Investigation of polar mobile organic compounds (PMOC) removal by reverse osmosis and nanofiltration: rejection mechanism modelling using decision tree

B. Teychene, F. Chi, J. Chokki, G. Darracq, J. Baron, M. Joyeux and H. Gallard

ABSTRACT

Polar mobile organic compounds (PMOC) are highly polar chemicals and tend to accumulate in short water cycles. Due to their properties, PMOC might be partially eliminated by advanced water treatment technologies. The goal of this study is to investigate the rejection of 22 PMOC (highly mobile and persistent) by reverse osmosis (RO) and nanofiltration (NF) membranes. The impact of transmembrane pressure was evaluated through laboratory-scale cross-flow constant pressure filtration tests. Among the investigated experimental conditions, PMOC rejection with NF at eight bars is comparable to values obtained on RO at 15 bars. Negatively charged PMOC are highly rejected by both RO and NF membranes while guanidine-like compounds exhibit higher passage values and are strongly impacted by transmembrane pressure. In order to model the rejection mechanism, decision tree methodology was employed to link PMOC physicochemical properties to rejection values. Based on laboratory-scale results, decision trees were computed and emphasized that the NF rejection mechanism is governed by electrostatic interaction and sieving effects. In contrast, PMOC rejection on the RO membrane strongly depends on the topological polar surface area (TPSA) of the PMOC. This study suggests that micropollutant TPSA should be more investigated in order to describe RO removal efficiency. Moreover, it is shown that the decision tree is a powerful numerical tool in order to reveal the specific sequence leading to micropollutant removal by RO and NF membranes.

Key words | decision tree, deep learning, micropollutants, nanofiltration, PMOC, reverse osmosis

B. Teychene (corresponding author)
F. Chi
J. Chokki
H. Gallard
Institut de Chimie des Milieux et Matériaux de Poitiers (IC2MP, UMR CNRS 7285),
Poitiers,
France
E-mail: benoit.teychene@univ-poitiers.fr

G. Darracq
J. Baron
M. Joyeux
Eau de Paris, Direction de la Recherche & Développement et de la Qualité de l'Eau,
Ivry sur Seine,
France

INTRODUCTION

Polar mobile organic compounds (PMOC, referred to also as persistent mobile organic compounds) are highly polar chemicals and tend to accumulate due to their low degradation properties and slow adsorption kinetics in short water cycles (for example in de facto reuse systems) (Benotti *et al.* 2009; Reemtsma *et al.* 2016; Albergamo *et al.* 2019; Reemtsma *et al.* 2019). Depending on their toxic effects PMOC might be extremely problematic and discussions are ongoing to implement the REACH registration (**R**egistration, **E**valuation, **A**uthorization and **R**estriction of **C**hemicals) of the identified toxic PMOC

(Neumann 2017). In addition, due to their high polarity these compounds are challenging to analyse in water samples and innovative analytical methodologies have been recently developed in order to narrow the knowledge gap of PMOC (Bieber *et al.* 2017; Deeb *et al.* 2017; Montes *et al.* 2017; Gago-Ferrero *et al.* 2018). Recently, Schulze *et al.* (2019) classified the PMOC in three categories based on their detection frequency in EU water samples and awareness. For example *p*–toluenesulfonic acid (defined as a priority 2 compound) is frequently detected at high concentrations and its toxicity should be deeply evaluated

(Glowienke *et al.* 2005). Due to their properties, PMOC might be partially eliminated by advanced technologies such as activated carbon adsorption, membrane and advanced oxidation processes (Kalberlah *et al.* 2014; Gaur *et al.* 2018). However, only 41% of drinking water treatment plants in Europe use advanced treatment processes, underlining a potential risk for consumers regarding these compounds (Neumann & Schliebner 2017). Membrane processes such as reverse osmosis (RO) and nanofiltration (NF) might be good technological solutions to reject PMOC and to protect both consumers and water resources. Indeed, several research works have already demonstrated that RO and NF processes are able to remove compounds of emerging concern (CEC) such as endocrine-disrupting compounds (EDCs) and pharmaceuticals (PhACs)/personal care products (PPCPs) (Al-Rifai *et al.* 2011; Shojaee Nasirabadi *et al.* 2016; Kim *et al.* 2018). Globally, studies relate rejection efficiency with compound characteristics (such as solute molecular weight, size, geometry, charge, and hydrophobicity), water quality (pH, solute concentration, temperature, inorganic and organic water matrix), membrane properties (permeate flux, molecular weight cut-off, pore size, electrostatic charges, salt rejection) and operating filtration conditions (membrane fouling, porosity, cross-flow velocity and transmembrane pressure). As recently reviewed by Kim *et al.* (2018), the retention of CECs by RO membranes is governed by steric exclusion while high rejection might be observed when adsorption due to hydrophobic interactions and attraction due to electrostatic interactions occur for NF. In addition, more polar (dipole moment above 3 D) and hydrophilic compounds (log D < 2) are less rejected by NF membranes (Van der Bruggen *et al.* 1999; Nghiem *et al.* 2005).

