimal}$ obtained.
- This study demonstrates once again the success of the application of artificial neural networks in the modeling of the rejection of organic molecules, especially with membrane separation processes such as ultrafiltration, nanofiltration, reverse osmosis, and actually the forward osmosis process.

Acknowledgements

The authors gratefully acknowledge the Ministry of Higher Education of Algeria (PRFU Projects N° A16N01UN260120220004**)** and the group of Laboratory of Biomaterials and Transport Phenomena in university of Medea.

References

1. Zhao, S., Zou, L., Tang, C.Y., Mulcahy, D. (2012). Recent developments in forward osmosis: Opportunities and challenges. Journal of Membrane Science, 396, 1-21. https://doi.org/10.1016/j.memsci.2011.12.023
2. Lutchmiah, K., Verliefde, A.R.D., Roest, K., Rietveld, L.C., Cornelissen, E.R. (2014). Forward osmosis for application in wastewater treatment: A review. Water research, 58, 179197. https://doi.org/10.1016/j.watres.2014.03.045.
3. Mi, B., Elimelech, M. (2010), Organic fouling of forward osmosis membranes: Fouling reversibility and cleaning without chemical reagents. Journal of Membrane Science, 348(1-2), 337-345. https://doi.org/10.1016/j.memsci.2009.11.021.
4. Nguyen, H.T., Cho, K., Jang, A., Jeong, S. (2021). Cost analysis and scheduling of the desalination vessel using reverse osmosis technology. Membrane and Water Treatment, 12(4), 177-185. https://doi.org/10.12989/mwt.2021.12.4.177.
5. Cath, T.Y., Childress, A.E, and Elimelech, M. (2006). Forward osmosis: Principles, applications, and recent developments. Journal of Membrane Science, 281(1-2), 70-87. https://doi.org/10.1016/j.memsci.2006.05.048.
6. Shaffer, D.L., Werber, J.R., Jaramillo, H., Lin, S., Elimelech, M. (2015). Forward osmosis: where are we now?. Desalination, 356, 271-284. https://doi.org/10.1016/j.desal.2014.10.031.
7. Chung, T.S., Luo, L., Wan, C.F., Cui, Y., Amy, G. (2015). What is next for forward osmosis (FO) and pressure retarded osmosis (PRO). Separation and Purification Technology, 156, 856-860. https://doi.org/10.1016/j.seppur.2015.10.063.
8. Chung, T.S., Zhang, S., Wang, K.Y., Su, J., Ling, M.M. (2012). Forward osmosis processes: yesterday, today and tomorrow. Desalination, 287, 78-81. https://doi.org/10.1016/j.desal.2010.12.019.
9. Martinetti, C.R., Childress, A.E., Cath, T.Y. (2009). High recovery of concentrated RO brines using forward osmosis and membrane distillation. Journal of Membrane Science, 331(1-2), 31-39. https://doi.org/10.1016/j.memsci.2009.01.003.
10. McGinnis, R. L, and Elimelech, M. (2007). Energy requirements of ammonia–carbon dioxide forward osmosis desalination. Desalination, 207(1-3), 370-382. https://doi.org/10.1016/j.desal.2006.08.012
11. Zhao, S., Zou, L. (2011). Effects of working temperature on separation performance, membrane scaling and cleaning in forward osmosis desalination. Desalination, 278(1-3), 157-164. https://doi.org/10.1016/j.desal.2011.05.018.
12. Bowen, W.R., Jones, M.G., Welfoot, J.S., Yousef, H.N. (2000). Predicting salt rejections at nanofiltration membranes using artificial neural networks. Desalination. 129(2), 147-162. https://doi.org/10.1016/S0011-9164(00)00057-6.
13. Shetty, G.R., Malki, H., Chellam, S. (2003). Predicting contaminant removal during municipal drinking water nanofiltration using artificial neural networks. Journal of Membrane Science. 212(1-2), 99-112. https://doi.org/10.1016/S0376-7388(02)00473-8.
14. Dornier, M., Decloux, M., Trystram, G., Lebert, A. (1995). Dynamic modeling of crossflow microfiltration using neural networks. Journal of Membrane Science, 98(3), 263-273. https://doi.org/10.1016/0376-7388(94)00195-5.
15. Yangali-Quintanilla, V., Verliefde, A., Kim, T.U., Sadmani, A., Kennedy, M., Amy, G. (2009). Artificial neural network models based on QSAR for predicting rejection of neutral organic compounds by polyamide nanofiltration and reverse osmosis membranes. Journal of Membrane Science, 342(1-2), 251-262. https://doi.org/10.1016/j.memsci.2009.06.048.
16. Ammi, Y., Khaouane, L., Hanini, S. (2015). Prediction of the rejection of organic compounds (neutral and ionic) by nanofiltration and reverse osmosis membranes using neural networks. Korean Journal of Chemical Engineering, 32(11), 2300-2310. https://doi.org/10.1007/s11814-015-0086-y.
17. Abbasi, M., Rasouli, Y., Jowkar, P. (2018). Application of ANN modeling for oily wastewater treatment by hybrid PAC-MF process. Membrane and Water Treatment, 9(4), 285-292.

