Journal of ENVIRONMENTAL Treatment Techniques

J. Environ. Treat. Tech.
ISSN: 2309-1185

Journal web link: http://www.jett.dormaj.com

Mathematical Modelling for Predicting Rejection of Trace Organic Contaminants by the Nanofiltration Membrane NF270

Hai Quang Dang*

Faculty of Environment, the University of Danang, University of Science and Technology, Danang, Vietnam

Received: 21/01/2020 Accepted: 13/04/2020 Published: 20/09/2020

Abstract

This study implemented multiple linear regression model to predict rejection of trace organic contaminants (TrOCs) by the nanofiltration (NF) membrane NF270. Multiple regression analysis by the Statgraphics Centurion software were used to find an optimal mathematical modeling that combines interactions between molecular width, molecular height, molecular length, molecular weight and log D of TrOCs for predicting rejection. The result shows a relatively good agreement between the predicted rejection and the observed rejection and an acceptable R-squared correlation coefficient were found (R^2 = 91.42 %) for the best model. In conclusion, a unified general multiple linear regression equation was able to predict rejections of TrOCs during nanofiltration with the explanatory variables of molecular width, molecular height, molecular length and molecular weight. Moreover, the present approach is a basis to continue investigation using multiple regression analysis techniques for understanding rejection of TrOCs by the NF membranes.

Keywords: Trace organic contaminants (TrOCs), Nanofiltration, Multiple linear regression, Mathematical modeling, Rejection

1 Introduction

The demand for fresh water worldwide is increasing dramatically caused by continued population growth posing challenges in the last few decades by growing water stress, both in terms of water scarcity and quality deterioration. Some of most important problems in water supply are the necessity of fresh water production for drinking, domestic, agricultural, landscape or industrial uses, the requirement of higher performance methods for wastewater reclamation and reusing applications, as well as lower maximum levels of contaminants [1]. However, the potential presence of trace organic contaminants (TrOCs) such as endocrine disrupting chemicals (EDCs), pharmaceutically active compounds (PhACs) and disinfection by products (DBPs) in treated wastewater and other water sources has become a public concern because of their potential risks on ecological and human health in recent decades [2-4].

Recognizing these problems, the rejection of TrOCs in water treatment processes, which are associated with potentially adverse human health effects, is of increasing interest for membrane applications. Nanofiltration (NF) has been demonstrated to be appropriate technologies for removing most TrOCs [5,6]. An important driving force for the widespread implementation of NF membranes is their high removal efficiency for a large number of inorganic salts and TrOCs amongst the membrane processes [7-9]. This will have special significance because satisfactory elimination of micropollutants in water sources is of paramount importance for the protection of public health. It has been found that the physicochemical properties of TrOCs, such as molecular size, molecular weight (MW), hydrophobicity and charge caused by the functional groups, have significant effects on their rejection by NF membranes [10,11]. According to Kimura et al. [12], the molecular size of the TrOCs could be considered one of the most important factors influencing their rejection by the NF membranes. The molecular weight of the solutes is often used as an indication of size while the molecular size parameters such as molecular width, molecular length and molecular height have been confirmed to more appropriate predicators for describing size exclusion effects on the rejection of TrOCs by NF membranes [13-16].

A number of articles have proposed a mechanistic understanding of the rejection of TrOCs, others have tried to apply fitting parameter models to model rejection [17-19]. However, there have been few models to predict the rejection of TrOCs by NF membranes. It would therefore be of interest to have a statistical model that can use for predicting rejection of TrOCs. Among modeling approaches, multiple linear regression analysis is a relatively simple statistical method used to examine the correlation among variables. The present study is aimed at developing multiple regression model that can usefully estimate the rejection of TrOCs by NF membrane based on an integral approach that considers physicochemical properties of the compounds.

Corresponding author: Hai Quang Dang, Faculty of Environment, the University of Danang, University of Science and Technology, Danang, Vietnam. E-mail: dqhai@dut.udn.vn.

2 Materials and methods

2.1 Background of statistical analysis

Statistical analysis is useful for exploring and examining the basic features of the data prior to applying statistical tests and fitting statistical models. Because the fact that many factors influence some phenomena, it is necessary to calculate the interaction among phenomena. In order to attain this, multiple regression methods can be used. Multiple regression has taken a very significant place in statistical science. It is a method of analysis for assessing the strength of the relationship between a set of explanatory variables known as independent variables, and a single response or dependent variable. Applying multiple regression analysis to a set of data results in what are known as regression coefficients, one for each explanatory variable [20,21]. Regression analysis is a mathematical measure of the average relationship between two or more variables in terms of the original units of the data. The concept of regression analysis involves finding the best relationship between variables.

