

Modeling of RO/NF membrane rejections of PhACs and organic compounds: a statistical analysis

V. Yangali-Quintanilla^{1,2}, T.-U. Kim³, M. Kennedy¹, and G. Amy^{1,2}

¹UNESCO-IHE Institute for Water Education, Westvest 7, 2611 AX Delft, The Netherlands

²Delft University of Technology, Stevinweg 1, Delft, The Netherlands

³Pennsylvania State University at Harrisburg, Middletown, PA 17057, USA

Received: 4 January 2008 – Published in Drink. Water Eng. Sci. Discuss.: 8 January 2008

Revised: 27 May 2008 – Accepted: 11 July 2008 – Published: 17 July 2008

Abstract. Rejections of pharmaceutical compounds (Ibuprofen, Diclofenac, Clofibrac acid, Naproxen, Primidone, Phenacetin) and organic compounds (Dichloroacetic acid, Trichloroacetic acid, Chloroform, Bromoform, Trichloroethene, Perchloroethene, Carbontetrachloride, Carbontetrabromide) by NF (Filmtec, Saehan) and RO (Filmtec, Saehan, Toray, Koch) membranes were studied. Chloroform presented the lowest rejection due to small molar volume, equivalent width and length. Diclofenac and Primidone showed high rejections related to high molar volume and length. Dichloroacetic acid and Trichloroacetic acid presented good rejections caused by charge exclusion instead of steric hindrance mechanism influencing rejection. Bromoform and Trichloroethene showed low rejections due to small length and equivalent width. Carbontetrabromide, Perchloroethene and Carbontetrachloride with higher equivalent width than BF and TCE presented better rejections. A qualitative analysis of variables using Principal Component Analysis was successfully implemented for reduction of physical-chemical compound properties that influence membrane rejection of PhACs and organic compounds. Properties such as dipole moment, molar volume, hydrophobicity/hydrophilicity, molecular length and equivalent width were found to be important descriptors for simulation of membrane rejection. For membranes used in the experiments, we may conclude that charge repulsion was an important mechanism of rejection for ionic compounds. After analysis with Multiple Linear Regression, we also may conclude that membrane rejection of neutral compounds was well predicted by molar volume, length, equivalent width, hydrophobicity/hydrophilicity and dipole moment. Molecular weight was a poor descriptor variable for rejection modelling. We were able to provide acceptable statistical significance for important results.

1 Introduction

The presence of pharmaceutically activated compounds (PhACs) and endocrine disrupters compounds (EDCs) in surface waters has been reported, detailed and quantified in many studies (Ternes, 1998; Hirsch et al., 1999; Heberer, 2002; Kolpin et al., 2002). Although pharmaceuticals are generally taken up by the human organism, a significant amount of the original substance is often excreted by the human body, thus entering raw sewage and eventually ending up in the aquatic environment. The fate of pharmaceuticals in the environment has raised the interest of scien-

tists because the accumulation of them may result in environmentally significant concentrations with unknown effects. It has been studied that conventional water treatment presented limitations in removing PhACs and EDCs (Vieno et al., 2006; Adams et al., 2002). In that sense, many studies have investigated the removal of micropollutants i.e. PhACs, EDCs, by membrane treatment (NF, RO) and their separation mechanisms such as size/steric exclusion, hydrophobic adsorption, partition and electrostatic repulsion (Kiso et al., 2001a, b, 2002; Schäfer et al., 2003; Nghiem et al., 2004; Kimura et al., 2003, 2004; Kim et al., 2005). Characteristics such as MWCO, porosity, membrane morphology, charge, and hydrophobicity of the membrane influence rejection of compounds (Schaep and Vandecasteele, 2001; Childress and Elimelech, 2000); compound properties such as



Correspondence to:
V. Yangali-Quintanilla
(v.yangaliquintanilla@unesco-ihe.org)

molecular weight, molecular size, charge, dipole moment and hydrophobicity can be used as predictors of rejection in applications of membrane water treatment (Ozaki and Li, 2002; Van der Bruggen et al., 2000; Kiso et al., 2001a; Van der Bruggen et al., 1999). However, there are disagreements defining the importance of some descriptors, in that sense our objective is to investigate a group of compound properties and membrane characteristics with a statistical approach in order to model membrane rejection.

2 Background of statistics

2.1 Principal component analysis

A summary of principal component analysis is presented by Landau and Everitt (2004). Principal Component Analysis (PCA) is essentially a method of data reduction that aims to produce a small number of derived variables that can be used in place of the larger number of original variables to simplify subsequent analysis of the data. The principal component variables y_1, y_2, \dots, y_q are defined to be linear combinations of the original variables x_1, x_2, \dots, x_q that are uncorrelated and account for maximal proportions of the variation in the original data, i.e., y_1 accounts for the maximum amount of the variance among all possible linear combinations of x_1, \dots, x_q , y_2 accounts for the maximum variance subject to being uncorrelated with y_1 and so on. Explicitly, the principal component variables are obtained from x_1, \dots, x_q as follows:

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1q}x_q \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2q}x_q \\ &\vdots \\ y_q &= a_{q1}x_1 + a_{q2}x_2 + \dots + a_{qq}x_q \end{aligned} \quad (1)$$

Where the coefficients a_{ij} ($i=1, \dots, q, j=1, \dots, q$) are chosen so that the required maximal variance and uncorrelated conditions hold. Since the variances of the principal components variables could be increased without limit, simply by increasing the coefficients that define them, a restriction must be placed on these coefficients. The constraint usually applied is that the sum of squares of the coefficients is one in that way the total variance of all the components is equal to the total variance of all the observed variables. It is often convenient to rescale the coefficients in order that their sum of squares is equal to the variance of the component they define. In the case of components derived from the correlation matrix of the data, these rescaled coefficients give the correlations between the components and the original variables. Those values are often presented as the result of a principal components analysis. The coefficients defining the principal components are given by what are known as the eigenvectors of the correlation matrix, \mathbf{R} . PCA transforms the multivariate set into a set of artificial components (principal components) based on the symmetric correlation matrix. The procedure

of PCA starts with assigning eigenvalues to each component for transforming a set of multi variables into a set of components.