Recently, Albergamo *et al.* (2019) evaluated the robustness of the RO process against PMOC. Results showed a good rejection of investigated compounds with rejection ranged from 75% to 99%. In particular, small (~120 g mol^{-1}), neutral and hydrophilic compounds showed the highest passage of about 25% through low-pressure RO membranes (ESPA2). In contrast, neutral and hydrophobic compounds showed a weak inverse correlation between size and passage. Finally, a size–passage correlation was proposed based on statistical analysis, in order to predict neutral PMOC rejection by RO membranes.

Modelling the rejection of micropollutants (MPs) by RO or NF membranes is of great importance for developing robust high-pressure membrane technologies. Numerous articles have proposed a mechanistic understanding of membrane-to-CECs interactions leading to description 'diagrams' describing qualitatively the rejection depending on the physicochemical parameters of the investigated solutes and membranes (Bellona & Drewes 2005; Kim *et al.* 2018). Fewer attempts have been made to 'predict' the removal of CECs (Ammi *et al.* 2018). Recent work has been conducted to investigate the use of multiple linear regressions (MLR), artificial neural networks (ANN), principal component analysis (PCA), partial least squares model (PLS), quantitative structure activity relationship (QSAR) or bootstrap aggregated neural networks (BANN) in order to model and predict the RO and NF removal efficiency (Yangali-Quintanilla *et al.* 2008; Ammi *et al.* 2015, 2018; Flyborg *et al.* 2017). Yangali-Quintanilla *et al.* (2009, 2010) developed a model for organic contaminant removal on NF and RO membranes using ANN models based on QSAR methodology. Their model underlines the importance of the geometric properties of MPs, the membranes' molecular weight cut-off (MWCO), salt rejection and MP hydrophobicity (log D) values in rejection mechanisms. Recently, Ammi *et al.* (2018) obtained robust and accurate prediction of the removal of CECs using the BANN method involving organic compound diameter in water and membrane properties (MWCO, salt rejection) as input model parameters.

ANN, MLR and PLS have several drawbacks regarding prediction of new data (data which were not used to train the network) (Ammi *et al.* 2018). In addition, ANN might be considered as a 'black box' that delivers results without a clear explanation on how the output was obtained. Therefore it is difficult to understand how the decisions were made based on the output of the model. In contrast, classification and regression tree, which are similar to ANN, have the advantage of expressing regularities explicitly and thus being convenient for inspecting the different decisions that have been made for modelling (Walley & Džeroski 1996). Decision tree algorithms are based on a divide-and-conquer approach (Quinlan 1986). The algorithms work top-down, splitting recursively the input data into node and leaves. The leaves correspond to model output obtained by the 'chain-decision' based on the parameters defined in the

input dataset. For example, decision trees have already been used to model the oxidation rate of CEC by ozone (Gupta & Basant 2017), and in the analysis of membrane fouling (Chen *et al.* 2012).

Therefore the decision tree might be useful to model CEC removal by RO and NF membranes and to obtain a prediction of removal efficiency and a description of the rejection mechanism, respectively. Moreover, the decision tree methodology returns results in a similar way to the description diagram often reported in literature review, which is convenient to describe CEC rejection mechanisms.

To the best of our knowledge, no study has been reported in the literature that has used decision tree methodology to describe and model the rejection efficiency of CECs and more specifically the removal of highly polar, persistent and mobile organic compounds (PMOC). Therefore, the aim of the present study is to investigate the rejection of highly ranked PMOC (high persistency and mobility detected in the Schulze *et al.* (2019) study) by RO and NF membranes. The impact of transmembrane pressure was evaluated through laboratory-scale cross-flow constant pressure filtration tests. Decision tree methodology was used to mine the experimental data and describe the rejection mechanism of PMOC.

MATERIALS AND METHODS

PMOC properties and chemical analysis

The properties of the selected PMOC are listed in Table S1 (Supplementary Material). All PMOC properties (MW (g mol^{-1}), log D, polarizability (Å^3), topological polar surface area (Å^2)) were computed thanks to the Chemaxon website. PMOC geometry properties (width, length) was computed thanks to Chem3D software.

PMOC analysis was performed by liquid chromatography–mass spectrometry (LC-MS) using the Dionex Ultimate 3000 standard system (Thermo Scientific, USA). Separation was conducted on a Kinetex F5 column (Phenomenex, USA) at a flow rate of 0.3 ml min^{-1}. The liquid chromatography was coupled to a quadrupole time-of-flight mass spectrometer (Impact HD, Bruker, USA) with an electrospray ionization source operated in both positive and negative modes. Every sample was analysed

in triplicate. PMOC calibration curves were plotted from 10^{-9} to 10^{-6} mol L^{-1} for every PMOC in dechlorinated water and MilliQ water for feed and permeate quantification, respectively. Limit of detection (LOD) and quantification (LOQ) were calculated for every compound using the 'chemCal' packages on RStudio (version 0.2.1).