In case of researching relationship between two phenomena and in case of prediction of the value of dependent variable, first to identify variables and then to find out random sample n size for the chosen values of dependent variables. Suppose that k phenomenon is identified as independent variable (predictor), or x_i, i = 1, 2,…, k, and Y as dependent random variable. The whole multiple linear model can be presented as one equation for dependent variable Y_i:

$$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon_i \quad (1)$$

where Y_i is dependent random variable, x_1, x_2, …, x_k are values of independent variable, β_0, β_1, …, β_k are model parameters (regression coefficient), and ε_i is a supporting element, or a random error which has normal distribution, zero mean and constant variance. The whole regression model can be estimated by the sample regression model using least squares fitted (prediction) equation (obtained by minimizing Error Sum of Squares, SSE):

$$y_i = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k \quad (2)$$

where y_i is adjustable or foreseen value of dependent variable Y_i, x_1, x_2, …, x_k are values of independent variables, and b_0, b_1, …, b_k are estimations of unknown parameters β_0, β_1, …, β_k. Once a regression model has been constructed, it may be important to confirm the goodness of fit of the model and the statistical significance of the estimated parameters. Commonly used techniques to verify the goodness of fit include the R square, Adjusted R square, Multiple R, and hypothesis testing. Statistical significance can be verified by a Fisher distribution (F-test) of the overall model, followed by tests of the individual parameters using Student's t-distribution (t-test) [22].

The fit of a multiple regression model can be judged with calculation of the multiple correlation coefficient, Multiple R, defined as the correlation between the observed values of the response variable and the values predicted by the model. The squared value of R (R^2) gives the proportion of the variability of the response variable accounted for by the explanatory variables. Adjusted R square used to compare models with different sets of independent variables in terms of predictive capabilities. Analysis of variance (ANOVA) will provide an F-test of the null hypothesis that each of b_0, b_1, b_2,… b_k, is equal to zero, or in other words that R^2 is zero [21].

2.2 Nanofiltration membrane

The NF270 membrane (Dow-Filmtec, Minneapolis, MN) was selected for this study. According to the manufacturer, it is a thin-film composite polyamide membrane that is widely used for water and wastewater treatment application. This is a loose NF membrane with a relatively high permeability (of approximately 11 L/bar m^2 h). At pH 4 and above, this membrane is negatively charged [23]. The flat sheet membrane samples were stored dry before use.

2.3 Trace organic contaminants, analytical chemicals and reagents

The target TrOCs for this research have been chosen from the major classes of EDCs, PhACs and DBPs. They have diverse physicochemical properties such as hydrophobicity, charge, solubility, and molecular size. A stock solution was prepared at a concentration of 1 mg/mL in pure methanol. A working solution of these TrOCs was also prepared in pure methanol. Both these solutions were stored in a freezer at -18 ºC prior to use.

Chemical solutions and feed waters were prepared with Milli-Q water. Both the solvents used for solid phase extraction and analysis of samples including methanol and dichloromethane, purchased from Sigma-Aldrich (Sydney, Australia). Internal standard of bisphenol A-d_{16} and N,O-bis (trimethylsilyl) trifluoroacetamide (BSTFA) containing 1 % of trimethylchlorosilane (TMCS). Pyridine used in the derivatization process. All reagents and chemicals were purchased from Sigma-Aldrich (Sydney, Australia).

2.4 Experimental protocol

A laboratory scale cross flow NF/RO system consisted of a stainless steel NF/RO membrane cell with an effective surface area of 40 cm^2 and channel height of 2 mm, a stainless steel feed reservoir, and a high pressure pump (Hydra-Cell, Wanner Engineering Inc., Minneapolis, MN) was used. The temperature of the feed solution was controlled by a chiller/heater (Neslab RTE 7, Thermo Scientific Inc., Waltham, MA, USA) equipped with a stainless steel heat exchanger coil which was submerged directly into the feed reservoir. A digital flow meter (Optiflow 1000, Agilent Technologies, Palo Alto, CA) connected to a PC was utilized to measure permeate flow, and the cross flow was monitored with a rotameter.

The rejection of TrOCs was performed in a background electrolyte solution containing 10 mM of NaCl, 1 mM of $CaCl_2$, and 1 mM of NaH_2PO_4 (pH 7) and conducted over 24 hours. Prior to each experiment, the NF270 membrane samples were gently washed with copious Milli-Q water to remove any preservatives. They were then compacted using Milli-Q water at 1,000 kPa for at least one hour until a stable permeate flux had been obtained. The background electrolyte solution was then added to the feed reservoir, and made up to the total feed volume of 10 litres. During the experiment, the feed reservoir temperature and cross flow velocity were kept constant at 20 ± 0.1 °C and 42 cm/s, respectively. The permeate flux was set to the manufacturer's quoted nominal membrane flux of 42 L/m^2h throughout the experiment. Both permeate and retentate were recirculated to the feed reservoir. A mixture of 25 target TrOCs was then added to the feed reservoir to obtain a concentration of 25 µg/L of each. The feed solution pH was kept constant during the experiments by periodically adding a small amount of 1 M of NaOH or 1 M HCl. Approximately 100 mL of feed and permeate samples were taken at specific times. Samples were stored in clean glass bottles,

wrapped in aluminium foil, stored in the fridge for subsequent extraction and GC/MS analysis. The effective rejection was defined as R (%):

$$R = 100 \times (1 - \frac{C_p}{C_f}) \quad (3)$$

where C_f and C_p were the feed and the permeate concentrations, respectively.

