

Desalination and Water Treatment www.deswater.com 15 (2010) 121–128 March

1944-3994 / 1944-3986 © 2010 Desalination Publications. All rights reserved.
doi: 10.5004/dwt.2010.1675

Prediction of reverse osmosis membrane fouling due to scale formation in the presence of dissolved organic matters using genetic programming

Jae-Seok Cho^a, Hwan Kim^b, June-Seok Choi^a, Sangho Lee^{a*}, Tae-Mun Hwang^a, Hyunje Oh^a, Dae Ryook Yang^c, Joon Ha Kim^d

^aKorea Institute of Construction Technology, 2311 Daehwa-Dong, Ilsan-gu, Kyonggi-do, Korea Tel. +82 (31) 910-0320; Fax +82 (31) 910-0295; email: s-lee@kict.re.kr

^bUniversity of Science & Technology, Daejeon, 305-333, Korea

^eDepartment of Chemical & Biological Engineering, Korea University, Seoul, 136-701, Korea

^dDepartment of Environmental Science and Engineering, Gwangju Institute of Science and Technology (GIST), Gwangju, 500-712, Korea

Received 12 November 2009; Accepted in revised form 24 December 2009

ABSTRACT

Scale formation of soluble salts is one of the major factors limiting the performance of reverse osmosis (RO) membranes for desalination. However, it is difficult to predict membrane fouling due to scale formation in a complicated feed water containing dissolved organics such as humic substances. This study aims at prediction of the complicated fouling phenomenon by scale formation in the presence of dissolved organic matters. Experimental studies with model solutions were conducted in a small batch filtration device. Humic acid and calcium sulfate were used as model dissolved organic matters and scale-forming salts. A genetic programming technique was applied to predict the effect of dissolved organic matters on scale formation.

Keywords: Desalination; Scale formation; Dissolved organic matters; Humic substances; Fouling; Genetic programming

1. Introduction

Reverse osmosis (RO) membrane process has been considered a promising technology for water treatment and desalination. RO membrane removes ions and organic chemicals with lower energy consumption than other competing technologies such as evaporation and electrodialysis. In addition, its treatment efficiency and performance are stable as well as predictable. RO has been proven to be adequate for producing pure water in various applications [1,2].

Since scale formation is a serious constraint in designing and operating RO systems, elaborate models to simulate these scale and fouling problems are required to help the design engineer to predict the effects of such problems on the performance of the RO plants. Unfortu-

Nevertheless, seawater and brackish water always have the tendency for scale formation and fouling problems due to dissolved salts and finely suspended solids. Scales are hard mineral deposits that precipitate from the feed stream onto the membrane surface. The scaleforming salts are ubiquitous in most water environments, including calcium sulfate (CaSO₄), calcium carbonate (CaCO₃), and silica (SiO₂) [3].

^{*} Corresponding author.

Presented at SeaHERO Academic Workshop on Development and Optimization of SWRO Desalination Systems Korea University, Seoul, Korea, November 13–14, 2009

nately, scale formation is complex and hard to predict in RO systems. Two different mechanisms of scale formation are involved, including such as surface (heterogeneous) crystallization and bulk (homogeneous) crystallization. Concentration polarization plays an important role in scale formation in membrane systems [4,5]. Other factors affect the crystallization process such as pH [6], temperature [7], and the presence of dissolved organics [8]. Although mechanistic models are useful to understand scale formation mechanisms, they have limited ability to predict fouling due to scale formation in practical applications.

In this study, we developed a data-driven model to explore the effect of $CaSO_4$ scale formation on RO membrane process. The genetic programming (GP) technique was applied to overcome the limitation of mechanistic models. Using the model, we investigated the effect of background organic matters on RO membrane fouling due to scale formation. The relative importance of each operating parameter was also analyzed using this modeling approach.