Principal component analysis is a multivariate technique that transforms a set of related (correlated) variables into a set of unrelated (uncorrelated) variables that account for decreasing proportions of the variation of the original observations. The aim is to reduce the complexity of the data by decreasing the number of variables that need to be considered. When the derived variables (the principal components) account for a large proportion of the total variance of the observed variables, they can be used both to provide a convenient summary of the data and to simplify subsequent analyses (Ho, 2006).

Components produced in the initial extraction phase are often difficult to interpret. This is because the procedure in this phase ignores the possibility that variables identified to load on or represent components may already have high loadings (correlations) with previous components extracted. This may result in significant cross-loadings in which many components are correlated with many variables. This makes interpretation of each component difficult, because different components are represented by the same variables. The rotation phase serves to “sharpen” the components by identifying those variables that load on one component and not on another. The ultimate effect of the rotation phase is to achieve a simpler, theoretically more meaningful component pattern ((Ho, 2006).

2.2 Multiple linear regression

Multiple linear regression 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 (Landau and Everitt, 2004). The multiple regression model for a response variable, y , with observed values, y_1, y_2, \dots, y_n (where n is the sample size) and q explanatory variables, x_1, x_2, \dots, x_q with observed values, $x_{1i}, x_{2i}, \dots, x_{qi}$ for $i=1, \dots, n$, is:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_q x_{qi} + \varepsilon_i \quad (2)$$

The regression coefficients, $\beta_0, \beta_1, \dots, \beta_q$, are generally estimated by least squares. The term ε_i is the residual or error for individual i and represents the deviation of the observed value of the response for this individual from that expected by the model. These error terms are assumed to have a normal distribution with variance σ^2 . The fit of a multiple regression model can be judged with calculation of the multiple correlation coefficient, R , defined as the correlation between the observed values of the response variable and the values predicted by the model. The value of R^2 gives the proportion of the variability of the response variable accounted for

Table 1. Physical-chemical properties of compounds.

| Compound | Abbr. | MW (g/mol) | Dipole (Debye) | Molar Vol. (cm ³ /mol) | log K_{ow} | pK_a | Mole. length (nm) | Mole. width (nm) | Mole. depth (nm) | eqwidth (nm) | Class.* |
|-----------------------|-------|---------------|-------------------|--------------------------------------|--------------|--------|----------------------|---------------------|---------------------|-----------------|---------|
| Dichloro-acetic acid | DCAA | 129 | 2.40 | 82 | 0.92 | 1.48 | 0.7 | 0.69 | 0.52 | 0.60 | HL-ion |
| Trichloro-acetic acid | TCAA | 163 | 1.50 | 100 | 1.33 | 0.70 | 0.89 | 0.67 | 0.41 | 0.52 | HL-ion |
| Ibuprofen | IBP | 206 | 1.29 | 130 | 3.97 | 4.91 | 1.31 | 0.77 | 0.64 | 0.70 | HP-ion |
| Diclofenac | DCF | 296 | 1.48 | 182 | 4.51 | 4.15 | 1.13 | 0.91 | 0.45 | 0.64 | HP-ion |
| Clofibric acid | CFA | 214 | 0.79 | 155 | 2.57 | 3.64 | 0.95 | 0.66 | 0.41 | 0.52 | HP-ion |
| Naproxen | NPX | 230 | 2.74 | 167 | 3.18 | 4.15 | 1.26 | 0.69 | 0.54 | 0.61 | HP-ion |
| Chloroform | CF | 119 | 1.12 | 80 | 1.97 | N/A | 0.53 | 0.5 | 0.35 | 0.42 | HL-neu |
| Primidone | PMD | 218 | 4.31 | 164 | 0.91 | N/A | 0.97 | 0.88 | 0.48 | 0.65 | HL-neu |
| Phenacetin | PAC | 179 | 2.40 | 146 | 1.58 | N/A | 1.39 | 0.68 | 0.42 | 0.53 | HL-neu |
| Bromoform | BF | 253 | 1.00 | 87 | 2.40 | N/A | 0.69 | 0.65 | 0.48 | 0.56 | HP-neu |
| Trichloro-ethene | TCE | 131 | 0.95 | 90 | 2.29 | N/A | 0.78 | 0.66 | 0.36 | 0.49 | HP-neu |
| Perchloro-ethene | PCE | 165 | 0.11 | 102 | 3.40 | N/A | 0.78 | 0.77 | 0.45 | 0.59 | HP-neu |
| Carbontetra-chloride | CT | 154 | 0.30 | 96 | 2.83 | N/A | 0.64 | 0.64 | 0.57 | 0.60 | HP-neu |
| Carbontetra-bromide | CTB | 332 | 0.01 | 112 | 3.42 | N/A | 0.69 | 0.63 | 0.65 | 0.64 | HP-neu |

* HL, hydrophilic; HP, hydrophobic; ion, ionic; neu, neutral.

by the explanatory variables. Analysis of variance (ANOVA) will provide an F-test of the null hypothesis that each of $\beta_0, \beta_1, \dots, \beta_q$, is equal to zero, or in other words that R^2 is zero.