https://doi.org/10.12989/mwt.2018.9. 4.285.
18. Pardeshi, P.M., Mungray, A.A., Mungray, A.K. (2016). Determination of optimum conditions in forward osmosis using a combined Taguchi–neural approach. Chemical Engineering Research and Design, 109, 215-225. https://doi.org/10.1016/j.cherd.2016.01.030.
19. Jawad, J., Hawari, A.H., Zaidi, S. (2020). Modeling of forward osmosis process using artificial neural networks (ANN) to predict the permeate flux. Desalianation, 484, 114427. https://doi.org/10.1016/j.desal.2020.114427.
20. Ammi, Y., Khaouane, L., Hanini, S. (2021). Stacked neural networks for predicting the membranes performance by treating the pharmaceutical active compounds. Neural Computing and Applications, 33(19), 12429-12444. https://doi.org/10.1007/s00521-021-05876-0.
21. Barello, M., Manca, D., Patel, R., Mujtaba, I.M. (2014). Neural network based correlation for estimating water permeability constant in RO desalination process underfouling. Desalination, 345, 101-111. https://doi.org/10.1016/j.desal.2014.04.016.
22. Darwish, N.A., Hilal, N., Al-Zoubi, H., Mohammad, A.W. (2007). Neural networks simulation of the filtration of sodium chloride and magnesium chloride solutions using nanofiltration membranes. Chemical Engineering Research and Design, 85(4), 417-430. https://doi.org/10.1205/cherd06037.
23. Madaeni, S., Shiri, M., Kurdian, A. (2015). Modeling, optimization, and control of reverse osmosis water treatment in kazeroon power plant using neural network. Chemical Engineering Communications, 202(1), 6-14. https://doi.org/10.1080/00986445.2013.828606.
24. Si-Moussa, C., Hanini, S., Derriche, R., Bouhedda, M., Bouzidi, A. (2008). Prediciton of high-pressure vapor liquid equilibrium of six binary systems, carbon dioxide with six esters, using an artificial neural network model. Brazilian Journal of Chemical Engineering, 25(1), 183-199.
25. Maouz, H., Hanini, L.K.S. Ammi, Y. (2019). Modeling the molecular weight and number average molecular masses during the photo-thermal oxidation of polypropylene using neural networks. Moroccan Journal of Chemistry, 7(1), 17-27. https://doi.org/10.48317/IMIST.PRSM/morjchem-v7i1.12522
26. Alturki, A.A., McDonald, J.A., Khan, S.J., Price, W.E., Nghiem, L.D. Elimelech, M. (2013). Removal of trace organic contaminants by the forward osmosis process. Separation and Purification Technology, 103, 258-266. https://doi.org/10.1016/j.seppur.2012.10.036.
27. Arcanjo, G.S., Costa, F.C., Ricci, B.C., Mounteer, A.H., DE Melo, E.N., Cavalcante, B.F., Amaral, M.C. (2020). Draw solution solute selection for a hybrid forward osmosis-membrane distillation module: effects on trace organic compound rejection, water flux and polarization. Chem Eng Journal, 400, 125857. https://doi.org/10.1016/j.cej.2020.125857.
28. Cui, Y., Liu, X.Y., Chung, T.S., Weber, M., Staudt, C., Maletzko, C. (2016). Removal of organic micro-pollutants (phenol, aniline and nitrobenzene) via forward osmosis (FO) process: evaluation of FO as an alternative method to reverse osmosis (RO). Water Resaerch, 91, 104-114. https://doi.org/10.1016/j.watres.2016.01.001.
29. Heo, J., Kim, S., Her, N., Park, C.M., Yu, M., Yoon, Y. (2020). Removal of contaminants of emerging concern by FO, RO, and UF membranes in water and wastewater. Contaminants of Emerging Concern in Water and Wastewater, 139-176. https://doi.org/10.1016/B978-0-12-813561-7.00005-5.
30. Im, S.J., Lee, H., Jang, A. (2021). Effects of co-existence of organic matter and microplastics on the rejection of PFCs by forward osmosis membrane. Environmental Research, 194, 110597. https://doi.org/10.1016/j.envres.2020.110597.
31. Jamil, S., Loganathan, P., Kazner, C., Vigneswaran, S. (2015). Forward osmosis treatment for volume minimisation of reverse osmosis concentrate from a water reclamation plant and removal of organic micropollutants. Desalination, 372, 32-38. https://doi.org/10.1016/j.desal.2015.06.013.
32. Kim, Y., Li, S., Chekli, L., Woo, Y.C., Wei, C.H., Phuntsho, S., Shon, H.K. (2017). Assessing the removal of organic micro-pollutants from anaerobic membrane bioreactor effluent by fertilizer-drawn forward osmosis. Journal of Membrane Science, 533, 84-95. https://doi.org/10.1016/j.memsci.2017.03.027.
33. Kong, F.X., Yang, H.W., Wang, X.M., Xie, Y.F. (2014). Rejection of nine haloacetic acids and

