Optimization of a Membrane Filtration Process for Drinking Water Treatment using Fluorescence-based measurements

R. H. Peiris*, H. Budman*, C. Moresoli* and R.L. Legge*

*Department of Chemical Engineering, University of Waterloo, 200 University Avenue West, Waterloo, ON, N2L 3G1 Canada
(Tel: 519-888-4567 ext. 36980; e-mail: hrpeiris@engmail.uwaterloo.ca; hbudman@engmail.uwaterloo.ca; cmoresoli@uwaterloo.ca; rllegge@engmail.uwaterloo.ca)
(Submitted to invited Energy session at the DYCOPS 2010 Symposium)

Abstract: Membrane fouling control is of paramount importance for sustainable operation of membrane-based drinking water treatment processes. Natural organic matter (NOM) is considered as the major membrane foulant and therefore its characterization is important for implementing fouling control strategies. This study proposes a fluorescence-based modeling approach for estimating and predicting the fouling dynamics in a bench-scale ultrafiltration (UF) membrane cross flow set-up for drinking water treatment. Principal component analysis (PCA) was used to extract the information that is relevant for membrane fouling from fluorescence excitation-emission matrix measurements captured during UF operation. PCA extracted principal components (PCs) that were related to major NOM membrane foulants. The model predictions were based on PC scores of retentate and permeate captured at time = 15 min of the UF experiments. The proposed fluorescence-based modeling approach is able to forecast different fouling behaviours with good accuracy. This proposed approach was then used for optimization of the UF process in which membrane back-washing times were estimated in order to achieve minimum energy consumption while ensuring maximum production of drinking water.

Keywords: drinking water treatment, fluorescence spectroscopy, membrane filtration, optimization, prediction, principal component analysis, real-time optimization.

1. INTRODUCTION

Membrane-based technologies are increasingly used to achieve improved removal of pathogenic organisms and comply with water quality related regulatory limits in drinking water treatment applications. However, membrane fouling is a major constraint for maintaining efficient membrane-based drinking water treatment processes. Membrane fouling in drinking water applications is mainly caused by the presence of natural organic matter (NOM) and colloidal/particulate matter in water (Saravia et al., 2006; Jermann et al., 2007). NOM-related fouling increases operational costs as a result of permeate flux decline and/or increased energy consumption due to higher trans-membrane pressure (TMP) requirements. In addition, the need for frequent chemical cleaning of fouled membranes leads to the deterioration of membrane performance, shortened service life and increased costs.

Preventing or reducing membrane fouling while ensuring a high production of water flux is therefore essential to reduce the energy demand and other operational costs associated with fouling for sustainable operation of membrane-based drinking water treatment facilities. This could be achieved by optimizing the operation of the membrane filtration processes (Seidel and Elimelech, 2002).

1.1 Characterization of Membrane Foultants

The individual and combined effects of different NOM fractions, such as humic-, protein- and polysaccharide-like substances, as well as colloidal/particulate matter present in natural water contribute to different membrane fouling behaviour. Characterization of these components in the raw water serving as feed to the membrane operations is essential for the understanding and identification of their changes that occur during the membrane filtration process. This information can then be used for the development of fouling control and optimization strategies.

In this study, the fluorescence excitation-emission matrix (EEM) approach has been used as a method of characterizing NOM membrane foulants in water as it is able to capture specific fluorescence features that correspond to humic- and protein-like materials (Henderson et al., 2009). The light scattering regions captured in the fluorescence EEMs can also provide information related to the particulate/colloidal matter present in water (Peiris et al., 2010). The fluorescence EEMs of natural water contain a large number of intensity readings recorded at different excitation and emission wavelength combinations. Unlike other available NOM characterization methods (Huber et al., 1992; Her et al., 2003; Gray et al., 2007), this approach is able to differentiate the...
1.2 Fouling Monitoring and Optimization Approach

In a previous study, principal component analysis (PCA) was successfully used to de-convolute spectral information captured within fluorescence EEMs into principal components (PCs) that were related to humic-like, protein-like and colloidal/particulate matter present in natural water (Peiris et al., 2010). This approach, which was based on the PC scores generated by the PCA of fluorescence EEMs, is suitable for rapid monitoring of the performance of a membrane-based drinking water treatment system with high sensitivity. The same approach was therefore used in this study to generate PC scores that correspond to the fluorescence EEMs captured over the course of the membrane filtration operation. These PC scores were then used as states within a system of differential equations in a dynamic model to estimate and predict ultrafiltration (UF) membrane fouling dynamics. Then, based on these predicted dynamics, an optimization approach is proposed for the estimation of optimal membrane back-washing times corresponding to minimum energy consumption while ensuring maximum production levels of drinking water.