Filtration test procedure

The filtration experiments were carried out with an Osmonic cell (SEPA CF II, General Electrics, USA) using flat sheet membrane coupons (140 cm^2). The membrane coupon was placed into the cell sandwiched between a permeate and feed spacer, similar to the ones found in spiral wound modules (extracted from a BW30-4040 Fimtech membrane module, thickness equal to 34 mil). As shown in Figure S1 (Supplementary Material), the filtration pilot consists of (1) a high-pressure pump (Wanner G13, Axflow, France), (2) a 30-litre feed-tank with a heating/cooling jacket in order to keep a constant temperature (at 20 °C) during the filtration test and finally (3) different valves in order to control the feed pressure and the recirculation flow. The permeate flow was recorded using an electronic balance (VIC Series, Acculab France) connected to a computer (4). Before every filtration test, the tested membrane was compacted for 8 h at 30 bars using pure water (18.2 MΩ·cm at 25 °C, Millipore, France) and then a permeability test was performed in order to check the integrity of the membrane.

The rejection was evaluated at three transmembrane pressures (TMP) equal to 15, 10 and 8 bars at constant crossflow velocity of about 120 L·h^{-1} (Re = 630 laminar regime according a theoretical calculation in such a system (Schock & Miquel 1987)). Experiment duration was set at 78 h for every investigated TMP in recycling mode (closed loop for permeate and concentrate). A new virgin membrane was used for every test. The selected RO and NF membranes supplied by the DOW Chemical Company were the BW30 and the NF90, respectively.

The feed water consisted of 14 L of dechlorinated drinking water (pH 7.8, hardness = 30°F, TOC = 0.8 mg L^{-1}, conductivity ≈ 800 μS/cm) and was used to prepare a feed solution spiked at 10^{-6} mol L^{-1} of every PMOC. Experiments were conducted at high PMOC concentration in

order to speed up steady-state conditions (due to potential membrane adsorption).

The rejection was calculated by Equation (1):

$$Robs(\%) = 1 - Cp/Cf \tag{1}$$

where Cp and Cf represent the permeate and feed PMOC concentrations measured at 78 h, respectively.

Decision tree methodology

PMOC rejection values obtained after 78 h of filtration for every filtration test on both membranes were processed with RStudio software using the deep learning package (Classification and Regression Training package (caret), version 6.0.80). Decision trees were obtained by linking PMOC physicochemical parameters (Table S1) to rejection values on both membranes based on ANOVA methodology with a complexity parameter (cp) equal to 0.05.

RESULTS AND DISCUSSION

First of all, among the 22 investigated PMOC, four compounds (1,4-diazabicyclo[2.2.2]octane, ametryn, melamine, dimethylbenzylamine) showed inconsistent results in feed water due to a matrix effect on LC/MS analysis and were not taken into account.

The PMOC concentrations in the feed water obtained after 3 h and 78 h of filtration are reported in Figure S2 (Supplementary Material). First of all, it can be seen that the ϵ-caprolactam and tris(1–chloro-2-propyl) phosphate (TCPP) compounds returned high concentration values after 3 h of filtration at 15 bars on the BW30 membrane. This might be due to a high matrix effect due to the dechlorinated water. However, it can be seen that TCPP showed significant feed water concentration variation after 78 h of filtration at 10 and 15 bars on the NF90 membrane and at the three TMP for the BW30 membrane. According to Table S1, TCPP is neutral, hydrophobic (log D at pH 8 = 3.36) and exhibits the highest MW (equal to 327.563 Da), therefore it might adsorb on the membranes.

As shown in Figure S2 and according to analytical errors, adsorption of PMOC is low on both membranes whatever the investigated TMP. However, the 3,5-ditertbutyl

salicylic acid and the sulfanilic acid compounds showed the highest feed concentration variation after 78 h of filtration, which might be attributed to adsorption phenomena whatever the investigated membranes. For instance, the average variation concentration over the three TMP for the 3,5–ditertbutyl salicylic acid was equal to $3.4 \pm 0.4 \ 10^{-7}$ and $5.4 \pm 0.4 \ \text{mol L}^{-1}$ on the NF90 and BW30 membranes, respectively. Sulfanilic acid shows an average concentration variation on the NF90 at the three TMP of about $3.4 \pm 0.7 \ 10^{-7} \ \text{mol L}^{-1}$ while adsorption was only observed at ten and 15 bars on the BW30 membrane (concentration variation of $3.9 \pm 0.7 \ 10^{-7} \ \text{mol L}^{-1}$). In addition, the acetoguanamine compound showed strong feed concentration variation at eight and ten bars on the BW30 membranes (concentration variation equal to $6.0 \pm 0.7 \ 10^{-7} \ \text{mol L}^{-1}$).