coupled reverse draw solute permeation in forward osmosis. Desalination, 341, 1-9. https://doi.org/10.1016/j.desal.2014.02.019.
34. Lee, H., Im, S.J., Park, J.H., Jang, A. (2019). Removal and transport behavior of trace organic compounds and degradation byproducts in forward osmosis process: Effects of operation conditions and membrane properties. Chemical Engineering Journal, 375, 122030. https://doi.org/10.1016/j.cej.2019.122030.
35. Li, C., Li, H., Yang, Y., Hou, L.A. (2020). Removal of pharmaceuticals by fouled forward osmosis membranes: Impact of DOM fractions, Ca2+ and real water. Science of The Total Environment, 738, 139757. https://doi.org/10.1016/j.scitotenv.2020.139757.
36. Linares, R.V., Yangali-Quintanilla, V., Li, Z., Amy, G. (2011). Rejection of micropollutants by clean and fouled forward osmosis membrane. Water Research, 45(20), 6737-6744. https://doi.org/10.1016/j.watres.2011.10.037.
37. Madsen, H.T., Bajraktari, N., Hélix-Nielsen C, Van der Bruggen, B., Søgaard, E.G. (2015). Use of biomimetic forward osmosis membrane for trace organics removal. Journal of Membrane Science, 476, 469-474. https://doi.org/10.1016/j.memsci.2014.11.055.
38. Rastgar, M., Shakeri, A., Karkooti, A., Asad, A., Razavi, R., Sadrzadeh, M. (2020). Removal of trace organic contaminants by melamine-tuned highly cross-linked polyamide TFC membranes. Chemosphere, 238, 124691. https://doi.org/10.1016/j.chemosphere.2019.124691.
39. Xie, M., Nghiem, L.D., Price, W.E., Elimelech, M. (2012). Comparison of the removal of hydrophobic trace organic contaminants by forward osmosis and reverse osmosis. Water Research, 46(8), 2683-2692. https://doi.org/10.1016/j.watres.2012.02.023.
40. Xie, M., Nghiem, L.D., Price, W.E., Elimelech, M. (2014). Impact of organic and colloidal fouling on trace organic contaminant rejection by forward osmosis: Role of initial permeate flux. Desalination, 336, 146-152. https://doi.org/10.1016/j.desal.2013.12.037.
41. Xie, M., Price, W.E., Nghiem, L.D., Elimelech, M. (2013). Effects of feed and draw solution temperature and transmembrane temperature difference on the rejection of trace organic contaminants by forward osmosis. Journal of Membrane Science, 438, 57-64. https://doi.org/10.1016/j.memsci.2013.03.031.
42. Zheng, L., Price, W.E., McDonald, J., Khan, S.J., Fujioka, T., Nghiem, L.D. (2019).New insights into the relationship between draw solution chemistry and trace organic rejection by forward osmosis. Journal of Membrane Science, 587, 117184. https://doi.org/10.1016/j.memsci.2019.117184.
43. Ammi, Y., Khaouane, L., Hanini, S.(2018). A model based on bootstrapped neural networks for modeling the removal of organic compounds by nanofiltration and reverse osmosis membranes. Arabian Journal for Science and Engineering, 43(11), 6271-6284. https://doi.org/10.1007/s13369-018-3484-8.
44. Dolar, D., Zokić, T.I., Košutić, K., Ašperger, D., Pavlović, D.M. (2012). RO/NF membrane treatment of veterinary pharmaceutical wastewater: comparison of results obtained on a laboratory and a pilot scale. Environmental Science and Pollution Research, 19(4), 1033-1042. https://doi.org/10.1007/s11356-012-0782-7.
45. Santos, J.L., de Beukelaar, P., Vankelecom, I.F., Velizarov, S., Crespo, J.G. (2006). Effect of solute geometry and orientation on the rejection of uncharged compounds by nanofiltration. Separation and Purification Technology, 50(1), 122-131. https://doi.org/10.1016/j.seppur.2005.11.015.
46. Garson, D.G. (1991), Interpreting Neural-Network Connection Weights. AI Expert, 6(7), 47-51.
47. Gevrey, M., Dimopoulos, I., Lek, S. (2003). Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecological Mdelling, 160(3), 249-264. https://doi.org/10.1016/S0304-3800(02)00257-0.
48. Ammi, Y., Hanini, S., Khaouane, L. (2021). An artificial intelligence approach for modeling the rejection of anti-inflammatory drugs by nanofiltration and reverse osmosis membranes using kernel support vector machine and neural networks. Comptes Rendus Chimie, 24(2), 243-254. https://doi.org/10.5802/crchim.76.
49. Maouz, H., Khaouane, L., Hanini, S., Ammi, Y., Hamadache, M., Laidi, M. (2020). QSPR studies of