3 Materials and methods

3.1 Compounds, membranes and properties

A list of selected compounds is presented in Table 1; this table also shows physical-chemical estimations of compound properties, such as: molecular weight (MW), dipole moment, water-octanol partition coefficient (log K_{ow}), acid dissociation constant (pK_a), molar volume (MV), length, width and depth. Compounds were classified in ionic and neutral considering the acid dissociation constant, compounds with a pK_a greater than 7 or not available (N/A) pK_a were selected as neutral, otherwise were ionic, this classification is related to acidity of waters, our ionic compounds dissociate at pH 8. In our case the pK_a of all our compounds is less than 7; though the ionic species were dissociated at pH 8. Compounds were classified as hydrophilic (HL) when log K_{ow} was less than 2; and as hydrophobic (HP) when log $K_{ow} > 2$, there are some defined ranges of log K_{ow} related to lipophilicity (hydrophobicity) of compounds (Conell, 1989). Values of pK_a were calculated with online software Sparc (Sparc, 2006). Octanol-water partition coefficients expressed as log K_{ow} correspond to neutral molecule forms and were calculated with software Kowwin (Kowwin, 2006). We acknowledge that hydrophobicity/hydrophilicity in terms of HP and HL is not different of log K_{ow} , it was only a definition for the sake of a simplified classification and subsequent analysis. It is important to mention that some available experimental database values of pK_a and log K_{ow} were used. Calculation of dipole moments was carried out with commercial software Chem3D Ultra 7 (Chemof-

ice, 2002). Molar Volume (MV) of compounds was calculated as quotient of molecular weight (g/mol) and liquid density of compound (g/cm³), unknown liquid density values were calculated by Grain's Method (Lyman et al., 1990). Molecular size variables, represented by length, width and depth, were calculated with software Molecular Modeling Pro (ChemSW, 2006). Afterwards, an equivalent width was defined as (width \times depth)^{0.5} to represent the equivalent width of molecules.

Anhydrous sodium sulfate and sodium chloride were obtained from Fisher Scientific. Potassium hydroxide, potassium chloride, and a buffer solution based on potassium phosphate, used for water quality maintenance, were also purchased from Fisher Scientific. Sulfuric acid was purchased from Mallinckrodt Chemicals (USA). Methyl tert-Butyl Ether (MtBE), used for GC-ECD analysis, and HPLC grade methanol were obtained from Fisher Scientific. BF, PCE, CT, CTB, and DCAA were obtained from Aldrich (USA). CF, TCE, and TCAA were purchased from Fisher Scientific (USA). IBP and PMD were purchased from Sigma, DCF from TCI America, PAC from Fluka. CFA, CBM and NPX were ordered from Aldrich.

The four FilmTec membranes tested were BW-400 (RO), LE-440 (RO), XLE-440 (LPRO), and NF-90. Additional tested membranes were Saehan RE-BLR (RO), Saehan NE-90 (NF), Toray UTC-70UB (LPRO), and Koch CTA (RO). These membranes were composed of MPD (Metaphenylene diamine (Benzene-1,3-diamine)) and TMC (Trimesoyl chloride (1,3,5-Benzenetricarbonyl trichloride)) except Koch CTA (cellulose triacetate). Membrane characteristics were determined by different methods. Membrane hydrophobicity was characterized by contact angle measurement using a goniometer (Model 100-00, Ramé-hart, Inc., Surface Science Instrument, Landing, New Jersey). A stirred cell filtration unit (Amicon) was used to test disk specimens of membranes

Table 2. Characteristics of membranes.

| Membrane | BW-400 | LE-440 | XLE-440 | NF-90 | RE-BLR | NE-90 | UTC-70 | CTA |
|--------------------------------------|---------|---------|-----------|---------|---------|---------|-----------|----------|
| Company | FilmTec | FilmTec | FilmTec | FilmTec | Saehan | Saehan | Toray | Koch |
| Group (type)* | PA (RO) | PA (RO) | PA (LPRO) | PA (NF) | PA (RO) | PA (NF) | PA (LPRO) | CTA (RO) |
| MWCO | n/a | n/a | n/a | 200 | n/a | 200 | n/a | n/a |
| NaCl Rej. (%) | 98 | 98 | 98 | 90 | 99.5 | 90 | 99.6 | 98 |
| Contact Angle (°) | 57 | 42 | 40 | 60 | 47 | 52 | 54 | 47 |
| Zeta Potential (mV), pH 8, 10 mM KCl | -4.5 | -23 | -19 | -31 | -21 | -24 | -15 | -19 |
| J_0 (L/m ² hr) | 18 | 18 | 16 | 26 | 15 | 22 | 14 | 18 |
| Oper. Pressure (kPa) | 620 | 550 | 410 | 280 | 480 | 240 | 340 | 550 |
| PWP (L/m ² day kPa) | 0.68 | 0.77 | 0.92 | 2.23 | 0.77 | 2.17 | 0.99 | 0.79 |

* PA polyamide; CTA cellulose triacetate

in a dead-end mode for determining the molecular weight cutoff (MWCO) of membranes using polyethylene glycols (PEGs), that was done only for the NF membrane (NF-90 and NE-90). The pure water permeability (PWP (L/day-m²-kPa)) of each membrane was measured, reflecting the capacity for water to pass through the membrane normalized by transmembrane pressure. For study of electrokinetic properties, an electrophoresis method for zeta potential measurement has been used with ELS-8000 (Otsuka Electronics, Japan); in the electrophoresis cell consisting of membrane and quartz cells, asymmetric electro-osmotic flow occurs due to the accumulation of ions on the membrane surface during the electrophoresis method (Shim et al., 2002). A summary of membrane characteristics used for PCA, is shown in Table 2.