2. Theory

2.1. Mechanistic model for RO fouling due to scale formation

Scale formation is a complex process in which two pathways for crystallization are involved. This can be expressed using the resistance-in-series model and crystallization kinetic theory. Based on this approach, the permeate flux is given by combining the surface blockage and cake filtration models [9]:

$$J = L_v (\Delta P - \pi) = \frac{\Delta P - \pi}{\eta (R_m + R_c)} \times \frac{A - A_b}{A}$$
(1)

where L_v is the solvent transport parameter; ΔP is the transmembrane pressure; π is the osmotic pressure; η is the permeate viscosity; R_m is the membrane resistance; R_c is the resistance due to cake formation; A is the membrane area; and A_b is the membrane area occupied by surface crystals. Here, A_b and R_c are given by [4,9]:

$$A_b = \frac{\beta m_s}{A} \tag{2}$$

$$R_c = \frac{\alpha m_c}{A} \tag{3}$$

where β is the area occupied per unit mass; m_s is the weight of scale formed directly on membrane surface; α is the specific cake resistance; and m_c is the accumulated weight of precipitated scale.

The rate of scale formation depends on both nucleation and crystal growth kinetics. The rate of nucleation is related to the induction time (τ), which is defined as time to induce formation of detectable crystals. This can be calculated using the kinetic information on the crystal [9]. Once the nucleation occurs, crystals grow on the surface of membrane through surface crystallization and/or bulk crystallization. The surface and bulk crystal growth rate of scale-forming salt can be written as [5,9]:

$$\frac{dm_s}{dt} = k_s \left(A - A_b\right) \left(c_w - c_s\right)^n = k_s A \left(A - \frac{\beta m_s}{A}\right) \left(c_w - c_s\right)^n \quad (4)$$

$$\frac{dm_c}{dt} = k_c s_p \Psi \left(c_b - c_s \right)^m = k_b \left(c_b - c_s \right)^m \tag{5}$$

where k_s is the rate constant of surface crystallization; c_s is the saturation concentration; n is the order of reaction rate; k_c is the rate constant of bulk crystallization; s_p is surface area of active sites on bulk crystals; c_b is the bulk phase concentration; ψ is the deposition probability of crystal particles; m is the order of reaction rate; and k_c is the apparent rate constant of bulk crystallization (= $k_c s_p \psi$).

The mechanistic model allows better understanding of crystallization mechanisms in RO systems. However, it is difficult to obtain all kinetic and filtration parameters under various conditions, leading to poor prediction of experimental data using the model. Moreover, the kinetic parameters may be changed if the feed solution contains dissolved organics. This is why data-driven models are required to predict and control scale formation and fouling in RO systems.

2.2. Application of genetic programming

Genetic programming (GP) is an evolutionary algorithm-based methodology inspired by biological evolution to find computer programs that perform a user-defined task [10]. GP evolves computer programs represented in memory as tree structures, which can be easily evaluated in a recursive manner. Every tree node has an operator function and every terminal node has an operand, making mathematical expressions easy to evolve and evaluate.

Using GP, a model to predict the complicated phenomena can be developed if experimental data is enough to evolve (or train) it. On the other hand, it is difficult to find physical meaning of model structures. Thus, GP models may not be used for fundamental studies but process control and simulation. The following steps are involved for developing the optimum model based on GP algorithm [11]:

- 1. Initialize the population: A GP system (a software tool to make GP models) creates a population of programs randomly.
- Run a tournament: The GP system picks four programs randomly out of the population of programs. It compares them and picks two winners and two losers based on fitness.
- 3. Apply the search operators: The GP system then applies search operators like crossover and muta-

tion to the winners and produces two "Children" or "Offspring."

- Replace the losers. After the search operators have been applied to the copies of the winners (the offspring), these offspring replace the two losers in the tournament. The winners of the tournament are unchanged.
- 5. Repeat until Termination. The GP system then repeats steps 2 through 4 until the run is terminated.

In this study, a commercial GP system, Discipulus (RML Technologies, USA), was used to make models for simulating scale formation in RO systems.

3. Experimental

The test system shown in Fig. 1 was used to measure filtration characteristics for RO membrane. The stirred cell was made of stainless steel to improve chemical stability. The diameter of the stirred cell was 54 mm and the working volume was 100 ml. A magnetic stirrer (Stirrer assembly 8200, Millipore, USA) was positioned just above the membrane. The length of the stirring bar was 52 mm. The working pressure was controlled by a high pressure nitrogen cylinder and by a gas pressure regulator. The stirring speed was controlled by a magnetic stirrer plate. The temperature of the feed solution was adjusted to 20–25°C and the effect of temperature on viscosity as well as density was corrected. Since the experiment was performed in a short time (normally less than 30 min), the variations of the temperature during an experiment were smaller than ±1°C.