2. MATERIALS AND METHODS

Grand River water (GRW) (Southwestern Ontario, Canada) was filtered using a 200 micron filter (038A-2080; Keller Products, Inc. Acton, MA) and used as the membrane feed in UF experiments. The dissolved organic content (DOC) of the membrane feed ranged from 3.9 – 6.5 mg/L and its turbidity values were in the range of 1.2 – 3.8 NTU.

2.1 Bench-scale Membrane Filtration Set-up

UF of GRW was performed using a bench-scale cross flow set-up shown in Fig. 1. Flat sheet UF membranes (Polysulfone - YMEWS3001; GE Osmonics) with a 60 kDa molecular weight cut-off (MWCO) were used. The initial pure water flux at TMP = 15 psi (103.4 kPa) was ~ 2.4 L/min.m². A new membrane was used for each filtration run.

Filtered GRW was fed to the membrane set-up at 0.6 L/min. The retentate was circulated back to the feed tank. The TMP was maintained at 15 psi (103.4 kPa) and the temperature of the feed tank was maintained at ~ 25 °C using a temperature controller. The permeate water flux was recorded with a balance connected to a computer using a LabView-based interface (version 8.0; National Instruments, Austin, TX). The filtration consisted of a two step operation cycle: (1) permeation period and (2) back-washing for 20 s. For non-optimized conditions, the permeation period was 1 h while for optimized back-washing, the permeation period was varied to accommodate the back-washing times calculated based on the optimization approach discussed later. Back-washing of the membrane was implemented by forcing the permeate flow in the opposite direction through the membrane using pressurized Nitrogen gas at 10 psi (68.9 kPa). Fluorescence EEMs of both retentate and permeate were recorded at 15 min intervals during the course of the filtration as explained below.

The fluorescence EEMs were recorded using a Varian Cary Eclipse Fluorescence Spectrofluorometer (Palo Alto, CA) by scanning 301 individual emission spectra (300 – 600 nm) at sequential 10 nm increments of excitation wavelengths between 250 and 380 nm. A detailed description of the fluorescence analysis procedure and the selection of the spectrofluorometer parameter settings used in this study for obtaining reproducible fluorescence signals, especially for low NOM concentrations levels, is found in Peiris et al. (2008; 2009). To eliminate water Raman scattering and to reduce other background noise, fluorescence spectra for Milli-Q water, obtained under the same conditions, were subtracted from all spectra. During the course of the fluorescence analyses, the Raman scattering peak intensity recorded for Milli-Q water at Ex/Em ~ 348 nm/396 nm was less than 1%, confirming that there were no significant fluctuations in the performance of the spectrofluorometer lamp or other hardware. The temperature of the water samples were maintained at ~ 25 °C during the analysis.

2.3 Fluorescence Data Pre-treatment and PCA

The fluorescence EEM of each sample contained 4214 excitation and emission coordinate points. The fluorescence intensity values corresponding to all 4214 coordinate points (spectral variables) of each EEM were rearranged following the fluorescence EEM data rearrangement procedure described in Peiris et al. (2010). This resulted in a n x 4214
PCA was applied to matrix X to generate PC scores as explained in Persson and Wedborg (2001) and Peiris et al. (2010). Essentially, PCA extracts a smaller set of underlying new variables that are uncorrelated, orthogonal and mathematically represented by linear combinations of original variables in the X matrix. These new variables, referred to as PCs, are calculated to account for much of the variance present in the X matrix (Wold et al., 1987; Eriksson et al., 2001) and therefore are able to describe major trends in the original spectral data in X matrix. PCA decomposes the data matrix X as the sum of the outer product of vectors $s_i$ and $p_i$ plus a residual matrix $E$ as presented in equation (1).

$$X = \sum_{i=1}^{n} s_i \cdot p_i + E$$  

Before performing PCA analysis, X matrix was auto-scaled, i.e. adjusted to zero mean and unit variance by dividing each column by its standard deviation. To determine the number of PCs that are statistically significant in capturing the underlying features in X matrix, the leave-one-out cross validation method (Eriksson et al., 2001) was implemented. All computations were performed using the PLS Toolbox 5.2 (Eigenvector Research, Inc., Manson, WA) within the MATLAB 7.8.0 (R2009a) computational environment (MathWorks, Natick, MA).

