APPLICATION OF EVOLUTIONARY POLYNOMIAL REGRESSION IN ULTRAFILTRATION SYSTEMS CONSIDERING THE EFFECT OF DIFFERENT PARAMETERS ON OILY WASTEWATER TREATMENT

Amin Reyhani¹, Mahmood Hemmati², Fatemeh Rekabdar³*, and Mehdi Ahmadi⁴

¹ Department of Chemical and Petroleum Engineering, Sharif University of Technology, Tehran, Iran.
² Deputy of Technology and International Affair, Research Institute of Petroleum Industry (RIPI).
³ Polymer Science and Technology Division, Research Institute of Petroleum Industry (RIPI), Tehran, Iran.
⁴ Department of Chemical Engineering, Sahand University of Technology, Tabriz, Iran.

ABSTRACT

In the present work, the effects of operating conditions including pH, transmembrane pressure, oil concentration, and temperature on fouling resistance and the rejection of turbidity for a polymeric membrane in an ultrafiltration system of wastewater treatment were studied. A new modeling technique called evolutionary polynomial regression (EPR) was investigated. EPR is a method based on regression algorithm, which combines the best properties of the conventional numerical regression technique. This paper employs the capability of EPR as a powerful tool to develop a formula with a variable number of polynomial coefficients. Herein, the evolutionary polynomial regression approach is adopted on two parametric studies, i.e. total fouling resistance and rejection rate. These parameters are all evaluated as a function of some mentioned independent variables. Maximum average error and minimum average error are obtained to be equal to 4.25% and 0.05%, respectively. Therefore, EPR is a practical and useful method to describe a membrane performance.

Keywords: Ultrafiltration, Wastewater, Fouling, Rejection, Evolutionary Polynomial Regression.

INTRODUCTION

Nowadays, the environmental issue is a critical subject of concern in different manufacturing industries [1]. Oily wastewaters generated by industrial centers, particularly refineries, are one of the main problems of the environment [2]. Petrochemical processing, petroleum purifying, and natural gas and oil production are the most significant industries in which large quantities of wastewaters with high contents of oil are being generated. In these industries, the largest single wastewater stream is the produced water, which consists of heavy metals, oil, salts, and other organics [3]. Oily wastewaters and oil-water emulsions are the most important pollutants in the environment [4]. In recent years, membrane separation methods like microfiltration (MF), ultrafiltration (UF) [5], nanofiltration (NF), and reverse osmosis (RO) are being used for wastewater and produced water treatment [1]. The main drawback of membranes is the fouling of their pore spaces,
which is caused because of concentration polarization and chemical interaction with water constituents. The investigation of the fouling is worthwhile because fouling causes dramatic flux reductions during operation, affects selectivity negatively, increases the operational cost, and requires frequent membrane replacement. Therefore, knowledge about the effect of operating conditions on fouling of membrane is essential [6]. One of the most important concerns in wastewater purification by membrane filtration is removal of solutes. Recently, modeling the permeation flux decline, fouling of membrane, and rejection rate evaluation have been the subject and challenging issue of many studies. Many researchers have represented their models in the cases of analytical modeling and modeling with machine learning such as genetic programming (GP) [7-9] and evolutionary polynomial regression (EPR) in the field of modeling [10,11].

In recent years, more attention has been given to the modeling based on genetic programming. In genetic programming, there is no need to have knowledge about neither the physics of the problem nor the design of the model. Shokrkar et al. [7] studied the treatment of oily wastewaters with synthesized ceramic microfiltration membranes and proposed a new approach for modeling flux membrane using GP. The results obtained from the genetic programming model demonstrated acceptable agreement with the experimental data with an average error of less than 5%. Hwang et al. [8] modeled and predicted membrane fouling rate in a pilot-scale drinking water production system using GP to discover the mathematical function for the pattern of the membrane fouling rate. The model adopted the input parameters for operating conditions (flow rate and filtration time) and feed water quality (turbidity, pH, and temperature). The proposed model successfully simulated the pattern of membrane resistance during the operational period. Okhovat et al. [9] developed robust models based on experimental data to predict the membrane rejection of arsenic, chromium, and cadmium ions in a NF pilot-scale system using GP. The results of the proposed GP models showed excellent consistency with the experimental results. The performance and precisions of the proposed GP models were quite satisfactory.