According to Table S1, the 3,5-ditertbutyl salicylic acid and sulfanilic acid compounds are both hydrophilic and negatively charged, therefore they should not strongly adsorb on the membranes. In contrast, the acetoguanamine compound is hydrophilic and positively charged at the investigated pH, which might explain the adsorption observed on the BW30 membranes at low TMP.

These results suggest that PMOC are only slightly adsorbed on the BW30 and NF90 for a short operating time (78 h). However, more investigations are needed to fully investigate the potential adsorption of PMOC on the investigated membranes.

The rejection values obtained at 78 h on both membranes at the three investigated TMP (eight, ten and 15 bars) are reported in Table 1. The rejection values for the NF90 membrane are high and decrease with the TMP. At eight bars, only nine compounds could be quantified while a total of 17 compounds were quantified at 15 bars on the NF90 membrane. The TCPP could not be quantified in the permeate, confirming the important role of the sieving effect and its potential adsorption on the NF90 membrane material. Similarly, the sulfanilic acid exhibited a permeate concentration lower than the LOQ, confirming its adsorption on the NF90.

Among the quantified compounds, rejection values obtained on the NF90 membrane ranged from $92.5\% \pm 2.9\%$ to $99.7\% \pm 0.1\%$ and from $88.7\% \pm 2.3\%$ to $98.3\% \pm 0.8\%$ at eight and 15 bars, respectively. These results are in accordance with previous studies, which showed that the rejection of CEC decreased with the applied pressure (Zhu

Table 1 | Rejection values of PMOC on NF and RO membranes at the three investigated TMP

	Rejection values (%)					
	NF			**RO**		
	8 bars	**10 bars**	**15 bars**	**8 bars**	**10 bars**	**15 bars**
Ditolylguanidine	98.1% ± 0.7	96.6% ± 1.1	90.5% ± 3.0	94.3% ± 1.8	96.4% ± 1.2	96.3% ± 1.4
1.3-Diphenylguanidine	94.0% ± 1.4	90.9% ± 1.9	88.7% ± 2.3	93.8% ± 1.4	96.1% ± 1.0	95.3% ± 1.5
Amantadine	99.1% ± 0.4	98.4% ± 0.5	89.8% ± 0.9	94.1% ± 0.7	96.8% ± 0.5	95.8% ± 0.9
Bisphenol S	99.3% ± 0.2	98.9% ± 0.3	95.9% ± 0.9	98.0% ± 0.5	97.2% ± 0.6	99.1% ± 0.3
Acetoguanamine	92.5% ± 2.9	90.1% ± 3.2	91.4% ± 2.9	93.5% ± 2.9	<LOQ	<LOQ
Acesulfame potassium	99.6% ± 1.0	99.4% ± 0.2	95.3% ± 0.9	97.9% ± 0.5	97.3% ± 0.6	98.9% ± 0.3
2-Toluenesulfonic acid	<LOQ	<LOQ	95.7% ± 0.5	98.1% ± 0.5	96.3% ± 0.5	<LOQ
3.5-Ditertbutyl salicylic acid	98.6% ± 0.2	99.5% ± 0.1	97.0% ± 0.2	98.9% ± 0.2	98.5% ± 0.2	99.2% ± 0.1
Naphthalenesulfonic acid	<LOQ	99.5% ± 0.1	96.0% ± 0.5	98.1% ± 0.2	97.3% ± 0.4	<LOQ
p-Toluenesulfonic acid	<LOQ	99.6% ± 0.2	96.2% ± 0.7	98.4% ± 0.4	97.3% ± 0.6	99.4% ± 0.3
Sulfanilic acid	<LOQ	<LOQ	98.3% ± 0.8	<LOQ	97.9% ± 1.1	<LOQ
Trifluoromethanesulfonic acid	99.7% ± 0.1	99.5% ± 0.1	95.4% ± 0.3	98.0% ± 0.2	97.4% ± 0.2	98.9% ± 0.1
Benzyltrimethylammonium chloride	97.9% ± 0.8	95.8% ± 0.9	90.4% ± 1.0	94.3% ± 0.9	96.2% ± 0.9	95.1% ± 1.3
ε-Caprolactam	<LOQ	<LOQ	92.5% ± 3.4	<LOQ	<LOQ	<LOQ
Methyl sulfate sodium salt	<LOQ	98.7% ± 0.5	96.4% ± 0.6	98.7% ± 0.4	91.7% ± 1.4	<LOQ
Sodium acryloyldimethyltaurate	<LOQ	99.7% ± 0.1	96.8% ± 0.4	98.5% ± 0.2	95.9% ± 0.5	99.0% ± 0.2
Sodium *O*-xylene-4-sulfonate	<LOQ	<LOQ	96.2% ± 0.3	98.3% ± 0.3	96.1% ± 0.4	<LOQ
Tris(1-chloro-2-propyl)phosphate	<LOQ			<LOQ		
Total quantified compounds	9	13	17	15	15	10

LOQ: limit of quantification in permeate; standard error in percentage points (pp)

Feed water: dechlorinated drinking water; PMOC: 10^{-6} mol L^{-1}; duration: 78 h; temperature: 20 °C.