A commercially available RO membrane (Filmtec, USA) of the thin film composite (TFC) type was used for the filtration tests. The pure water flux was measured to be 60–70 L/m²-h at 1000 kPa and the ion rejection was over 98%. All experiments were performed using new membranes. Prior to the filtration all membranes were thoroughly cleaned to remove remaining organic and



Fig. 1. Schematic diagram of a batch filtration system.

inorganic materials by successively filtering 0.1 M sodium hydroxide and 0.1 M hydrochloric acid solution at 10 bar. Finally the membrane was cleaned with ultrapure water.

The permeate flux was expressed in terms of concentration factor (VCF). The concentration factor, defined as a ratio of the feed volume to concentrate volume, indicates the extent of concentration:

$$VCF = \frac{V_f}{V_c} = 1 + \frac{V_p}{V_c}$$
(6)

where V_{p} , $V_{c'}$, and V_{p} are defined as the volume of feed, concentrate, and permeate, respectively. VCF is proportional to permeate recovery.

Humic acid (Aldrich, USA) was obtained from Aldrich Chemical and used as a model dissolved organic matters (DOM). To prepare the stock solution, the powdered form of humic acid were dissolved in deionized water and filtered by a 0.45 µm filter. The final concentration of humic acid stock solution was set to be 35 mg/L as TOC. The stock solution was kept in a refrigerator at 4°C before use. Hydrochloric acid and sodium hydroxide (Ajax Chemical, Australia) were used to adjust pH of the solution. The concentration of DOM was measured using a total organic carbon analyzer (DC-180, Rosemount, USA), which is based on a persulfate-ultraviolet light oxidation method [12]. To determine the molecular weight of DOM, the feed solution was divided into the four fractions using three cellulose acetate hydrophilic ultrafiltration membranes (Amicon, USA) with different molecular weight cut-off; 30,000, 10,000 and 5,000 Dalton.

A saturated solution of CaSO₄ (2000 mg/L) was used as a model scale-forming salt. Prior to filtration test, the solution was prefiltered using a 0.45 μ m filter. The concentrations of CaSO₄ were determined by an ion chromatography (DIONEX 4000I, USA) and by a conductivity meter (Model 170, Orion, USA). The turbidity measurement for the feed and retentate was made on a turbidimeter (HF, DRT-100B, USA).

4. Results and dscussion

4.1. RO flux decline due to scale formation

A series of experiments were performed to investigate the effect of scale formation on RO flux under various conditions. The details on the experimental matrix are summarized in Table 1. The applied pressure ranged from 6 bar to 15 bar and the stirring speed ranged from 170 rpm to 600 rpm. The range of humic acid concentration was between 0 mg/L and 10 mg/L.

Fig. 2 shows the flux in batch cell filtration of saturated $CaSO_4$ solutions with various humic acid concentrations as a function of time. The operating conditions for each run are given by Table 1. In all cases, the flux decreased with time because $CaSO_4$ in solution crystallized and the crystals fouled the membrane. Nevertheless, the

Table 1	
Experimental	matrix

Run	Applied pressure (bar)	Stirring speed (rpm)	Time (min)	VCF	Humic acid concentration (mg/L)
1	12	600	0–57	1–6.2	0
2	12	600	0–57	1-6.4	0.2
3	12	600	0–53	1–7.1	0.4
4	12	600	0–58	1–6.7	0.6
5	12	600	0-61	1–5.9	0.8
6	12	600	0-52	1-8.4	1
7	12	600	0-52	1–7.9	2
8	12	600	0–58	1–12.3	4
9	6	170	0–116	1-5.7	0
10	6	600	0–97	1-5.1	0
11	6	1000	0-142	1–7	0
12	6	170	0–95	1–7.7	0.5
13	6	170	0–98	1–9.8	1
14	6	170	0-128	1-32.5	2
15	6	170	0–97	1–28	10
16	15	300	0–49	1–6	0
17	15	400	0–45	1–5.2	0
18	15	500	0–45	1–5.6	0

patterns for flux decline are quite different depending on the operating conditions. For instance, the initial flux is high for high applied pressure (12 bar and 15 bar) but the rate of flux decline also is high. The stirring speed and humic acid concentration also affect the rate of flux change with time.