In the evolutionary polynomial regression, there is a hybrid of capacities of conventional numerical regression and genetic programming. Savic et al. [10] modeled the number of collapses and blockages in two sewer systems using EPR to develop the set of formulas. Two approaches were implemented; first, two types of sewer failures (collapses and blockages) were recorded during a 5-year period. Second, the pipe data (age, size, etc.) was considered. The value of coefficient of determination (COD) was close to 1, which showed the best consistency between the experimental data and models.

In this study, it has been envisaged that the EPR model can be a potential candidate in order to figure out the dependence of the total fouling resistance and rejection rate of turbidity upon our variables including pH, transmembrane pressure (TMP), oil concentration, and temperature.

### EPR in Brief

Numerical regression as a powerful data analyzing method is commonly used to estimate the best fitting model for a set of experimental data. However, the type of a function (exponential, logarithmic, linear, etc.) must be selected before the fitting procedure commences. On the other hand, genetic programming as a simple, but strong, strategy based on artificial intelligence is utilized for the computer learning inspired by natural evolution to find a suitable mathematical model to fit a set of points. The computer generates and evolves a
whole population of functional expressions. The automated induction of mathematical descriptions of data using GP is usually referred to as symbolic regression [7-9]. Evolutionary polynomial regression (EPR) as a synergistic technique is a recently developed data-hybrid regression method by Giustolisi and Savic. This method integrates the best characteristics of genetic programming with that of numerical regression. In brief, the EPR consists of the following set of equations [11, 12]:

\[ Y = a_0 + \sum_{j=1}^{m} a_j (x_i)^{ES(j,1)} \cdots (x_k)^{ES(j,k)} \cdot \]

\[ f ((x_i)^{ES(j,k+1)}) \cdots f ((x_k)^{ES(j,k)}) ] \quad \text{Case 0} \]

\[ Y = a_0 + \sum_{j=1}^{m} a_j (x_i)^{ES(j,1)} \cdots (x_k)^{ES(j,k)} \cdot \]

\[ f ((x_i)^{ES(j,k+1)}) \cdots (x_k)^{ES(j,k)}) ] \quad \text{Case 1} \]

\[ Y = a_0 + \sum_{j=1}^{m} a_j (x_i)^{ES(j,1)} \cdots (x_k)^{ES(j,k)} \cdot \]

\[ g (a_0 + \sum_{j=1}^{m} a_j (x_i)^{ES(j,1)} \cdots (x_k)^{ES(j,k)} ) \quad \text{Case 2} \]

\[ Y = g (a_0 + \sum_{j=1}^{m} a_j (x_i)^{ES(j,1)} \cdots (x_k)^{ES(j,k)} ) \quad \text{Case 3} \]

where, \( x_i \) is the vector of the \( k \) candidate inputs; \( ES \) represents the matrix of exponents (coded as integers in the genetic algorithm (GA)); \( f \) and \( g \) are functions defined by the user; \( a_j \) are constant values and \( m \) stands for the length of the expressions. The \( g \)-function in Case 3, due to the upcoming parameter evaluation, is required to be invertible. The technique permits for the single exponent \( ES(j,h) \) to equal zero. This means that the \( h \)-th input variable takes a constant value equal to 1, and is thus deselected (1) since it does not significantly affect the output. As a statistical perspective, this means that such a variable \( (x_i) \) is not significant enough to be considered explaining the phenomenon [11]. The least squares (LS) method is utilized to estimate parameters \( a_j \) in the EPR procedure [13]. The LS ensures a two-way correspondence between the pseudo-polynomial structure and its coefficients. In addition to the common LS search, the user can dictate the LS to search for structures that only contain positive coefficients \( (a_j>0) \). This is particularly useful in modeling the systems in which there is a high probability that the negative coefficient values \( (a_j<0) \) are chosen to balance the particular realization of the errors related to the finite training data set [14].