2.5 Analytical methods

The Oasis HLB SPE cartridges (6 mL, 200 mg, Waters, Milford, MA, USA) for extraction of the TrOCs in feed and permeate samples were used in this investigation. The feed and permeate samples of 100 mL were allowed to reach room temperature and adjusted by 4 M sulphuric acid to pH range between 2 and 3. Before the samples were extracted, the SPE cartridges were conditioned sequentially by 7 mL dichloromethane and methanol (1:1, v/v), 7 mL methanol, and about 2 x 7 mL reagent water on a vacuum manifold at a flow rate of 2 mL/min. Subsequently, the samples were passed through the cartridges with a flow rate of 2 mL/min. The loaded cartridges were washed with 6 x 7 mL of Milli-Q water and dried under vacuum for 30 minutes along with a stream of nitrogen. The SPE columns containing the TrOCs were eluted with 7 mL methanol followed by 7 mL dichloromethane and methanol (1:1, v/v) at a flow rate of 1 - 5 mL/min. The elution volume was then evaporated to dryness under a gentle stream of nitrogen in a water bath at 40 °C. An amount of 200 µL methanol solution containing 5 µg bisphenol A-d_{16} was utilized to dissolve the extracted residues, and was transferred into 1.5 mL vials before further evaporation to dryness under a gentle nitrogen stream. Finally, the derivatization of the dried residues in the vials was performed by adding 100 µL of BSTFA (N,O-bis (trimethylsilyl) trifluoroacetamide) (1 % TMCS (trimethylchlorosilane)) and 100 µL of pyridine (dried with KOH solid). The conditions of the derivatization reaction were 30 min at 60 - 70 °C. The derivatives were allowed to cool to room temperature before analysis by GC-MS [24].

A Shimadzu GCMS-QP5000 system consisting of a Shimadzu AOC 20i autosampler and a Phenomenex Zebron ZB-5 (5 % diphenyl - 95 % dimethylpolysiloxane) capillary column (30 m × 0.25 mm ID, df = 0.25 µm) was used to determine the concentrations of the organic compounds. Helium was used as the carrier gas at a constant flow rate of 1.3 mL/min. The GC oven temperature program was conducted as follows: 100 °C for 1 min, first ramp 10 °C/min to 175 °C, 3 min at 175 °C, second ramp 30 °C to 210 °C, third ramp 2 °C/min to 228 °C, fourth ramp 30 °C to 260 °C, fifth ramp 3 °C/min to 290 °C, 3 min at 290 °C. The injector port and the temperature of the GCMS interface were set at 280 °C. A sample volume of 1 µL was injected in splitless mode.

The MS was obtained by electron impact ionisation in full scan mode from 50 to 600 of m/z, and later on in selected ion monitoring (SIM) mode for qualitative determinations. The most abundant ions of each organic compound were selected from its spectrum for quantitation, in accordance with previous studies [25,26]. A series of standard TrOCs at 1, 10, 50, 100, 500, and 1000 ng/mL and a bisphenol A-d_{16} internal standard were prepared for the instrument calibration. The calibration curves obtained for each compound had correlation coefficients greater than 0.99. The detection limits and quantification limits for analytes were estimated with the signal to noise (s/n) ratio higher than 3 and higher than 10, respectively.

A Metrohm model 744 pH Meter was calibrated before beginning of an experiment and utilized to measure the feed solution pH for the duration of the experiment.

3 Results and discussion

3.1 Properties of trace organic contaminants and their rejection efficiency

The major physicochemical properties of the target TrOCs and their rejection efficiency are shown in Table 1. The standard deviation of data obtained from two independent experiments. The compounds selected for this investigation exhibited considerably difference in their physicochemical properties. These compounds have low molecular weight, ranging between 138.12 and 361.82 g/mol for salicylic acid and bezafibrate, respectively. However, they are markedly different in their dissociation constants (pK_a), molecular dimension (width, height and length) and hydrophobicity properties. Most TrOCs are weak acids and will dissociate into an ionic form at pH above the pK_a. The molecular widths, heights and lengths of these TrOCs are from 0.354 to 0.435, 0.505 to 1.313 and 0.615 to 1.179 nm, respectively. The difference in molecular weight and dimension can play a major role in the rejection of TrOCs.

On the other hand, it is striking to note that the intrinsic hydrophobicity of TrOCs was an important factor in determining their rejection by a NF process [27,28]. The logarithm of the effective octanol-water distribution coefficient, log D, is a good parameter which can be used to evaluate the hydrophobicity of TrOCs at any pH value [29,30]. According to Wells [31] and Alturki et al. [32], organic compounds with log D equal to 3 or higher are generally referred to as hydrophobic. By contrast, organic compounds with log D below 3 are referred to as hydrophilic.