carbonyl, hydroxyl, polyene indices, and viscosity average molecular weight of polymers under photostabilization using ANN and MLR approaches. Kem. Ind, 69(1-2), 1-16. https://doi.org/10.15255/KUI.2019.022.

50. Hamadache, M., Benkortbi, O., Hanini, S., Amrane, A. (2018). QSAR modeling in ecotoxicological risk assessment: application to the prediction of acute contact toxicity of pesticides on bees (Apis mellifera L.). Environmental Science and Pollution Research, 25(1), 896-907. https://doi.org/10.1007/s11356-017-0498-9.
51. Zhang, Y., Gao, X., Smith, K., Inial, G., Liu, S., Conil, L.B., Pan, B. (2019).Integrating water quality and operation into prediction of water production in drinking water treatment plants by genetic algorithm enhanced artificial neural network. Water research. 164, 114888. https://doi.org/10.1016/j.watres.2019.114888.
52. Ammi, Y., Khaouane, L., Hanini, S. (2020). A comparison of "neural networks and multiple linear regressions" models to describe the rejection of micropollutants by membranes. Kemija u industriji: Časopis kemičara i kemijskih inženjera Hrvatske, 69(3-4), 111-127. https://doi.org/10.15255/KUI.2019.024.
53. Rossel, R.V., McGlynn, R., McBratney, A. (2006). Determining the composition of mineral-organic mixes using UV–vis–NIR diffuse reflectance spectroscopy. Geoderma, 137(1-2), 70-82. https://doi.org/10.1016/j.geoderma.2006.07.004.
54. Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASAB, 50(3), 885-900. https://doi.org/10.13031/2013.23153.
55. Zeybek, M. (2018). Nash-sutcliffe efficiency approach for quality improvement. J Appl Math Comput (JAMC), 2(11), 496-503. https://doi.org/10.26855/jamc.2018.11.001.
56. Gupta, D.K., Prasad, R., Kumar, P., Mishra, V.N., Vishwakarma, A.K., Srivastava, P.K. (2015). Support vector regression for retrieval of soil moisture using bistatic scatterometer data at X-band. Internet Journal of Geology and Envirionmental Engineering, 9(10), 1201-1204. https://doi.org/10.5281/zenodo.1108400.
57. Briand, C., Bressy, A., Ghassan, C., Deroubaix, J.F., Deshayes, S., Deutsch, J.C., Zeglil, Z. (2018). Que Sait-On des Micropolluantsd dans les Eaux Urbaines? ARCEAU IdF; AFB-Agence française pour la biodiversité, France.
58. Mailler, R. (2015). Devenir des micropolluants prioritaires et émergents dans les filières conventionnelles de traitement des eaux résiduaires urbaines des grosses collectivités (files eau et boues), et au cours du traitement tertiaire au charbon actif, Doctoral dissertation, Université Paris-Est, Paris, France.
59. Soulier, C., Gabet, V., Lardy, S., Lemenach, K., Pardon, P., Esperanza, M., ... Budzinski, H. (2011). Zoom sur les substances pharmaceutiques: présence, partition, devenir en station d'épuration. TSM. Techniques Sciences Méthodes–Génie Urbain, Génie Rural, 1, 63-77. https://doi.org/10.1051/tsm/201101063
60. Caron, G., Ermondi, G. (2006). Lipophilicity, Polarity and Hydrophobicity. In Comprehensive Medicinal Chemistry, 5, 425-452. Elsevier.