The model accuracy, or consistency with the observed data, is evaluated using the coefficient of determination (COD) as follows [15]:

\[ COD = 1 - \frac{\sum_{N} (y - y_{exp})^2}{\sum_{N} (y_{exp} - \text{avg}(y_{exp}))^2} \quad (2) \]

where, \( N \) is the number of experiments; \( y \) stands for the value predicted by the generated polynomial model, and \( \text{avg}(y_{exp}) \) is the average value of the corresponding observations. Equation (2) shows that COD is strictly connected with cost functions [15]. Figure 1 depicts the flowchart of evolutionary polynomial regression paces. At the top of the flowchart, the steps of the procedure have been shown. The prime steps consist of inputting data as EPR settings and that of the user defined. The successive steps are the evaluation of formula using least squares method. Eventually, the genetic algorithm is used for the evolutionary stage of EPR, which is employed to select the set of independent variables \( (X) \) that must form the model structure [14,16].

**MATERIALS AND METHODS EXPERIMENTAL SETUP**

Figure 2 shows the ultrafiltration module used in the experimental setup. It is made of stainless steel with an effective area of 44 cm\(^2\). As demonstrated, the membrane is held by a support.
Figure 1: The EPR flowchart [14,16]

Figure 2: Ultrafiltration module

Figure 3 shows the diagram of the filtration setup. The wastewater in the feed tank was pumped to a flat sheet membrane module. The needed flow velocity and pressure were attained by controlling the input electromotor power and the back pressure valve after the membrane cell. As temperature is one of the controlling factors, a cooling/heating system was used to maintain the required temperature. All the experiments conducted in cross flow operation were carried out in the concentration mode of filtration (CMF) for 150 minutes, and the weight of permeates were weighed. Each series of experiments was performed under various conditions under different pH, TMP, oil concentration, and temperature.

Synthetic Wastewater

The synthetic feed used in the experiments consisted of gas-oil, deionized water, and surfactant (Tween 85, Merck chemicals). 10 minutes before the addition of gas-oil to water, the surfactant was dissolved in water (which was the continuous phase of the emulsion), and gas-oil was then added gradually during 1.5 hours alongside mixing water. Various emulsions with 0.1, 0.3, 0.6, 0.8, and 1 (%) of gas-oil and a constant volume of surfactant (equal to 10% of oil) were prepared.

Membrane Characterization

A polyacrylonitrile membrane (YMMWSP-1905), which was purchased from Sepro Company, was used in this study. The technical characteristics of this membrane are given in Table 1. The fouling of the membrane pores was observed with a scanning electron microscope [17]. The scanning electron microscope was operated with a maximum voltage of 30 kV (model XL30, Philips). Figure 4 illustrates the cross section SEM image of a fresh polymeric membrane.
(PAN) before filtration. As it can be seen, at first, there is no cake layer on the membrane surface and no fouling in the pores.

Table 1: Characteristics of polyacrylonitrile membrane

<table>
<thead>
<tr>
<th>Commercial name</th>
<th>YMMWSP1905</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material</td>
<td>Polyacrylonitrile (PAN)</td>
</tr>
<tr>
<td>MWCO (kDa)</td>
<td>100</td>
</tr>
<tr>
<td>Typical Flux/bar (lit/m².h.bar)</td>
<td>130</td>
</tr>
<tr>
<td>pH range</td>
<td>1-10</td>
</tr>
</tbody>
</table>

Turbidity Measurement

Turbidity parameter is a measure of the quantity of the colloidal and suspended non-perceptible particles within the sample. A turbidimeter measures turbidity by calculating the intensity of the light being scattered by the particles present in the feed. Turbidimeter measures the turbidity in terms of Nephelo turbidity unit (NTU) according to D-1889-94 standard.