2015; Licona *et al.* 2018; Albergamo *et al.* 2019). Indeed, convective transport dominates solute transport through NF membranes (Kim *et al.* 2007). Therefore, lower rejections at 15 bars are due to higher convective flow through the NF membrane. Acetoguanamine (the least rejected compound at eight bars and adsorbed on the NF90) is positively charged and has a molecular weight equal to 125 g mol^{-1} while trifluoromethanesulfonic acid (the most rejected compound at eight bars) exhibits an MW equal to 150 g mol^{-1} and is negatively charged. Except for the ε–caprolactam and TCPP, which are uncharged, all non-quantifiable compounds (<LOQ in Table 1) exhibit a molecular weight higher than the NF90 molecular weight cut-off (MWCO: 200 g mol^{-1}) and/or are negatively charged. Similar results were found in the literature; rejection values were observed equal to 71% and 99.5% for caprolactam on NF70

and BW30 at 15 bars, respectively (Van der Bruggen *et al.* 1999; Ahmed *et al.* 2015). Similarly, TCPP was found to be strongly rejected (above 99% at five bars) by NF membranes in water reuse application (Bellona & Drewes 2007).

Interestingly, the TMP has a strong negative impact on the rejection of the positively charged organic amine compounds such as di-tolyl-guanidine, 1,3-diphenylguanidine and amantadine (compound not adsorbed on NF90 material, Figure S2). The NF90 membrane is considered hydrophilic and is negatively charged at the investigated pH (contact angle: 41.4° ± 0.5° and zeta potential: −20 mV) (Licona *et al.* 2018). Therefore, these results suggest that the main rejection mechanisms of PMOC on the NF90 are sieving and electrostatic repulsion effects.

Regarding the RO BW30 membrane, the PMOC rejection values are globally high (higher than 93%), higher

than the values obtained on the NF90 membrane, and increase with the TMP. Whatever the investigated pressure, the hydrodynamic ratio is lower than 1 (Jo/k equals 0.4 and 0.8 at eight and 15 bars, respectively) indicating that the diffusion phenomenon governs the mass transfer (Jo/k, calculation details given in Porter (1972)).

Among the quantified compounds, the highest rejected compound is the 3,5–ditertbutyl salicylic acid with an MW equal to 250 g mol^{-1} and negatively charged. Similar to the NF membrane, the lowest rejected compounds at eight bars after 78 h of filtration are acetoguanamine and 1,3-diphenylguanidine. However, with the increase in TMP, acetoguanamine could not be quantified in the permeate sample. The TCPP is strongly rejected due to its high MW (327 g mol^{-1}) and potential adsorption.

As shown here, the description of the rejection mechanism of organic micropollutants is complex and depends on several parameters (for example MW, electrostatic charge, log D). Therefore, in order to model PMOC rejection by NF and RO membranes, the decision tree calculation was performed using various molecule descriptors as input (log D, MW, length, width, polarizability, etc.) and rejection values as output.

Decision trees were computed on both membranes showing similar hydrodynamic conditions (similar Jo/k corresponding to NF90-8 bars and BW30-15 bars) and are reported on Figure 1. As shown in Figure 1(a), the decision tree consists of three nodes and four leaves corresponding to the PMOC rejection values obtained at eight bars on the NF90 membrane. The rejection values are best discriminated by the MW of quantified PMOC in the permeate with

a threshold value equal to 137 g mol^{-1}. The first leaf consists of one PMOC (acetoguanamine) with a small MW (125 g mol^{-1}). The second node regroups eight PMOC with higher MW and splits compounds into two groups according to their log D. Compounds with low log D (lower than 1.66 corresponding to six PMOC, see Table S1) are highly rejected (last leaf on the right, Figure 1(a)). Compounds with higher log D (log D > 1.66) are further split according to their MW through the third node. Interestingly in this third node, the MW threshold value corresponds to the MWCO of the NF90 (approximately 200 g mol^{-1}). The two resulting leaves correspond to the 1,3-diphenylguanidine (rejection: 94%) and ditolylguanidine (rejection: 98.1%), respectively. Therefore, this result shows that guanidine-like compounds' rejection by the NF membrane is strongly influenced by their MW.