3.2 Apparatus, analysis, equipment setup and experimental conditions

A membrane filtration unit SEPA cell of flat-sheet type (Osmonics, USA) was used for cross-flow tests. The system was composed of a polymeric membrane, a membrane holder, pumps with a gear type pump head, needle valves (for the feed, retentate, and permeate streams), pressure sensors, flow meters and a reservoir of 120 L. Either varying the pump head speed or controlling the needle valve in the retentate stream controlled the feed flow rate, the corresponding cross-flow velocity, and the trans-membrane pressure. The feed water temperature was maintained in the range (20–25°C). The total membrane surface area in the test cell was approximately 135.8 cm² and the total cross flow area in the test cell was approximately 1.45 cm². All permeates and concentrates were not recirculated but wasted in these experiments. Changes in pressures permitted variation of the J_0/k ratio, a hydrodynamic condition embodying initial flux (J_0) toward the membrane and back-diffusion (k , mass transfer coefficient, a function of diffusion coefficient and cross-flow velocity) away from the membrane (Cho et al., 2000). The system recovery ratio was 10%. Feed concentrations of compounds were 0.1 mg/L. Experiments were carried out at a pH

of 8 and conductivity of 300 μ S/cm conditioned with KCl. The membrane was pre-compacted with deionized water for two hours before starting filtration experiments, experiments were conducted for 48 h. The J_0/k ratio was 1 for all experiments reported in this publication. Information about pressures and permeate fluxes used for the experiments are presented in Table 2. It is important to mention that the statistical analysis only corresponds to a unique hydrodynamic regime in terms of $J_0/k=1$. Kimura et al. (2003) established an experimental protocol for filtration of some hydrophobic compounds, demonstrating that a “quasi-saturation” of the tested membrane was reached after about 20 h of operation using a feed solution of 100 μ g/L concentration, suggesting that for low concentration feed, the filtration time should be extended and a large volume of feed should be circulated in order to achieve sufficient membrane saturation (Kimura et al., 2003). In that sense, in our case, 48 h produced adequate membrane saturation. The issue of fouled membranes was not addressed during this study, although we think that our approach may be applicable to fouled membranes using the same descriptors for compound properties but with changed membrane characteristics.

EPA sample vials (40 mL) with a screw cap lined with Teflon were used for sample collection and extraction. Autosampler vials used were 2.0 mL amber glass vials with a crimp cap and a Teflon-faced seal. Disposable Pasteur pipettes (9 inch) were used to transfer extracted samples. Micropipettes (10–100 mL, 100–1000 mL, and 1–5 mL) with disposable tips were purchased from Fisher Scientific (USA). A mini-vortexer (VWR Scientific, USA) was used as an orbital mixer. A Brinkmann bottle top dispenser was used for adding solvent. An analytical balance (Mettler Toledo AT201) was capable of weighing to 0.01 mg. A diazomethane reaction chamber was used for a step of HAAs analysis. The GC used was an HP 6890 series GC system with a micro electron capture detector, an HP7683 autoinjector, an autosampler tray module, and an HP PC. A DB-1 capillary column coated with dimethylpolysiloxane (30 m \times 0.25 mm \times 1 mm) was employed.

Table 3. Membrane rejection results after 48 h.

| Membrane | Compounds and rejections (%) | | | | | |
|----------|------------------------------|-----|----|----|-----|-----|
| | CF | TCE | BF | CT | PCE | CTB |
| LE-440 | 3 | 2 | 7 | 69 | 76 | 89 |
| NF-90 | 0 | 3 | 0 | 35 | 39 | 70 |
| BW-400 | 4 | 0 | 16 | 78 | 71 | 96 |
| XLE-440 | 10 | 3 | 16 | 69 | 71 | 83 |
| RE_BLR | 34 | 33 | 50 | 90 | 87 | 99 |
| NE90 | 1 | 0 | 17 | 52 | 57 | 70 |
| UTC-70 | 9 | 6 | 34 | 84 | 57 | 79 |
| CTA | 27 | 6 | 2 | 13 | 2 | N/A |

| Membrane | Compounds and rejections (%) | | | | | | | |
|----------|------------------------------|------|-----|-----|-----|-----|-----|-----|
| | DCAA | TCAA | IBP | DCF | CFA | NPX | PMD | PAC |
| LE-440 | 77 | 95 | 91 | 91 | 87 | 75 | 83 | 56 |
| NF-90 | 89 | 87 | 86 | 90 | 86 | 89 | 82 | 45 |

3.3 Statistical analysis

A demanding evaluation of our study may question the statistical robustness of our statistical analysis because we have used limited amount of data (64 cases) representing rejections of 14 compounds. However this study was conducted to show that the methodology was correct and statistics could be applied; we think that future work may allow inclusion of more data or use of new data that will produce results similar to our findings. Moreover, we were able to provide acceptable statistical significance for our more important results. Statistical analysis was used to determine which physical-chemical properties of compounds contributed most significantly to membrane rejection. The statistical software package SPSS 14 (SPSS, 2005) was used for analysis. The performed statistical methods included principal component analysis (PCA) and multiple linear regression (MLR). The first step for data analysis was carried out with PCA to reduce the number of variables (physical-chemical properties) that are representative for a group of compounds. Components produced in the initial extraction phase are often difficult to interpret. This may result in significant cross-loadings in which many components are correlated with many variables. This makes interpretation of each component difficult, because different components are represented by the same variables. The rotation phase serves to “sharpen” the components by identifying those variables that load on one component and not on another. Orthogonal rotation with varimax method was used for separation of components, it assumes that the components are independent; the rotation process maintains the reference axes of the components at 90°. The subsequent analysis of the reduced number of variables was performed with MLR. The selected method for linear regression was *stepwise*. Stepwise regression is the most sophisticated of statistical methods for multiple linear regression.