Theoretical models based on the solution-diffusion model and the film theory may be applied to analyze these scale formation phenomena in the RO system. However, it is difficult to obtain all the model parameters required for these models using the experimental data. In addition, these parameters are dependent on the operating conditions such as $CaSO_4$ and humic acid concentrations, leading to a failure of reasonable prediction of new set of experimental data. Thus, a model based on GP was considered and applied to analyze the experimental results in Fig. 2.

4.2. Development of GP model

A GP model was designed to predict RO flux as a function of input parameters including time, VCF, applied pressure, stirring speed, and humic acid concentration. To apply GP algorithm for model development, some of the data set in Table 2 (Run-1, Run-3, Run-5, Run-7, Run-8, Run-10, Run-11, Run-12, Run-15, Run-17, and Run-18) were used for training the model. Then, the model was validated by other data set (Run-2, Run-6, Run-9, and Run-14). The final model code was generated in JAVA (see Appendix). This two-step validation seems to be



Fig. 2. Experimental results for RO test runs under various conditions.

appropriate to develop a predictive model because it can demonstrate how well the evolved model works on data it did not train on.

Fig. 3 shows how the fitness changed with the generation of model programs. The best program should have smallest fitness. The fitness did not decrease significantly over the run number of 10000, suggesting that no more runs are required to have better accuracy.



Fig. 3. Fitness test results for programs generated by GP algorithm.

Fig. 4 and Fig. 5 compare the experimental results with model calculations. The calculations by the present model are in good agreement with the experimental data. The

 R^2 values for model training and validation were 0.951 and 0.954, respectively.

4.3. Application of GP model

In Fig. 6, the model calculations were compared with experimental data for Run-4, Run-13, and Run-16. The model predicts the effect of scale formation on flux well except for Run-13, in which the model slightly underpredicts the flux. Nevertheless, the model captures the overall trend of flux decline. The *R*² value for model application was 0.832.

There are possible reasons for deviation of model from experimental data in Run-13. First, there were some experimental errors, leading to high initial flux than expected. Based on the calculation of membrane resistance, the initial flux for Run-13 should be less than 40 L/m²-h. Second, the model itself has limited ability to predict the experimental data for this condition (low pressure, low concentration of DOM) because the data for model



Fig. 4. Comparison of experimental data for model prediction: training of model.



Fig. 5. Comparison of experimental data for model prediction: model validation.



Fig. 6. Comparison of experimental data for model prediction: model application.

training was insufficient. It is likely that the selection of experimental set for model tanning is also important.

4.4. Impact of input parameters

Using the model, the relative importance of each input parameter can be analyzed. Table 2 lists the impact of input parameters on permeate flux. The frequency shows how many times each input appears in a way that

Table 2 Impact of input parameters

contributes to the fitness of the programs. The impact shows the relative importance of the input parameter. A value of 1.0 represents the largest impact value possible. It should be noted that the impact is defined as the relative importance of input parameters on permeate flux rather than the scale formation rate. Thus, the impact of an input parameter indicates how much it affects not only the absolute value of flux but also the changes in flux (or membrane fouling).

It is evident from Table 2 that VCF and applied pressure have higher impact than other input parameters (total impact of these two parameters is 0.88). These results qualitatively match with the mechanistic model shown in Eqs. (1), (4), and (5), where flux is proportional to the applied flux and inversely proportional to c_b and c_w (which are related to VCF). In addition, the impact of humic acid concentration is substantial compared with stirring speed and time. Humic acid is likely to interfere with the nucleation and growth of crystals, thereby affecting the kinetic properties [8].