It can be observed that rejection efficiency of TrOCs varied considerably depending on the different physicochemical characteristics of the compounds, ranging from 49.27 to 98.30 %. In general, larger MW and molecular dimension compounds showed higher rejections than small MW and molecular dimension compounds on size exclusion grounds. Because of the large MW and molecular dimension, TrOCs do not significantly penetrate into the membrane pores, resulting in their adsorption occurring mainly at the membrane surface. Consequently, the diffusion of these compounds across the membrane is very limited, leading to the high rejection efficiencies observed. These results are consistent with the observations of Yangali-Quintanilla and coworkers [33,34], who also demonstrated a strong correlation between molecular weight, width, length and the rejection for hydrophilic compounds (such as acetaminophen, phenacetine, caffeine, metronidazole, phenazone and sulfamethoxazole) by NF200 and NF90 membranes, and that rejection of these compounds may be attributed to the domination of the size exclusion effect. Agenson et al. [13] demonstrated that size exclusion effect represented by molecular weight and molecular width of solutes played a major influence on the rejection efficiency in membrane separation. Additionally, Van der Bruggen and Vandecasteele [35] suggested that the rejection of neutral TrOCs can be predicted using the molecular weight of the compound, a higher rejection for the compounds with larger molecular weight was obtained.

Table 1: Physicochemical properties and rejection efficiency of the selected trace organic contaminants

Trace organic compounds	Log K_{ow} [a]	pK_a [a]	Molecular Width (nm) [b]	Molecular Height (nm) [b]	Molecular Length (nm) [b]	Molecular weight (MW) (g/mol)	Log D at pH 7 [c]	Rejection ± STDEV (%)
Salicylic acid	2.011	3.01	0.354	0.577	0.615	138.12	-1.130	49.27 ± 3.21
Ibuprofen	3.502	4.41	0.354	0.561	0.900	206.28	0.940	80.61 ± 2.13
Gemfibrozil	4.302	4.75	0.354	0.670	0.972	250.33	2.070	93.08 ± 2.45
Diclofenac	4.548	4.18	0.354	0.767	0.829	296.15	1.770	98.03 ± 1.05
Carbamazepine	1.895	13.94	0.354	0.676	0.818	236.27	1.890	81.32 ± 2.88
Pentachlorophenol	5.115	4.68	0.354	0.640	0.659	266.34	2.850	92.74 ± 3.96
4-tert-butylphenol	3.397	10.13	0.354	0.505	0.735	150.22	3.400	51.28 ± 5.99
4-tert-octylphenol	5.180	10.15	0.354	0.595	0.822	206.32	5.180	80.77 ± 2.73
4-n-nonylphenol	6.142	10.15	0.354	0.519	1.179	220.35	6.140	88.95 ± 2.76
Triclosan	5.343	7.80	0.354	0.602	0.926	289.54	5.280	92.28 ± 1.49
Bisphenol A	3.641	10.29	0.354	0.570	0.876	228.29	3.640	78.61 ± 4.68
Estrone	3.624	10.25	0.340	0.693	0.697	270.37	3.620	82.40 ± 4.83
17β-estradiol	4.146	10.27	0.340	0.693	0.697	272.38	4.150	84.38 ± 1.10
Estriol	2.527	10.25	0.340	0.693	0.751	288.38	2.530	95.72 ± 0.40
17α-ethinylestradiol	4.106	10.24	0.356	0.693	0.788	296.40	4.110	95.36 ± 1.94
17β-estradiol acetate	5.111	10.26	0.354	0.842	0.947	314.42	5.110	98.30 ± 1.03
Caffeine	-0.628	0.52	0.412	0.676	0.750	194.19	-7.110	89.52 ± 3.75
Primidone	0.829	12.26	0.426	0.740	0.734	218.25	0.830	92.07 ± 1.86
Trimethoprim	0.594	7.04	0.420	0.766	1.047	290.32	0.310	93.64 ± 3.27
Sulfamethoxazole	0.659	5.81	0.412	0.595	1.032	253.28	-0.560	88.15 ± 1.75
Amitriptyline	4.410	9.18	0.435	0.933	0.871	277.40	4.410	91.05 ± 0.84
Bezafibrate	2.504	3.29	0.420	1.313	0.773	361.82	-1.210	98.23 ± 0.05
Linuron	3.125	12.13	0.412	0.668	0.902	249.09	3.120	81.02 ± 5.36
Formononetin	2.860	6.99	0.412	0.760	1.015	268.26	2.550	85.88 ± 5.27
Genistein	3.114	6.51	0.354	0.706	1.033	270.24	2.500	87.15 ± 5.52

[a] Scifinder Scholar, [b] calculated using Molecular Modeling Pro™ Plus software, [c] calculated by the equation: $\log D_{(pH)} = \text{Log } K_{ow} - \text{Log } (1+10^{(pH-pKa)})$.