Each variable is entered in sequence and its value assessed. If adding the variable contributes to the model then it is retained, but all other variables in the model are then re-tested to see if they are still contributing to the success of the model. If they no longer contribute significantly they are removed. Therefore, the method should ensure that we end up with the smallest possible set of predictor variables included in the model.

4 Results and discussion

4.1 Rejection of compounds by membranes

Table 3 presents results of rejection for compounds corresponding to the membranes with which the experiments were carried out. All results are from samples taken after 48 hours of experimental running. Two membrane types were selected for experiments with all compounds: LE-440 (RO) and NF-90 (NF).

4.2 PCA for compounds and membranes

Our hypothesis is that a reduced number of variables will appropriately explain rejection qualitatively based on physical-chemical properties of compounds and membrane characteristics. In order to validate it, we formulated the following questions: i) which physical-chemical properties significantly represent a compound in a smaller set of variables? ii) which compounds depending on their properties were better rejected? iii) which membrane characteristics can describe rejection better? Variables considered for PCA were molecular weight (MW), dipole moment (dipole), molar volume (MV), water-octanol partition coefficient expressed as $\log K_{ow}$, hydrophobicity/hydrophilicity (HP) a dummy variable that takes the value of 0 when the compound

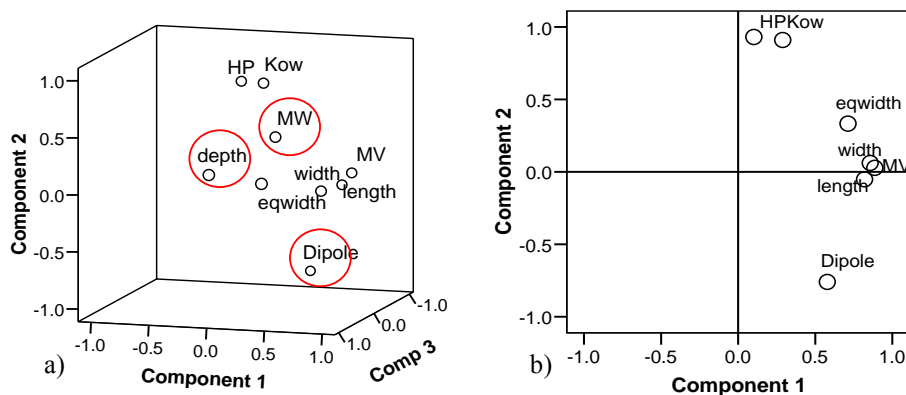


Figure 1. (a) Rotated component loadings for all variables; (b) Rotated component loadings for reduced variables.

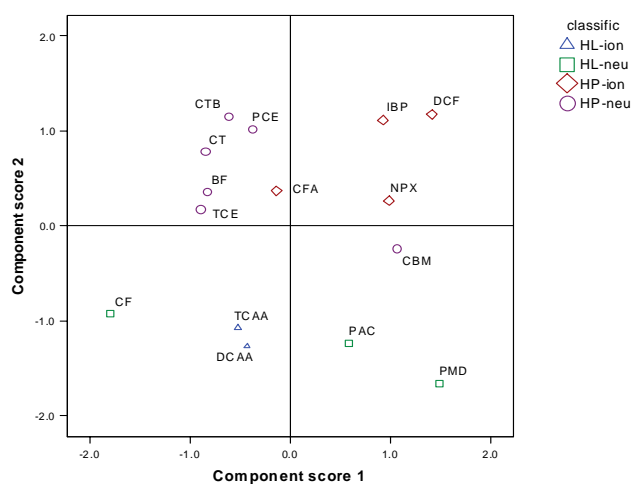


Figure 2. Component scores for compounds.

is hydrophilic (HL) and 1 when the compound is hydrophobic (HP), molecular length (length), molecular width (width), molecular depth (depth) and equivalent width (eqwidth). After the first analysis considering all mentioned variables, three principal components accounted for 82.7% of the total variance. Figure 1a shows the rotated component loadings graph of all variables in three components, we recognize that MW and depth presented component cross-loadings. A second run of PCA, without considering variables MW and depth, resulted in Fig. 2b, in this case component 1 represents size with variables width, length, eqwidth and MV; and component 2 defines clearly hydrophobicity/hydrophilicity, with variables HP/HL and $\log K_{ow}$. The effect of the variable dipole is not well defined (see Fig. 1a and b), we can observe cross-loadings for dipole; nevertheless, we decided to include dipole moment as variable for our statistical modelling with the hypothesis that dipole may influence rejection. According to Table 4, the first and second principal components (scaled eigenvectors), explained the largest part of the total variance, they have eigenvalues of 3.2 and 2.2,

respectively; this accounts to 45.6% and 31.8% of the total variance. The third principal component has a variance of 0.7 and accounts for a further 9.9% of the variance.