### 3. MODEL DEVELOPMENT AND OPTIMIZATION

The PCs extracted from the fluorescence data (matrix X) during UF were found to be respectively correlated to different NOM foulants such as humic-like, protein-like and particulate/colloidal matter present in water. Thus, the scores of each PC can be thought as a qualitative measurement of the corresponding foulant component. The evolution of the PC scores ($s_i$) that correspond to these PCs during the UF of water is related to the membrane fouling behaviour (Peiris et al., 2010). Therefore the PC scores ($s_i$) associated with the retentate and permeate of the UF process were used to model the fouling behaviour of the UF membrane operation.

#### 3.1 PC-based Modeling of Membrane Fouling

Since it is impossible at this point to identify the individual foulant species in natural water, it was decided instead to perform a balance on the PC scores which are physically related to groups of foulants as mentioned above. Accordingly, the accumulation of NOM foulants on the surface and in the pores of the membrane was calculated based on the PC score balance for a given group of foulant which is analogous to a mass balance performed on the control volume of the solution occupied by the membrane. The accumulation of the NOM foulant ($j$) that contributes to fouling can therefore be represented as follows:

$$\frac{ds_{j,M}}{dt} = \frac{1}{kV_M} \left[ (1 - \omega)A \frac{\Delta P}{\mu R_i}(s_{j,R} - s_{j,P}) - wL_j \right]$$  

(2a)

$$L_j = m_{\text{wash}} \text{eff} j e^{-qR_j} s_{j,M}$$  

(2b)

for $j=1, 2, 3, ..., N$ and $w = 0 \text{ or } 1$

Where $s_i$ is the PC score related to the $j^{th}$ NOM foulant. $N$ is the number of PCs generated by PCA which are statistically significant and deemed to be important for capturing the information related to the major groups of NOM foulants as explained in section 2.3. Subscripts R, P and M denote retentate, permeate and the membrane, respectively. $V_M$ is the volume of the solution occupied by the membrane and $k$ is a parameter that specifies the active portion of $V_M$ (i.e. actual portion of $V_M$ that participates in the filtration). The membrane area, TMP and the water viscosity are represented by symbols $A$, $\Delta P$ and $\mu$ respectively. $m_{\text{wash}}$ is the mass flow rate used for periodic membrane back-washing, $w$ is a binary variable that models permeation through the membrane ($w=0$) or back-washing ($w=1$). eff represents the efficiency at which the $j^{th}$ NOM foulant fraction (i.e. $j^{th}$ PC) is removed during the back-washing. $q$ is a parameter describing the decay of efficiency in back-washing over time due to irreversible fouling due to the $j^{th}$ NOM foulant; accumulated membrane foulant material that cannot be removed by membrane back-washing results in irreversible membrane fouling. $R_i$ is the membrane resistance at time $= t$, which is given in terms of the scores as follows:

$$R_i = R_0 + \sum_{j=1}^{N} \beta_j s_{j,M} + \beta_{\text{inter}} S_{\text{protein},M} \times S_{\text{coll./partic.},M}$$  

(3)

$R_0$ is the initial membrane resistance of the membrane before fouling occurs. $\beta_j$, $j=1, 2, 3, ..., N$ are the model parameters. $\beta_{\text{inter}}$ is also a model parameter related to the interaction between protein and colloidal/particulate matter (represented by $S_{\text{protein},M}$ and $S_{\text{coll./partic.},M}$ respectively) that contributes to membrane fouling. The existence of this interaction was found to be significant in a separate correlation analysis study (results not shown for brevity) and found to be very important for improving the model predictions in this study.

Also, from continuity considerations, the net NOM foulant transfer rate across the membrane was assumed to be equal to the back diffusion rate of foulants from the membrane foulant layer to the bulk retentate phase which can be expressed by:

$$D_j (s_{j,R} - s_{j,M}) = \frac{\Delta P}{\mu R_i} (s_{j,R} - s_{j,P})$$  

(4)
Where $D_j$ is the molecular diffusivity of the $j^{th}$ NOM foulant fraction.

The permeate water flux through the membrane at time $t = 1$ is as follows:

$$J_t = \frac{\Delta P}{\mu R_j}$$  \hspace{1cm} (5)

### 3.2 Model Calibration and Validation

Experimental permeate water flux data obtained by UF runs performed using GRW with different DOC content and turbidity values within the ranges as indicated above were used to calibrate the model given by the system of equations (2a), (2b), (3), (4) and (5). The model calibration involved the estimation of the model parameters $k$, $\beta_1$, $\beta_2$, $\beta_3$, ..., $\beta_N$, $\beta_{\text{inter.}}$, $\epsilon_{R}$, $\epsilon_{F}$, $\epsilon_{\text{eff.}}$, ..., $\epsilon_{\text{eff.}}$ and $q$. This was achieved by minimizing the sum of squares error (SSE) between experimental and model predictions of permeate water flux by using the MATLAB function “ga”, a genetic algorithm code available within the MATLAB 7.8.0 computational environment. The model predictions were generated by solving the above state space model based on the fluorescence EEMs of retentate and permeate captured at time $t = 15$ min of the UF experiments using MATLAB ordinary differential equation (ODE) solver “ode23.” PC scores ($S_{R, 15 \text{min}}$ and $S_{P, 15 \text{min}}$) that are related to these fluorescence measurements were used for the estimation of the predicted permeate water flux into the future, i.e. for a total time horizon of 4 h. The model validation was achieved using additional experimental permeate water flux data that were not used in the calibration.