It is however noteworthy that there was no correlation between rejection and log D of these TrOCs. These observations can be attributed to the fact that in addition to the effect of the log D, there are a number of other factors which may influence TrOCs rejection such as MW, molecular dimension, charge, and so on. Correlation between the physicochemical properties and the rejection for the TrOCs will be discussed in detail in the following section.

3.2 Multiple linear regression model for trace organic contaminants rejection

Physicochemical properties of the TrOCs and their rejection efficiency were used as indicators in the model development. Database using for multiple regression analysis are presented in the Table 2. Statistical analysis of multiple regression by Statgraphics Centurion software was used to construct the best optimal mathematical modeling for trace organic contaminants rejection by the NF270 membrane. The whole multiple regression model can be presented as one equation for dependent variable Y:

$$Y = b_0 + b_1X1 + b_2X2 + b_3X3 + b_4X4 + b_5X5 + b_{12}X1X2 + b_{13}X1X3 + b_{23}X2X3 + b_{123}X1X2X3$$

where Y (rejection efficiency) is the dependent variable; X1 (molecular width), X2 (molecular height), X3 (molecular length), X4 (molecular weight), X5 (log D) are the independent variables; b_0, b_1, b_2, b_3, b_4, b_5, b_{12}, b_{13}, b_{23}, b_{123} are the model parameters (regression coefficient).

Results of multiple regression analysis are summarised in the Table 3 and Table 4. In addition, it was discovered that regression statistics included R-squared = 91.42 percent, R-squared (adjusted for d.f.) = 86.27 percent and Standard error of Est. = 4.56.

The output shows the results of fitting a multiple linear regression model to describe the relationship between Y and independent variables. The equation of the fitted model is as follows:

$$Y = -1886.76 + 5020.79X1 + 2267.82X2 + 2092.90X3 + 0.267792X4 - 0.837608X5 - 5891.10X1X2 - 5464.52X1X3 - 2505.03X2X3 + 6448.92X1X2X3$$

In determining whether the model can be simplified, the highest P-value on the independent variables is 0.1038, belonging to X5 (Table 3). Since the P-value is greater or equal to 0.05, that term is not statistically significant at the 95.0 % or higher confidence level. Consequently, it should consider removing X5 from the model. Whereas, the regression coefficients (include X1, X2, X3, X4, X1X2, X1X3, X2X3, and X1X2X3) contribute significantly to the model (exist in multiple regression equation) due to their P-values are less than 0.05. Since the P-value in the ANOVA table is less than 0.05, there is a statistically significant relationship between the variables at the 95.0 % confidence level.

The R-squared statistic indicates that the model as fitted explains 91.42 % of the variability in Y. This value shows that more than 91.42 % of the variability in the percent of TrOCs rejection is accounted for by knowing width, height, length and molecular weight of organic compounds.

Table 2: Database using for multiple regression analysis

Trace organic compounds	Molecular Width (nm) (X1)	Molecular Height (nm) (X2)	Molecular Length (nm) (X3)	Molecular weight (MW) (g/mol) (X4)	Log D at pH 7 (X5)	X1X2	X1X3	X2X3	X1X2X3	Rejection (%) Y
Salicylic acid	0.354	0.577	0.615	138.12	-1.130	0.204258	0.217710	0.354855	0.125619	49.27
Ibuprofen	0.354	0.561	0.900	206.28	0.940	0.198594	0.318600	0.504900	0.178735	80.61
Gemfibrozil	0.354	0.670	0.972	250.33	2.070	0.237180	0.344088	0.651240	0.230539	93.08
Diclofenac	0.354	0.767	0.829	296.15	1.770	0.271518	0.293466	0.635843	0.225088	98.03
Carbamazepine	0.354	0.676	0.818	236.27	1.890	0.239304	0.289572	0.552968	0.195751	81.32
Pentachlorophenol	0.354	0.640	0.659	266.34	2.850	0.226560	0.233286	0.421760	0.149303	92.74
4-tert-butylphenol	0.354	0.505	0.735	150.22	3.400	0.178770	0.260190	0.371175	0.131396	51.28
4-tert-octylphenol	0.354	0.595	0.822	206.32	5.180	0.210630	0.290988	0.489090	0.173138	80.77
4-n-nonylphenol	0.354	0.519	1.179	220.35	6.140	0.183726	0.417366	0.611901	0.216613	88.95
Triclosan	0.354	0.602	0.926	289.54	5.280	0.213108	0.327804	0.557452	0.197338	92.28
Bisphenol A	0.354	0.570	0.876	228.29	3.640	0.201780	0.310104	0.499320	0.176759	78.61
Estrone	0.340	0.693	0.697	270.37	3.620	0.235620	0.236980	0.483021	0.164227	82.40
17β-estradiol	0.340	0.693	0.697	272.38	4.150	0.235620	0.236980	0.483021	0.164227	84.38
Estriol	0.340	0.693	0.751	288.38	2.530	0.235620	0.255340	0.520443	0.176951	95.72
17α-ethinylestradiol	0.356	0.693	0.788	296.40	4.110	0.246708	0.280528	0.546084	0.194406	95.36
17β-estradiol acetate	0.354	0.842	0.947	314.42	5.110	0.298068	0.335238	0.797374	0.282270	98.30
Caffeine	0.412	0.676	0.750	194.19	-7.110	0.278512	0.309000	0.507000	0.208884	89.52
Primidone	0.426	0.740	0.734	218.25	0.830	0.315240	0.312684	0.543160	0.231386	92.07
Trimethoprim	0.420	0.766	1.047	290.32	0.310	0.321720	0.439740	0.802002	0.336841	93.64
Sulfamethoxazole	0.412	0.595	1.032	253.28	-0.560	0.245140	0.425184	0.614040	0.252984	88.15
Amitriptyline	0.435	0.933	0.871	277.40	4.410	0.405855	0.378885	0.812643	0.353500	91.05
Bezafibrate	0.420	1.313	0.773	361.82	-1.210	0.551460	0.324660	1.014949	0.426279	98.23
Linuron	0.412	0.668	0.902	249.09	3.120	0.275216	0.371624	0.602536	0.248245	81.02
Formononetin	0.412	0.760	1.015	268.26	2.550	0.313120	0.418180	0.771400	0.317817	85.88
Genistein	0.354	0.706	1.033	270.24	2.500	0.249924	0.365682	0.729298	0.258171	87.15