According to the decision tree plotted in Figure 1(a) the rejection mechanism is mostly governed by the sieving effect (first descriptor MW) and for high-MW compounds electrostatic and hydrophobic interactions lead to higher rejection. Moreover, the decision tree showed that the MW variable might be used two times (node 1 and node 4 in Figure 1(a)) to discriminate rejection values for large and hydrophobic PMOC, which it is not possible to observe using MLR or ANN modelling, for example. This decision tree (Figure 1(a)) confirms previous results obtained on NF membranes showing that low-MW compounds are less rejected while high rejection is obtained for high-MW and charged compounds (Van der Bruggen *et al.* 1999).

The decision tree built on the PMOC rejection obtained on the BW30 at 15 bars is simpler (consisting of one node
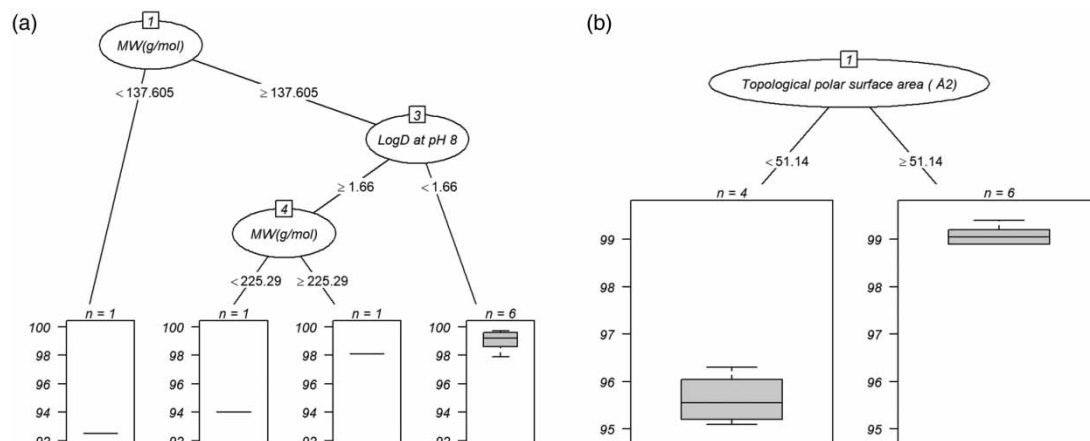

Figure 1 | Decision tree obtained on rejection data versus PMOC properties for (a) NF90-8 bars and (b) BW30-15 bars.

and two leaves) and shows that solely the topological polar surface area (TPSA) of PMOC could discriminate rejection data with a threshold value equal to 51 Å² (Figure 1(b)). TPSA is usually used to optimize pharmaceutical permeation into the cell membrane and also its accessibility to solvent (Barret 2018). Therefore according to Figure 1(b), high TPSA will induce high rejection around 99% due to low diffusion within the membrane material. Rejection values plotted against the TPSA values for the BW30 membranes at 15 bars clearly showed the two clusters and the threshold value of 51 Å² (Figure S3, Supplementary Material). This suggests that 'TPSA cut-off' should be further considered to model CEC rejection on RO membranes. Further experiments should be done to fully confirm this assumption.

In terms of modelling approach, such non-linear variation could not be well modelled using MLR and suggests that the decision tree is a powerful numerical tool to determine the best descriptor and to model non-linear variation such as QSAR methodology.

CONCLUSIONS

The present study investigates the rejection at laboratory-scale of 22 PMOC, detected in EU water resources, on RO and NF membranes at three different TMP. Globally, PMOC are highly rejected by membranes whatever the considered experimental conditions. Due to the PMOC properties (mostly hydrophilic and negatively charged), low adsorption was observed on both membranes. Filtration tests demonstrated that the less rejected compounds are positively charged aromatic amine compounds and TMP has a strong impact on compound passage.

Decision tree methodology was employed to model the rejection mechanism. Results confirm that PMOC rejection on NF membranes depends on the hydrophilicity/electrostatic interactions and sieving effects. In contrast, for RO membranes at 15 bars, the TPSA was the best descriptor of PMOC rejection, confirming a diffusive transport through thin-film composite membranes. This result suggests that the TPSA of CEC should be more used in future studies to model RO performances.

In terms of modelling, decision trees are able to reveal a specific sequence of decisions leading to the best description of the rejection data, which is not possible to obtain using MLR or ANN approaches. In addition, decision tree modelling is less time-consuming and easier to interpret than ANN or QSAR models. To conclude, future studies should focus on the training of decision tree models based on larger datasets in order to investigate the CEC removal prediction by NF and RO membranes for water treatment application.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this paper is available online at https://dx.doi.org/10.2166/ws.2020.020.

REFERENCES

Ahmed, H., Karakulski, K. & Pehlivan, E. 2015 Pilot plant studies on the recovery of caprolactam by reverse osmosis. *International Journal of Enhanced Research in Science Technology & Engineering.* **4** (3), 154–160.