Calculations

All the equations needed for the calculation of total fouling resistance and rejection rate are given below. The resistance against permeation can be obtained from Darcy’s law [18]:

\[ J = \frac{\Delta P}{\mu \sum R} \]  

(3)

where, \( J \) is out-flowing permeate flux; \( \Delta P \) represents the pressure difference applied on the two sides of membrane; \( \mu \) and \( \sum R \) stand for viscosity and the sum of resistances against fluid permeation, respectively. \( \sum R \) can be considered as the sum of the inherent resistance of the membrane against the flow \( (R_m) \) and the resistance arising from the membrane blockage and cake formation \( (R_f) \). \( R_m \) can be calculated by Equation 4; for this purpose, the flux of the distilled water must be measured through the unused membrane. Under this condition, there is no blockage in the membrane, i.e. \( R_f = 0 \) [18].

\[ J_{wi} = \frac{\Delta P}{\mu R_m} \]  

(4)

where, \( J_{wi} \) is the permeate flux of the distilled water. By progress of the filtration process and the blockage of membrane, the resistance arising from the blockage and cake formation must also be taken into calculation. Therefore, Equation 3 must be written in the following form [18]:

\[ J_{ww} = \frac{\Delta P}{\mu R_t} \]  

(5)

where, \( J_{ww} \) is the permeate flux of the distilled water after the filtration process. \( R_t \) is the sum of the \( R_m \) and \( R_f \)-fouling resistance arisen from filtration. And rejection percent was calculated as follows:

\[ Rejection (\%) = \left( 1 - \frac{T_p}{T_f} \right) \times 100 \]  

(6)

where, \( T_p \) and \( T_f \) are the turbidity values of permeate and feed respectively.

RESULTS AND DISCUSSION

EPR Settings and Models for Total Fouling Resistance Analysis Using MATLAB Toolbox

The EPR settings used in supplying the model search are reported in Table 2. The type of regression is set static for \( R_t \) variations. The polynomial structure has been reported in Table
2. It is considered to enable EPR to select building blocks such as \( f(\sum_{j=1}^{K} x_j \cdots x_j^{2K}) \), where \( x_j \)'s are input variables including pH, TMP (bar), oil concentration (% v/v), and temperature (°C).

As it was mentioned, many of the experiments were performed in a way that the total resistance fouling depended on pH, TMP, oil concentration, and temperature. In each series of the samples, one of the variables varied while the others were kept constant. The EPR predicts the best matching between \( R_t \) and these independent variables based on EPR settings reported in Table 2. The results of EPR as a number of \( a_j \)'s in the set of predicted formulas are given in Table 3.

**Table 2: EPR settings**

<table>
<thead>
<tr>
<th>Regression type</th>
<th>Static</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polynomial structure</td>
<td>Case 0 of Equations 1</td>
</tr>
<tr>
<td>Function of type</td>
<td>No Function</td>
</tr>
<tr>
<td>Number of ( a_j )</td>
<td>See table 3</td>
</tr>
<tr>
<td>Range of exponents</td>
<td>([0, 0.5, 1])</td>
</tr>
<tr>
<td>Offset (( a_0 ))</td>
<td>Yes</td>
</tr>
<tr>
<td>Constant estimation method</td>
<td>Least squares</td>
</tr>
<tr>
<td>Number of generations</td>
<td>10</td>
</tr>
</tbody>
</table>

**Effect of pH on Total Fouling Resistance and Rejection**

Figure 5 compares the prediction quality of EPR model and the experimental measurements. Values of TMP, oil concentration, and temperature were chosen to be 3.5 bar, 0.30 (% v/v), and 30 °C, respectively. Increasing pH increases the Zeta potential value; thus, the thickness of the cake generated by the filtration is reduced due to inter-droplet repulsion and the feed solution is stable, which causes the fouling to decrease (Figure 5.a). In this case, the average error according to error formula is 4.25% (error (%) = \((Y_{exp.-Y_{model}})/Y_{exp.}\times100\); thereafter, all calculated errors are reported based on this equation. Furthermore, Figure 5.b shows that rejection increases by rising pH because of supramolecular forces. In basic media, coagulation and aggregation of oil droplets occur and thus rejection percentage increases. Esch [19] may attribute this phenomenon to the presence of supramolecular forces between oil droplets. The average error obtained is equal to 0.16%.

**Effect of TMP on Total Fouling Resistance and Rejection**

The variation of fouling resistance against TMP is illustrated in Figure 6.a. In this case, the values of pH, oil concentration, and temperature are set at 7, 0.30 (% v/v), and 30 °C respectively. EPR model has a good correspondence with the experimental values and predicts the enhancement of fouling resistance with an increase in TMP. The Higher TMP values cause a higher compaction of the filter cake on the membrane surface (Figure 7.b) and more compression in internal fouling [18,20]; average error for this part is 2.92%.