The adjusted R-squared statistic, which is more suitable for comparing models with different numbers of independent variables, is 86.27 %. The standard error of the estimate shows the standard deviation of the residuals to be 4.56. This value can be used to construct prediction limits for new observations.

Table 3: Regression coefficients

Parameter	Estimate	Standard Error	T Statistic	P-Value
Constant	-1886.76	551.468	-3.42134	0.0038
X1	5020.79	1445.26	3.47397	0.0034
X2	2267.82	826.314	2.7445	0.0151
X3	2092.90	643.065	3.25458	0.0053
X4	0.267792	0.0496485	5.39375	0.0001
X5	-0.837608	0.483639	-1.73189	0.1038
X1X2	-5891.10	2121.08	-2.7774	0.0141
X1X3	-5464.52	1707.50	-3.20029	0.0060
X2X3	-2505.03	945.318	-2.64994	0.0182
X1X2X3	6448.92	2474.24	2.60642	0.0198

In conclusion, the best optimal mathematical modeling for estimating rejection of TrOCs by the NF270 membrane can be written as follows:

Y = -1886.76 + 5020.79X1 + 2267.82X2 + 2092.90X3 + 0.267792X4 - 5891.10X1X2 - 5464.52X1X3 - 2505.03X2X3 + 6448.92X1X2X3

Table 4: Analysis of variance (ANOVA)

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	3326.16	9	369.574	17.76	0.0000
Residual	312.144	15	20.8096		
Total (Corr.)	3638.31	24			

It can rewrite this equation as follows: rejection = -1886.76 + 5020.79 molecular width + 2267.82 molecular height + 2092.90 molecular length + 0.267792 molecular weight - 5891.10 molecular width x molecular height - 5464.52 molecular width x molecular length - 2505.03 molecular height x molecular length + 6448.92 molecular width x molecular height x molecular length.

From the multiple regression model, it can be observed that rejection will increase in the order of increasing width, height, length, and increasing molecular weight of the organic compounds. The graph comparing the actual (observed) rejection values and predicted rejection values for each organic compound are shown in Figure 1. The result shows a relatively good agreement between the predicted and observed TrOCs removal. Therefore the multiple regression model was considered reliable. The dataset is provided as supplementary data, a 95.0 % confidence interval indicates that very few modelled rejections were out of that interval. Most of the data points focus on the line of perfect fit, indicating that the correlation between rejection efficiency and width, height, length and molecular weight are quite significantly. The goodness of fit of a multiple regression model describes how well the regression model fits the data points. All the indices that exist to evaluate the goodness of fit summarize the discrepancy between the observed values and the predicted values under the regression model. They can only tell how good the model fits with the data used to build the models, not beyond the extent of the data set.

4 Conclusions

From the results obtained using the selected TrOCs, at the experimental conditions used, a multiple linear regression model equation was developed to merge information about interaction of molecular width, molecular height, molecular length and molecular weight to predict rejections of TrOCs during nanofiltration. Mathematical modeling was obtained as follows: Rejection = -1886.76 + 5020.79 molecular width + 2267.82 molecular height + 2092.90 molecular length + 0.267792 molecular weight - 5891.10 molecular width x molecular height - 5464.52 molecular width x molecular length - 2505.03 molecular height x molecular length + 6448.92 molecular width x molecular height x molecular length.