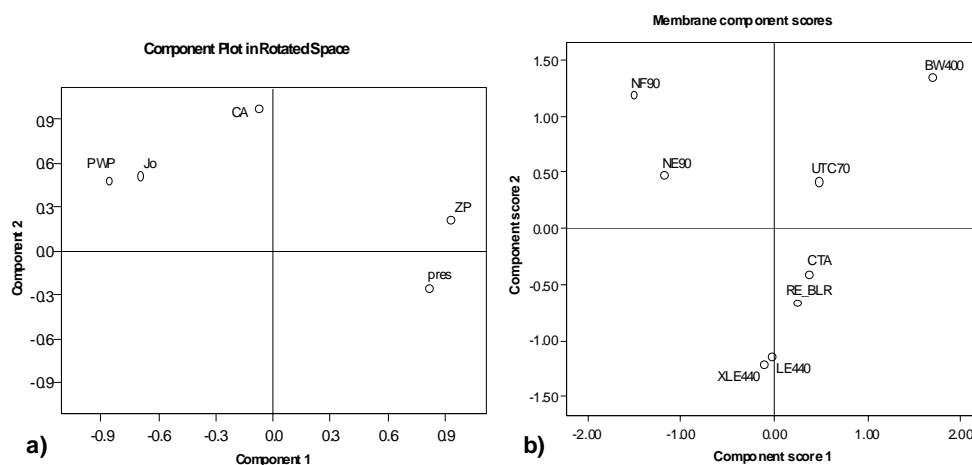
The cumulative % column shows how much of the total variance was accounted for by the components with eigenvalues greater than 1. The first two principal components accounted for 77.4% of the total variance. Thus, we can answer question i), we can reduce our first nine variables to seven variables that can be explained by the first two components of Table 4. Subsequently we answer question ii) which compounds depending on their properties were better rejected? We present Fig. 2, a graph of the first and second components scores for each compound. That plot helps to visualize the rejection patterns of the fifteen compounds. Scores on the x-axis (component 1) may indicate the overall level of rejection, while scores on the y-axis indicate possibly differences between compound properties. Clustering of compounds classified according to groups is a first notice from the graph. CF (chloroform) clearly appeared outside from the other compounds, presenting the lowest rejection (see Table 3) due to small molar volume, equivalent width and length. DCF (diclofenac) and PMD (primidone) showed high rejections related to high molar volume. CBM (carbamazepine) appears outside of the clustered group of hydrophobic-neutral compounds due to his high molar volume and dipole moment when compared to the rest of the compounds in his group. Dipole moment also influenced that CFA (clofibric acid) appeared outside of its group. Also notable was the clustering of DCAA (dichloroacetic acid) and TCAA (trichloroacetic acid), both compounds presented good rejections (Table 3); however their charge prevailed the mechanism of electrostatic instead of size exclusion mechanism influencing rejection. BF (bromoform) and TCE (trichloroethene) had low rejections due to small length and TCE and lower values of dipole moment than BF and TCE presented better rejections.

Table 4. Principal Components and total variance.

| Comp. | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|-------|---------------------|---------------|--------------|-------------------------------------|---------------|--------------|-----------------------------------|---------------|--------------|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 3.192 | 45.607 | 45.607 | 3.192 | 45.607 | 45.607 | 3.115 | 44.500 | 44.500 |
| 2 | 2.228 | 31.827 | 77.434 | 2.228 | 31.827 | 77.434 | 2.305 | 32.934 | 77.434 |
| 3 | 0.697 | 9.959 | 87.393 | | | | | | |
| 4 | 0.417 | 5.950 | 93.343 | | | | | | |
| 5 | 0.215 | 3.075 | 96.417 | | | | | | |
| 6 | 0.160 | 2.292 | 98.710 | | | | | | |
| 7 | 0.090 | 1.290 | 100.000 | | | | | | |

Table 5. Results of multiple linear regressions.

| Case | R^2 | Std. Error of Estimate | F | Sig. | Equation for rejection |
|-----------------|-------|------------------------|-------|-------|--|
| Ionic LE440 48h | 0.628 | 5.58 | 6.75 | 0.060 | 101.104-8.885Dipole |
| Ionic NF90 48h | 0.435 | 1.45 | 3.08 | 0.154 | 85.184+1.559Dipole |
| Neutr LE440 48h | 0.952 | 10.14 | 33.07 | 0.001 | -118.889+2.503MV-29.830Dipole-86.339length |
| Neutr NF90 48h | 0.868 | 12.56 | 19.74 | 0.002 | -163.619-30.423HP+386.160eqwidth |

**Figure 3.** (a) Component loadings for membrane variables; (b) Component scores for membranes.

Results of PCA for membrane characteristics presented in Fig. 3a suggest that no further reduction of variables is suitable. Therefore, the answer of question iii) is not possible considering only the set of variables we have. In general nanofiltration membranes (NF-90 and NE90) presented low membrane rejection when compared to reverse osmosis membranes, excluding CTA. Figure 3b present scores for each membrane. According to this, LE-440 performed better than XLE-440; NE 90 better than NF-90, and BW-400 performed better than UTC-70 and RE BLR, in this order BW-400>UTC-70>RE BLR. Although Fig. 3b describes well the

rejection performances of membranes, the information provided must be analyzed and disregard misleading results, it was the case for CTA (compare rejections Table 3), an explanation of it is that CTA is a cellulose triacetate membrane, which means they have differences with the polyamide membranes; we did not consider this variable (composition) for analysis.