### 3.3 Optimization of the UF process

The predicted permeate water flux can be used to understand the fouling of the membrane and reduced permeate water flux occurring over time for constant TMP operations (as demonstrated in this study). On the other hand, if constant permeate flux is desired, the TMP would increase as a result of fouling. In both situations, membrane fouling results in an increase in the energy requirement per unit amount of drinking water produced.

In this study, the UF membrane back-washing times were manipulated to optimize the UF process so that the energy requirement was minimized for the production of a unit amount of drinking water. This optimization approach was implemented by minimizing the following objective function (OF), (equation 6) subjected to the constraints listed in equations (9) and (10).

$$OF = \frac{\text{Energy consumption}}{\text{Water production}}$$  \hspace{1cm} (6)

Where energy consumption and the water production for time duration $\Delta t$ is given by:

$$\text{Energy consumption} = \frac{A(\Delta P)^2 \Delta t}{\mu R_j}$$  \hspace{1cm} (7)

$$\text{Water production} = J_t A \Delta t$$  \hspace{1cm} (8)

$$\begin{bmatrix}
1 & -1 & 0 & 0 \\
0 & 1 & -1 & 0 \\
0 & 0 & 1 & -1 \\
0 & 0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
t_1 \\
t_2 \\
t_3 \\
t_4
\end{bmatrix} \leq \begin{bmatrix}
-t_w \\
-t_w \\
-t_w \\
(t_F - t_w)
\end{bmatrix}$$

$$t_1 \geq t_d$$  \hspace{1cm} (10)

Where $t_1$, $t_2$, $t_3$, and $t_4$ are the times at which the back-washing of the UF membrane was implemented. The number of back-washing cycles was limited to four as it was sufficient to demonstrate the application of the proposed approach. The number of back-washes could be another parameter that could be included in this optimization approach; this will be addressed in future research.

Also, $t_w = 180$ s indicates the time that was needed for back-washing (20 s) plus the time required to connect and disconnect the Nitrogen gas supply for back-washing and adjusting the TMP of the UF membrane cell holder (160 s), which were performed manually. The total filtration time is indicated by $t_F (= 257$ min) and $t_d (= 15$ min) is the time at which the first set of fluorescence EEMs of the retentate and permeate for UF operation were obtained. This information was required for the model predictions as explained in section 3.2. The minimization of the OF (equation 6) subjected to the constraints (equations 9 and 10) were performed by using the MATLAB function “ga”- a genetic algorithm code available within the MATLAB 7.8.0 computational environment. “ga” was better able to handle the non-linear and non-smooth (due to back-washing) nature of the OF compared to other optimization approaches, and was therefore selected.

### 4. RESULTS AND DISCUSSION

#### 4.1 Typical fluorescence EEM features of GRW

![Fig. 2. Typical spectral features seen in the fluorescence EEM for GRW. First order Rayleigh scattering (RS) regions are indicated using dashed-lines.](image-url)
The spectral regions (α) and (β), indicated in the fluorescence EEM of GRW (Fig. 2), are representative of the presence of humic-like NOM (Coble et al., 1990; Peiris et al., 2008). The region (δ) is related to the presence of protein-like NOM in GRW (Peiris et al., 2010). Rayleigh scattering regions observed in the fluorescence EEM also provide information related to the particulate/colloidal matter present in water (Peiris et al., 2010).

During the UF of GRW, subtle changes in spectral information related to these NOM membrane foulants in both retentate and permeate are expected and related to UF membrane fouling. PCA was used to extract these changes in terms of the PC scores.