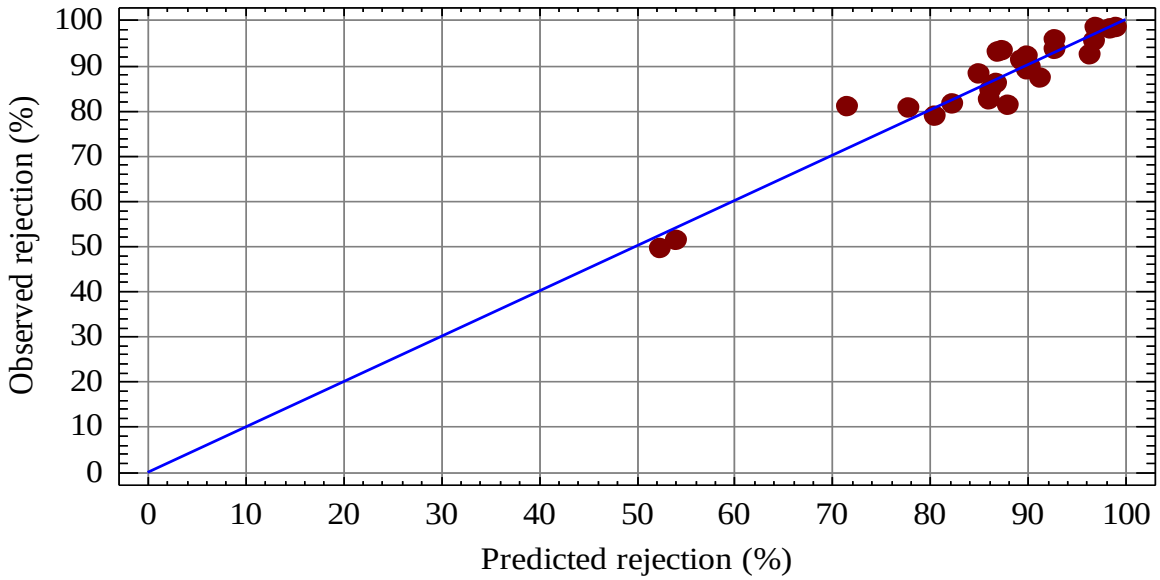

Figure 1: Predicted rejection variables versus observed rejection variables with 95.0 % confidence intervals

This model showed relatively good predictive power for rejection of TrOCs with R^2 = 91.42 %. Based on the model equation, rejection of TrOCs will increase in the order of increasing width, height, length, and increasing molecular weight of the compounds. The multiple linear regression model equation indicated good potential as a simplified modeling tool to predict the rejection of TrOCs during nanofiltration.

Acknowledgements

The author acknowledge the support of the Ministry of Agriculture and Rural Development of Vietnam (MARD) for funding and the University of Wollongong for supporting this research.