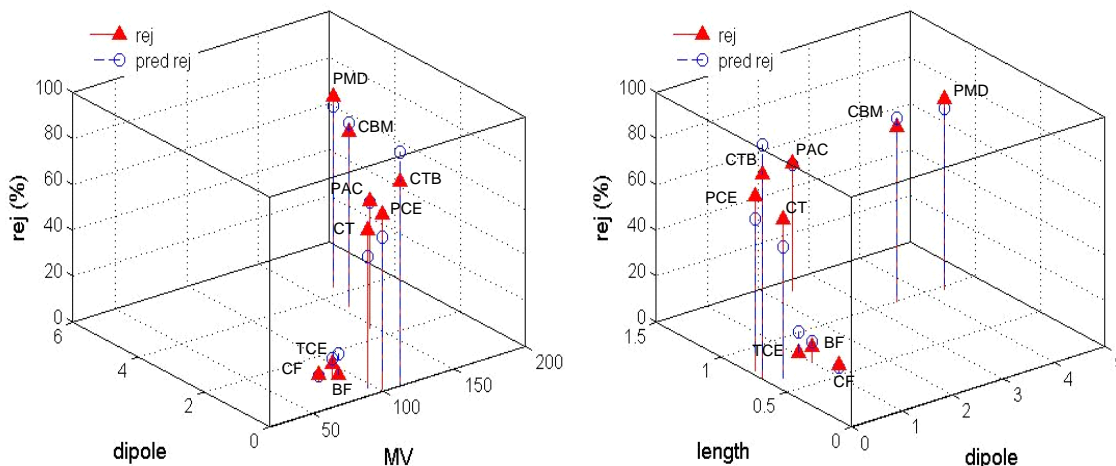


Figure 4. LE-440 membrane rejection modeling of neutral compounds.

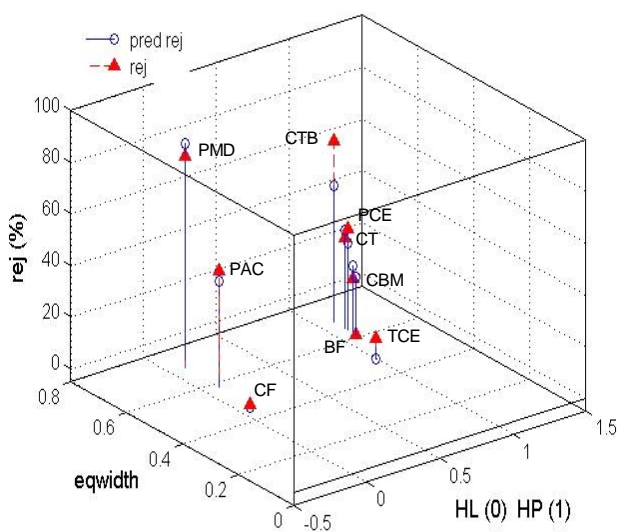


Figure 5. NF-90 membrane rejection modeling of neutral compounds.

4.3 Multiple linear regression

MLR analysis of rejections was separated for ionic and neutral compounds, and for two membranes (LE-440 and NF-90). Thus, MLR was performed for these cases: 1) ionic LE440 48 h; 2) ionic NF90 48 h; 3) neutral LE440 48 h; 4) neutral NF90 48 h. Rejection can be defined as a linear function with compound properties as variables. In general, we may assume that rejection can be described by the following linear equation.

$$\text{rej} = B_0 + B_1 \text{MV} + B_2 \text{HP} + B_3 K_{ow} + B_4 \text{Dipole} + B_5 \text{length} + B_6 \text{eqwidth} \quad (3)$$

Inclusion of all variables will give a R^2 value approaching 1. However the optimum output implies having the less number of variables in the equation that may explain or predict rejection with appropriate fit and significance. Thus, stepwise regression will achieve this by adding and removing variables until an acceptable statistical significance would possibly be reached. The summary of results is shown in Table 5. We can notice that our R^2 values varied between 0.435 and 0.952. Ionic compounds were less favoured in prediction and significance. The rejection prediction of ionic compounds by membrane NF-90 presented the lowest confidence (85%) and R^2 value (0.435). However, R^2 value (0.628) and significance (96%) improved for ionic compounds and LE-440 membrane. Thus, we may use the models and predictors under certain restrictions when referring to ionic compounds. The effect of charge repulsion between the membrane and those charged compounds was evident. On the other hand, neutral compounds showed acceptable R^2 values and good levels of confidence (>95%). An important observation was that MW was disregarded as predictor. According to Table 5, it seems that dipole moment may predict differences in rejection prediction for ionic compounds. The influence of molar volume (MV) was more notable for neutral compounds. A combination of MV, dipole moment and length was able to predict rejection of neutral compounds by LE-440 membrane (Fig. 4). The rejection prediction of neutral compounds for NF-90 membrane involved variables of equivalent width and hydrophobicity/hydrophilicity (Fig. 5). It seems that equivalent width replaced predictors MV and length as we compare NF and RO membranes.

5 Conclusions

A qualitative analysis of variables using principal component analysis was successfully implemented for reduction of physical-chemical compound properties that influence

membrane rejection of PhACs and organic compounds. Properties of dipole moment, molar volume, hydrophobicity/hydrophilicity, molecular length and equivalent width were found to be important descriptors for prediction of membrane rejection. For membranes used in the experiments we may conclude that charge repulsion was an important mechanism of rejection for ionic compounds. Molecular weight was found to be a poor variable descriptor for rejection simulation. We could model membrane rejection of neutral compounds with multiple linear regression; real rejection results were well represented by dipole moment, molar volume, length, equivalent width and hydrophobicity/hydrophilicity of compounds. A demanding evaluation of our study may question the statistical robustness of our statistical analysis because we have used limited amount of data (64 cases) representing rejections of 14 compounds. However this study was conducted to show that the methodology was correct and statistics could be applied; we think that future work may allow inclusion of more data or use of new data that will produce results similar to our findings. We were able to provide acceptable statistical significance for important results.