4.2 PCA of fluorescence data

PCA analysis was performed on X matrix that contained fluorescence data of 340 samples (retentate and permeate) that were obtained from 10 UF experiments with different feed water conditions as explained in section 2.3. This generated new and fewer numbers of variables or PCs that captured systematic trends present in the 4214 original spectral variables in the X matrix. The first four PCs alone, generated in this way, were able to capture nearly 90% of the total variance (i.e. PC – 1: 63.0 %, PC – 2: 16.4 %, PC – 3: 5.5 % and PC – 4: 4.7 %), present in the original spectral variables in X matrix. The remaining variance (~10%) is due to the combination of unexplained variance by the first four PCs and the instrumental noise in the fluorescence measurements. None of the additional PCs were statistically significant (<2% variance captured) and were not found to be related to membrane foulant fractions present in water as explained below.

The first four PCs were found to be related to different NOM membrane foulant fractions present in water; PC – 1, PC – 2, and PC – 3 were related to humic-like, colloidal/particulate and protein-like substances, respectively. PC – 4 was also found to be related to colloidal/particulate substances. This was verified by examining the loading plots corresponding to each PC, (generated from the loading values, i.e. \( \mathbf{P} \) values) as demonstrated in Peiris et al. (2010). For example, the loading peak of PC – 1 appeared in the same location where the fluorescence EEM regions related to humic-like NOM. Similar observations were made with PC – 2, PC – 3 and PC – 4 in relation to the NOM foulant fractions they represent (results not shown for brevity). The PCA model developed using these four PCs was then used to calculate the PC scores (i.e. \( s_1, s_2, s_3 \) and \( s_4 \)) that were used as states in the proposed PC-based dynamic fouling model explained in section 3.1.

4.3 Model Predictions

Fig. 3 demonstrates the model predictions and the experimentally measured permeate water flux of selected UF experiments that were not used in the model calibration. These experiments cover low, medium and high membrane fouling situations. The model predictions for these experiments were obtained using only the fluorescence-based PC scores of retentate and permeate obtained at time = 15 min of the UF. The prediction results indicate that the model was able to successfully predict different membrane fouling behaviours experienced by the UF operations. The root-mean-squared error between predictions and experimental values for high, medium and low fouling situations were 0.09, 0.07 and 0.08, respectively.

The results presented indicate that the proposed fluorescence-based membrane fouling modeling approach can be used for forecasting different fouling behaviours corresponding to changes in the membrane feed water quality. Therefore, this approach has application in the process control of membrane-based drinking water treatment systems where the model predictions of high membrane fouling situations could be detected well in advance and thus appropriate process optimization measures implemented to ensure sustainable operation of the treatment process.

4.4 Optimization of UF for Drinking Water Treatment

This section illustrates the application of the proposed modeling approach for the optimization of UF operations for drinking water applications. The model predictions of the fouling behaviour for UF of GRW (obtained on Oct. 25, 2009 and pre-filtered) with back-washing at normal time intervals (i.e. 1 hr) are shown in Fig. 4. When the back-washing times were optimized using the proposed optimization approach (section 3.3) the model predictions indicated an energy savings of 3.7% with a 4.3% increase in the total volume of drinking water production. The back-washing times generated by the optimization approach were \( t_1 = 61 \) min., \( t_2 = 90 \) min., \( t_3 = 118 \) min., and \( t_4 = 137 \) min. These model predictions were also experimentally validated (Fig. 4).

The optimization approach used in this study was limited to four back-washing cycles; however, it is possible to further improve the energy savings and the water production by employing additional back-washing cycles to the optimized conditions. For example, when two additional back-washing cycles were included as illustrated in Fig. 4, the model predictions indicated an increase in the energy savings and
the volume of drinking water production up to ~ 8.0% and ~ 9.8% respectively. The use of additional back-washing cycles will also limit the high fouling behaviour of the membrane that may occur even with the optimized back-washing (as was the case in this study). In addition, it would further extend the life span of the membrane and minimize the need for chemical cleaning to recover flux decline caused by irreversible fouling. Current research is investigating the use of the number of back-washing cycles as another optimization parameter.

**5. CONCLUSIONS**

The fluorescence-based membrane fouling model developed in this study was suitable for accurately predicting different fouling situations caused by changes in membrane feed water quality. The ability of this approach to forecast membrane fouling based on the fluorescence EEM measurements, captured at time = 15 min of the UF operation, should allow sufficient time for fouling control strategies to be implemented. This is especially applicable for forecasting high fouling events that are often harmful for membranes or challenging for the efficient production of drinking water to meet consumer demand. The potential of this approach for process optimization would be very useful for the sustainable operation of membrane-based drinking water treatment facilities in terms of minimizing the energy spent on a unit amount of drinking water produced.

**ACKNOWLEDGEMENTS**

The authors wish to acknowledge funding from the Canadian Water Network and the Natural Sciences and Engineering Research Council of Canada (NSERC), including an NSERC Postgraduate scholarship to R.H. Peiris.

**REFERENCES**