Albergamo, V., Blankert, B., Cornelissen, E. R., Hofs, B., Knibbe, W.-J., van der Meer, W. & de Voogt, P. 2019 Removal of polar organic micropollutants by pilot-scale reverse osmosis drinking water treatment. *Water Research* **148**, 535–545.

Al-Rifai, J. H., Khabbaz, H. & Schäfer, A. I. 2011 Removal of pharmaceuticals and endocrine disrupting compounds in a water recycling process using reverse osmosis systems. *Separation and Purification Technology* **77**, 60–67.

Ammi, Y., Khaouane, L. & Hanini, S. 2015 Prediction of the rejection of organic compounds (neutral and ionic) by nanofiltration and reverse osmosis membranes using neural networks. *Korean Journal of Chemical Engineering* **32** (11), 2300–2310.

Ammi, Y., Khaouane, L. & Hanini, S. 2018 A model based on bootstrapped neural networks for modeling the removal of organic compounds by nanofiltration and reverse osmosis membranes. *Arabian Journal for Science and Engineering* **43**, 6271–6284.

Barret, R. 2018 Importance and evaluation of the polar surface area (PSA and TPSA). In: *Medicinal Chemistry: Fundamentals* (R. Barret, ed.), Elsevier, Oxford, UK, pp. 89–95.

Bellona, C. & Drewes, J. E. 2005 The role of membrane surface charge and solute physico-chemical properties in the rejection of organic acids by NF membranes. *Journal of Membrane Science* **249**, 227–234.

Bellona, C. & Drewes, J. E. 2007 Viability of a low-pressure nanofilter in treating recycled water for water reuse applications: a pilot-scale study. *Water Research* **41** (17), 3948–3958.

Benotti, M. J., Trenholm, R. A., Vanderford, B. J., Holady, J. C., Stanford, B. D. & Snyder, S. A. 2009 Pharmaceuticals and endocrine disrupting compounds in US drinking water. *Environmental Science & Technology* **43** (3), 597–603.

Bieber, S., Greco, G., Grosse, S. & Letzel, T. 2017 RPLC-HILIC and SFC with mass spectrometry: polarity-extended organic molecule screening in environmental (water) samples. *Analytical Chemistry* **89** (15), 7907–7914.

Chen, J., Yang, Y.-C. & Wei, T.-Y. 2012 Application of wavelet analysis and decision tree in UTDR data for diagnosis of membrane filtration. *Chemometrics and Intelligent Laboratory Systems* **116**, 102–111.

Deeb, A. A., Stephan, S., Schmitz, O. J. & Schmidt, T. C. 2017 Suspect screening of micropollutants and their transformation products in advanced wastewater treatment. *Science of The Total Environment* **601–602**, 1247–1253.

Flyborg, L., Björlenius, B., Ullner, M. & Persson, K. M. 2017 A PLS model for predicting rejection of trace organic compounds by nanofiltration using treated wastewater as feed. *Separation and Purification Technology* **174**, 212–221.

Gago-Ferrero, P., Krettek, A., Fischer, S., Wiberg, K. & Ahrens, L. 2018 Suspect screening and regulatory databases: a powerful combination to identify emerging micropollutants. *Environmental Science & Technology* **52** (12), 6881–6894.

Gaur, N., Narasimhulu, K. & Pydi Setty, Y. 2018 Recent advances in the bio-remediation of persistent organic pollutants and its effect on environment. *Journal of Cleaner Production* **198**, 1602–1631.

Glowienke, S., Frieauff, W., Allmendinger, T., Martus, H.-J., Suter, W. & Mueller, L. 2005 Structure–activity considerations and in vitro approaches to assess the genotoxicity of 19 methane-, benzene- and toluenesulfonic acid esters. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **581** (1–2), 23–34.

Gupta, S. & Basant, N. 2017 Modeling the pH and temperature dependence of aqueousphase hydroxyl radical reaction rate constants of organic micropollutants using QSPR approach. *Environmental Science and Pollution Research* **24**, 24936–24946.

Kalberlah, F., Oltmanns, J. & Schwarz, M. 2014 *Guidance for the Precautionary Protection of Raw Water Destined for Drinking Water Extraction from Contaminants Regulated under REACH*. UFOPLAN Project FKZ 371265416, German Federal Environmental Agency, Dessau, Germany.

Kim, T. U., Drewes, J. E., Summers, R. S. & Amy, G. L. 2007 Solute transport model for trace organic neutral and charged compounds through nanofiltration and reverse osmosis membranes. *Water Research* **41** (17), 3977–3988.

Kim, S., Chu, K. H., Al-Hamadani, Y. A. J., Park, C. M., Jang, M., Kim, D., Yu, M., Heo, J. & Yoon, Y. 2018 Removal of contaminants of emerging concern by membranes in water and wastewater: a review. *Chemical Engineering Journal* **335**, 896–914.