**Table 3: Selected EPR results for \( R_t \) variations vs. independent variables**

<table>
<thead>
<tr>
<th>States</th>
<th>No. of ( a_j )</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_t ) vs. pH</td>
<td>3</td>
<td>( R_t = 27.2474+14.972332pH^{0.5}+2.432273pH )</td>
</tr>
<tr>
<td>( R_t ) vs. TMP</td>
<td>3</td>
<td>( R_t = 5.2077+1.6611TMP^{0.5}+1.61865TMP )</td>
</tr>
<tr>
<td>( R_t ) vs. oil Concentration</td>
<td>2</td>
<td>( R_t = 6.635081+2.4372C^{0.5} )</td>
</tr>
<tr>
<td>( R_t ) vs. temperature</td>
<td>2</td>
<td>( R_t = 0.200188-0.019481T^{1.1} )</td>
</tr>
</tbody>
</table>

* TMP = Transmembrane pressure
† C = Oil concentration
** T = Temperature
--The values of \( R_t \) in each part must be multiplied by 10^{12}
Figure 6.b indicates that rejection falls by increasing pressure. This behavior can be attributed to the permeation of oil droplets through gel layer under high pressures. Okhovat et al. [9] modeled the behavior of the rejection of arsenic, chromium, and cadmium by a nanofiltration pilot-scale system using GP and obtained average errors of the rejection of arsenic, chromium, and cadmium respectively equal to 0.216%, 0.836%, and 1.796%. These experiments were carried out in the range of 5 to 14 bar and the feed temperature was kept at 30 °C. Here, the average error of rejection versus temperature is 0.05%.

Figure 7 illustrates the surface of the membrane in two states. Figure 7.a shows the surface of the membrane without any cake layer before the filtration and Figure 7.b indicates the existence of an oil layer on the membrane surface at the end of the filtration process. The surface is covered by a cake layer acting as a resistance in the membrane filtration.

Figure 6: (a) Total fouling resistance variation vs. TMP; (b) Rejection rate vs. TMP (pH=7, oil concentration=0.30 (% v/v), and temperature=30 °C.)

Effect of Oil Concentration on Total Fouling Resistance and Rejection

Figure 8 shows a fine correspondence between the experimental measurements and the EPR model. At this stage, the oil concentration was varied and the other parameters were kept constant. At high concentrations, the average size of oil droplets was greater at the COD=0.96

COD=0.93
polarization layer than the feed.

Figure 7: SEM image of membrane (PAN-100 KDa) surface: (a) Before the filtration without a gel layer; (b) At the end of the filtration with a cake layer.

Previous studies also showed a similar behavior [21,22]. Thus, high concentrations of oil in the gel layer were the source of the tensile force caused by high viscosity. This force fuses the oil droplets together on the membrane surface. Figure 8.a implies that the total fouling resistance of the membrane increases with the formation of such a layer; average error is 1.71%. As Figure 8.b shows, rejection rate increases as a result of gel layer formation. The higher the oil concentration was, the greater the rejection rate became. Okhovat et al. [9] modeled the behavior of the rejection of arsenic, chromium, and cadmium by a nanofiltration pilot-scale system using GP and obtained average rejection errors of arsenic, chromium, and cadmium respectively equal to 0.216%, 0.836%, and 1.796%. These experiments were carried out at a feed temperature of 30 °C. The concentrations of arsenic and chromium in the feed were in the range of 100-400 µg.lit⁻¹ and the feed concentration of cadmium was 20-80 µg.lit⁻¹. At higher concentrations, rejection rate became independent of concentration. This was because the $R_t$ was almost constant due to internal blockage. The average error was 2.11%.

Figure 8: (a) Total fouling resistance variation vs. oil concentration; (b) Rejection rate vs. oil concentration (pH=7, TMP=3.5 bar, and temperature=30 °C).