References

1. Kasher, R. Membrane-based water treatment technologies: Recent achievements, and new challenges for a chemist. *Bulletin of the Israel Chemical Society*. 2009, 24(24), 10-18.
2. Williams, M., et al., *Endocrine Disrupting Chemicals in the Australian Riverine Environment: A pilot study on estrogenic compounds*. 2007, Land & Water Australia.
3. Li, Z.H. and T. Randak. Residual pharmaceutically active compounds (PhACs) in aquatic environment – status, toxicity and kinetics: a review. *Veterinarni Medicina*. 2009, 52(7), 295-314.
4. Sirivedhin, T. and K.A. Gray. Comparison of the disinfection by-product formation potentials between a wastewater effluent and surface waters. *Water Research*. 2005, 39, 1025-1036.
5. Wintgens, T., et al. The role of membrane processes in municipal wastewater reclamation and reuse. *Desalination*. 2005, 178, 1-11.
6. Bolong, N., et al. A review of the effects of emerging contaminants in wastewater and options for their removal. *Desalination*. 2009, 239(1-3), 229-246.
7. Shon, H.K., S. Vigneswaran, and S.A. Snyder. Effluent Organic Matter (EfOM) in Wastewater: Constituents, Effects, and Treatment. *Critical Reviews in Environmental Science and Technology*. 2006, 36(4), 327-374.
8. Sagne, C., et al. Screening of reverse osmosis membranes for the treatment and reuse of distillery condensates into alcoholic fermentation. *Desalination*. 2008, 219, 335-347.
9. Nghiem, L.D., et al. Estrogenic hormone removal from wastewater using NF/RO membranes. *Journal of Membrane Science*. 2004, 242(1-2), 37-45.
10. Van der Bruggen, B., et al. Influence of molecular size, polarity and charge on the retention of organic molecules by nanofiltration. *Journal of Membrane Science*. 1999, 156, 29-41.
11. Braeken, L., et al. Influence of hydrophobicity on retention in nanofiltration of aqueous solutions containing organic compounds. *Journal of Membrane Science*. 2005, 252, 195-203.
12. Kimura, K., et al. Rejection of organic micropollutants (disinfection by-products, endocrine disrupting compounds, and pharmaceutically active compounds) by NF/RO membranes. *Journal of Membrane Science*. 2003, 227, 113-121.
13. Agenson, K.O., J.I. Oh, and T. Urase. Retention of a wide variety of organic pollutants by different nanofiltration/reverse osmosis membranes: controlling parameters of process. *Journal of Membrane Science*. 2003, 225, 91-103.
14. Van der Bruggen, B., L. Braeken, and C. Vandecasteele. Evaluation of parameters describing flux decline in nanofiltration of aqueous solutions containing organic compounds. *Desalination*. 2002, 147, 281-288.
15. Chen, S.S., et al. Influences of molecular weight, molecular size, flux, and recovery for aromatic pesticide removal by nanofiltration membranes. *Desalination* 2004, 160, 103-111.
16. Xu, P., et al. Rejection of emerging organic micropollutants in nanofiltration-reverse osmosis membrane applications. *Water Environment Research*. 2005, 77(1), 40-48.
17. Cornelissen, E.R., et al. A nanofiltration retention model for trace contaminants in drinking water sources. *Desalination* 2005, 178, 179-192.
18. Kim, S.D., et al. Occurrence and removal of pharmaceuticals and endocrine disruptors in South Korean surface, drinking, and waste waters. *Water Research*. 2007, 41(5), 1013-1021.
19. Verliefde, A.R.D., et al., *Rejection of organic micropollutants by high pressure membranes (NF/RO)*. 2008, Netherlands: Water Management Academic Press.
20. Petrie, A., J.S. Bulman, and J.F. Osborn. Further statistics in dentistry Part 6: Multiple linear regression. *British Dental Journal*. 2002, 193(12), 675-682.
21. Yangali-Quintanilla, V., et al. A QSAR model for predicting rejection of emerging contaminants (pharmaceuticals, endocrine disruptors) by nanofiltration membranes. *Water research*. 2010, 44, 373-384.
22. Um, M.J., et al. Factor analysis and multiple regression between topography and precipitation on Jeju Island, Korea. *Journal of Hydrology*. 2011, 410, 189-203.
23. Nghiem, L.D., A.I. Schäfer, and M. Elimelech. Pharmaceutical Retention Mechanisms by Nanofiltration Membranes. *Environmental Science & Technology*. 2005, 39(19), 7698-7705.
24. Liu, R., J.L. Zhou, and A. Wilding. Simultaneous determination of endocrine disrupting phenolic compounds and steroids in water by solid-phase extraction-gas chromatography-mass spectrometry. *Journal of Chromatography A*. 2004, 1022, 179-189.
25. Sinnott, R.K., *Chemical Engineering Design*. Fourth ed. Vol. 6. 2005, Elsevier Butterworth-Heinemann, Linacre House, Jordan Hill, Oxford OX2 8DP, England. 1038.
26. Zhang, S.Y., et al. Simultaneous quantification of polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), and pharmaceuticals and personal care products (PPCPs) in Mississippi river water, in New Orleans, Louisiana, USA. *Chemosphere*. 2007, 66, 1057-1069.
27. Kiso, Y., et al. Effects of hydrophobicity and molecular size on rejection of aromatic pesticides with nanofiltration membranes. *Journal of Membrane Science*. 2001, 192, 1-10.
28. Schäfer, A.I., I. Akanyeti, and A.J.C. Semião. Micropollutant sorption to membrane polymers: A review of mechanisms for estrogens. *Advances in Colloid and Interface Science*. 2011, 164, 100-117.
29. Nghiem, L.D. and P.J. Coleman. NF/RO filtration of the hydrophobic ionogenic compound triclosan: Transport mechanisms and the influence of membrane fouling. *Separation and Purification Technology*. 2008, 62, 709-716.
30. Hua, W.S., D.B. Zhi, and H. Yu. Adsorption of bisphenol A by polysulphone membrane. *Desalination*. 2010, 253, 22-29.
31. Wells, M.J.M. Log Dow : Key to understanding and regulating wastewater-derived contaminants. *Environmental Chemistry*. 2006, 3(6), 439-449.
32. Alturki, A.A., et al. Combining MBR and NF/RO membrane filtration for the removal of trace organics in indirect potable water reuse applications. *Journal of Membrane Science*. 2010, 365, 206-215.
33. Yangali-Quintanilla, V., et al. Rejection of pharmaceutically active compounds and endocrine disrupting compounds by clean and fouled nanofiltration membranes. *Water Research*. 2009, 43, 2349-2362.
34. Yangali-Quintanilla, V., *Rejection of Emerging Organic Contaminants by Nanofiltration and Reverse Osmosis Membranes: effects of fouling, modelling and water reuse, PhD Thesis*, in *Municipal Water and Infrastructure, UNESCO – IHE Institute for Water Education*. 2010, Delft University of Technology: Delft, the Netherlands.
35. Van der Bruggen, B. and C. Vandecasteele. Removal of pollutants from surface water and groundwater by nanofiltration: overview of possible applications in the drinking water industry. *Environmental Pollution*. 2003, 122(3), 435-445.

Author Profile

Dr. Hai Quang Dang is currently working as lecturer and researcher in the University of Danang, University of Science and Technology. With many years' experience and research in the field of the environment, he has many studies and reports published about the work of the field related to the water quality and treatment, removal of trace organic contaminants, technology of treatment of wastewater and organic wastes.