Edited by: J. Verberk

References

- Adams, C., Wang, Y., Loftin, K., and Meyer, M.: Removal of Antibiotics from Surface Water and Distilled Water in Conventional Water Treatment Processes, *J. Env. Eng.*, 128, 253–259, 2002.
- Childress, A. E. and Elimelech, M.: Relating nanofiltration membrane performance to membrane charge (electrokinetic) characteristics, *Environ. Sci. Technol.*, 34, 3710–3716, 2000.
- Cho, J., Amy, G., and Pellegrino, J.: Membrane filtration of natural organic matter: comparison of flux decline, NOM rejection, and foulants during filtration with three UF membranes, *Desalination*, 127, 283–298, 2000.
- Conell, D. W.: Bioaccumulation of Xenobiotic Compounds, CRC Press, 219 pp., 75–110, 1989.
- Heberer, T.: Occurrence, fate, and removal of pharmaceutical residues in the aquatic environment: a review of recent research data, *Toxicol. Lett.*, 131, 5–17, 2002.
- Hirsch, R., Ternes, T., Haberer, K., and Kratz, K.-L.: Occurrence of antibiotics in the aquatic environment, *Sci. Total Environ.*, 225, 109–118, 1999.
- Ho, R.: Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS, Chapman & Hall/CRC Press, 393 pp., 2006.
- Kim, T.-U., Amy, G., and Drewes, J.: Rejection of trace organic compounds by high-pressure membranes, *Water Sci. Tech.*, 51, 335–344, 2005.
- Kimura, K., Amy, G., Drewes, J., and Watanabe, Y.: Adsorption of hydrophobic compounds onto NF/RO membranes: an artifact leading to overestimation of rejection, *J. Mem. Sci.*, 221, 89–101, 2003.
- Kimura, K., Toshima, S., Amy, G., and Watanabe, Y.: Rejection of neutral endocrine disrupting compounds (EDCs) and pharmaceutical active compounds (PhACs) by RO membranes, *J. Mem. Sci.*, 245, 71–78, 2004.
- Kiso, Y., Kon, T., Kitao, T., and Nishimura, K.: Rejection properties of alkyl phthalates with nanofiltration membranes, *J. Mem. Sci.*, 182, 205–214, 2001a.
- Kiso, Y., Sugiura, T., Kitao, T., and Nishimura, K.: Effects of hydrophobicity and molecular size on rejection of aromatic pesticides with nanofiltration membranes, *J. Mem. Sci.*, 192, 1–10, 2001b.
- Kiso, Y., Mizuno, A., Othman, R., Jung, Y. J., Kumano, A., and Arijji, A.: Rejection properties of pesticides with a hollow fiber NF membrane (HNF-1), *Desalination*, 143, 147–157, 2002.
- Kolpin, D. W., Furlong, E. T., Meyer, M. T., Thurman, E. M., Zugg, S. D., Barber, L. B., and Buxton, H. T.: Pharmaceuticals, Hormones, and other organic wastewater contaminants in U.S. Streams, 1999–2000: A National Reconnaissance, *Environ. Sci. Technol.*, 36, 1202–1211, 2002.
- Landau, S. and Everitt, B.: A Handbook of Statistical Analysis using SPSS, Chapman & Hall/CRC Press, 2004.
- Lyman, W. J., Reehl, W. F., and Rosenblatt, D. H.: Handbook of chemical property estimation methods, American Chemical Society, McGraw-Hill Inc., 17.1–17.25, 1990.
- Nghiem, L. D., Schäfer, A. I., and Elimelech, M.: Removal of natural hormones by nanofiltration membranes: measurement, modelling, and mechanisms, *Environ. Sci. Technol.*, 38, 1888–1896, 2004.
- Ozaki, H. and Li, H.: Rejection of organic compounds by ultra-low pressure reverse osmosis membrane, *Water Res.*, 36, 123–130, 2002.
- Schaep, J. and Vandecasteele, C.: Evaluating the charge of nanofiltration membranes, *J. Mem. Sci.*, 188, 129–136, 2001.
- Schäfer, A. I., Nghiem, L. D., and Waite, T. D.: Removal of the natural hormone estrone from aqueous solutions using nanofiltration and reverse osmosis, *Environ. Sci. Technol.*, 37, 182–188, 2003.
- Shim, Y., Lee, H.-J., Lee, S., Moon, S.-H., and Cho, J.: Effects of natural organic matter and ionic species on membrane surface charge, *Environ. Sci. Technol.*, 36, 3864–3871, 2002.
- Ternes, T. A.: Occurrence of drugs in Germany sewage treatment plants and rivers, *Water Res.*, 32, 3245–3260, 1998.
- Van der Bruggen, B., Schaep, J., Wilms, D., and Vandecasteele, C.: Influence of molecular size, polarity and charge on the retention of organic molecules by nanofiltration, *J. Mem. Sci.*, 156, 29–41, 1999.
- Van der Bruggen, B., Schaep, J., Wilms, D., and Vandecasteele, C.: A comparison of models to describe the maximal retention of organic molecules in nanofiltration, *Separ. Sci. Technol.*, 35, 169–182, 2000.
- Vieno, N., Tuhkanen, T., and Kronberg, L.: Removal of Pharmaceuticals in Drinking Water Treatment: Effect of Chemical Coagulation, *Environ. Technol.*, 27, 183–192, 2006.