Licona, K. P. M., Geaquinto, L. R. de O., Nicolini, J. V., Figueiredo, N. G., Chiapetta, S. C., Habert, A. C. & Yokoyama, L. 2018 Assessing potential of nanofiltration and reverse osmosis for removal of toxic pharmaceuticals from water. *Journal of Water Process Engineering* **25**, 195–204.

Montes, R., Aguirre, J., Vidal, X., Rodil, R., Cela, R. & Quintana, J. B. 2017 Screening for polar chemicals in water by trifunctional mixed-mode liquid chromatography–high resolution mass spectrometry. *Environmental Science & Technology* **51** (11), 6250–6259.

Neumann, M. 2017 Vorschlag für Kriterien und ein Bewertungskonzept zur Identifizierung von persistenten, mobilen und toxischen (PMT-) Stoffen zum Schutz des Rohwassers zur Trinkwassergewinnung unter der EU-Verordnung REACH (Proposal for criteria and an assessment concept for the identification of Persistent, Mobile and Toxic (PMT) substances to protect raw water for the production of drinking water under the EU regulation REACH). In: *Grundwasserschutz im Spannungsfeld zwischen Nachhaltigkeit und Ökonomie*, Zentralblatt für Geologie und Paläontologie, Teil I, Schweizerbart, Stuttgart, Germany, pp. 91–101.

Neumann, M. & Schliebner, I. 2017 *Protecting the Sources of Our Drinking Water. A Revised Proposal for Implementing Criteria and an Assessment Procedure to Identify Persistent, Mobile and Toxic (PMT) and Very Persistent, Very Mobile (vPvM) Substances Registered under REACH*. German Environment Agency, Dessau, Germany.

Nghiem, L. D., Schäfer, A. I. & Elimelech, M. 2005 Pharmaceutical retention mechanisms by nanofiltration membranes. *Environmental Science & Technology* **39**, 7698–7705.

Porter, M. C. 1972 Concentration polarization with membrane ultrafiltration. *Industrial & Engineering Chemistry Product Research and Development* **11**, 234–248.

Quinlan, J. R. 1986 Induction of decision trees. *Machine Learning* **1** (1), 81–106.

Reemtsma, T., Berger, U., Arp, H. P. H., Gallard, H., Knepper, T. P., Neumann, M., Quintana, J. B. & de Voogt, P. 2016 Mind the gap: persistent and mobile organic compounds – water contaminants that slip through. *Environmental Science & Technology* **50** (19), 10308–10315.

Schock, G. & Miquel, A. 1987 Mass transfer and pressure loss in spiral wound modules. *Desalination* **64**, 339–352.

Schulze, S., Zahn, D., Montes, R., Rodil, R., Benito Quintana, J., Knepper, T. P., Reemtsma, T. & Berger, U. 2019 Occurrence of emerging persistent and mobile organic contaminants in European water samples. *Water Research* **153**, 80–90.

Shojaee Nasirabadi, P., Saljoughi, E. & Mousavi, S. M. 2016 Membrane processes used for removal of pharmaceuticals, hormones, endocrine disruptors and their metabolites from wastewaters: a review. *Desalination and Water Treatment* **57** (51), 24146–24175.

Van der Bruggen, B., Schaep, J., Wilms, D. & Vandecasteele, C. 1999 Influence of molecular size, polarity and charge on the retention of organic molecules by nanofiltration. *Journal of Membrane Science* **156**, 29–41.

Walley, W. J. & Džeroski, S. 1996 Biological monitoring: a comparison between Bayesian, neural and machine learning methods of water quality classification. In: *Environmental Software Systems: Proceedings of the International Symposium on Environmental Software Systems, 1995* (R. Denzer, G. Schimak & D. Russell, eds), Springer, Dordrecht, The Netherlands, pp. 229–240.

Yangali-Quintanilla, V., Kim, T.-U., Kennedy, M. & Amy, G. 2008 Modeling of RO/NF membrane rejections of PhACs and organic compounds: a statistical analysis. *Drinking Water Engineering and Science* **1** (1), 7–15.

Yangali-Quintanilla, V., Verliefde, A., Kim, T.-U., Sadmani, A., Kennedy, M. & Amy, G. 2009 Artificial neural network models based on QSAR for predicting rejection of neutral organic compounds by polyamide nanofiltration and reverse osmosis membranes. *Journal of Membrane Science* **342** (1–2), 251–262.

Yangali-Quintanilla, V., Sadmani, A., McConville, M., Kennedy, M. & Amy, G. 2010 A QSAR model for predicting rejection of emerging contaminants (pharmaceuticals, endocrine disruptors) by nanofiltration membranes. *Water Research* **44** (2), 373–384.

Zhu, L. 2015 Rejection of organic micropollutants by clean and fouled nanofiltration membranes. *Journal of Chemistry* **2015**, 934318.

ChemCal package: https://pkgdown.jrwb.de/chemCal/.

Caret package: https://github.com/topepo/caret/.