Effect of Temperature on Total Fouling Resistance and Rejection

The EPR model predicted a decrease in fouling resistance and rejection rate by increasing...
temperature. The values of the independent variables \( \text{pH}, \text{TMP}, \) and oil concentration were set at 7, 3.5 bar, and 0.30 (\%, v/v), respectively. The comparison of the experimental data with the EPR model is demonstrated in the Figure 9.

This phenomenon affects the thickness of the gel layer and reduces it. As the temperature increases, the total fouling resistance of the membrane decreases (Figure 9.a) [20,23]. Hwang et al. [8] found an average error of 8% using genetic programming in polymeric membrane (PVDF) for the prediction of membrane fouling in the pilot-scale microfiltration system. The range of temperature in their study was between 2.79 °C and 22.58 °C. In our case, the average error was 2.39%. Figure 9.b implies that rejection rate decreases with increasing temperature due to an increase in oil permeation. From comparison the range of temperature in their study was between 2.79 °C and 22.58 °C. In our case, the average error was 2.39%. Figure 9.b implies that rejection rate between the experimental data and the EPR prediction, the average error was obtained to be 0.3%.

This phenomenon affects the thickness of the gel layer and reduces it. As the temperature increases, the total fouling resistance of the membrane decreases (Figure 9.a) [20,23]. Hwang et al. [8] found an average error of 8% using genetic programming in polymeric membrane (PVDF) for the prediction of membrane fouling in the pilot-scale microfiltration system. The range of temperature in their study was between 2.79 °C and 22.58 °C. In our case, the average error was 2.39%. Figure 9.b implies that rejection rate decreases with increasing temperature due to an increase in oil permeation. From comparison the range of temperature in their study was between 2.79 °C and 22.58 °C. In our case, the average error was 2.39%. Figure 9.b implies that rejection rate between the experimental data and the EPR prediction, the average error was obtained to be 0.3%.

\[ \text{COD} = 0.93 \]

\[ \text{Rejection} \times 100 = 99.5 \]

\[ \text{Number of } \sigma = 3 \]

Figure 9: (a) Total fouling resistance variation vs. temperature; (b) Rejection rate vs. temperature. (\( \text{pH} = 7, \text{TMP} = 3.5 \) bar, and oil concentration = 0.30 (\%, v/v)).

This phenomenon affects the thickness of the gel layer and reduces it. As the temperature increases, the total fouling resistance of the membrane decreases (Figure 9.a) [20,23]. Hwang et al. [8] found an average error of 8% using genetic programming in polymeric membrane (PVDF) for the prediction of membrane fouling in the pilot-scale microfiltration system. The range of temperature in their study was between 2.79 °C and 22.58 °C. In our case, the average error was 2.39%. Figure 9.b implies that rejection rate decreases with increasing temperature due to an increase in oil permeation. From comparison the range of temperature in their study was between 2.79 °C and 22.58 °C. In our case, the average error was 2.39%. Figure 9.b implies that rejection rate between the experimental data and the EPR prediction, the average error was obtained to be 0.3%.

\[ \text{COD} = 0.77 \]

\[ \text{Number of } \sigma = 3 \]

Figure 10: SEM image of PAN-100 kDa membrane as an asymmetric membrane: (a) Cross section; (b) Lateral cross section of a high porosity layer at the end of the filtration.
CONCLUSIONS

In this work, the effects of independent parameters including pH, TMP, oil concentration, and temperature on total fouling resistance and rejection rate in a polymeric membrane were investigated. Increasing temperature and pH reduced total fouling resistance and raising oil concentration and TMP increased total fouling resistance. Also, it was shown that an increase in oil concentration and pH enhanced the rejection of turbidity while increasing TMP and temperature decreased rejection rate. Evolutionary polynomial regression was used to predict the variation of total fouling resistance and rejection rate. The predicted values for total fouling resistance were compared with the measurements obtained by Darcy’s law. Likewise, the predicted values for rejection rate by the EPR were compared with the experimental data. The comparison demonstrated that the EPR was a suitable method for modeling membrane ultrafiltration process, because the maximum and minimum average errors obtained were 4.25% and 0.05%, respectively.

REFERENCES


Application of Evolutionary Polynomial Regression...