5. Conclusions

In this work, a GP model was developed to explore the effect of CaSO₄ scale formation on RO membrane process under complex conditions. Experimental results using a batch RO stirred cell indicated that the rates of flux decline were dependent on the applied pressure, stirring speed, and humic acid concentration. Nevertheless, it appears that the effect of these input parameters on flux is hard to be interpreted by simple mechanistic models. A GP algorithm was applied to develop a model for RO fouling due to scale formation. After training and validation, the model predicts the experimental data well. In the GP model, VCF, applied pressure, and humic acid concentration have higher impact than other input parameters.

Acknowledgements

This research was supported by grants 07seaheroB03-02 and 07seaheroB02-01-02 from the Plant Technology Advancement Program funded by the Ministry of Land, Transport and Maritime Affairs of the Korean government.

	Frequency	Average impact	Maximum impact
Humic acid concentration	1	0.11172	0.1645
Applied pressure	1	0.39201	0.8052
Stirring speed	0.53	0.04118	0.0670
Time	0.23	0.00046	0.00075
VCF	1	0.48631	0.49694
Time VCF	0.23 1	0.00046 0.48631	0.00075 0.49694

References

- R.J. Ray, S.B. McCray and D.D. Newbold, Small-scale membrane systems for the recovery and purification of water. Sep. Sci. Tech., 26(9) (1991) 1155–1176.
- [2] S. Lee and R.M. Lueptow, Reverse osmosis filtration for space mission wastewater: membrane properties and operating conditions. Journal of Membrane Science, 182 (2001) 77–90.
- [3] W. Omar, H. Al-Zoubi and J. Ulrich, Seeded crystallization of calcite and aragonite in seawater as a pretreatment scalecontrol process, a study of supersaturation limits. Desal. Wat. Treat., 3 (2009) 236–240.
- [4] S. Lee, J. Kim and C.H. Lee, Analysis of CaSO₄ scale formation mechanism in various nanofiltration modules. J. Membr. Sci., 163 (1999) 63–74.
- [5] S. Lee and C.H. Lee, Effect of operating conditions on CaSO₄scale formation mechanism in nanofiltration for water softening. Wat. Res., 34(15) (2000) 3854–3866.
- [6] P.G. Klepetsanis and P.G. Koutsoukos, Precipitation of calcium sulfate dihydrate at constant calcium activity. J. Crystal Growth, 98 (1989) 480–486.

Appendix

Model program generated by GP (Java code)

Definitions of function and input variables

Function : flux (L/m²-h) v[0] : humic acid concentration (mg/L) v[1] : applied ressure (bar) v[2] : stirring speed (rpm) v[3] : time (min) v[4] : concentration factor (–) float Function(float [] v) { double [] f=new double[8]; double tmp = 0; boolean cflag = false;

f[0]=f[1]=f[2]=f[3]=f[4]=f[5]=f[6]=f[7]=0;

f[0] += f[0];f[0]=Math.sqrt(f[0]); f[0]+=f[1]; if (cflag) f[0] = f[3]; f[0] += f[0];f[0]+=f[3]; f[0]+=-0.8057427406311035f; f[0]/=v[1]; tmp=f[2]; f[2]=f[0]; f[0]=tmp; if (!cflag) f[0] = f[2]; f[0]*=f[2]; tmp=f[2]; f[2]=f[0]; f[0]=tmp; f[2]-=f[0]; tmp=f[3]; f[3]=f[0]; f[0]=tmp; f[0]+=1.366016626358032f; f[0]*=Math.pow(2,trunc(f[1])); f[0]*=f[3]; f[0]+=f[0];

- [7] D.H. Troup and J.A. Richardson, Scale nucleation on a heat transfer surface and its prevention. Chem. Eng. Commun., 2 (1978) 167–180.
- [8] S. Lee, J.-S. Choi and C.-H. Lee, Behaviors of dissolved organic matter in membrane desalination. Desalination, 238 (2009) 109–116.
- [9] H.-J. Oh, Y.-K. Choung, S. Lee, J.-S. Choi, T.-M. Hwang and J.H. Kim, Scale formation in reverse osmosis desalination: model development. Desalination, 238 (2009) 333–346.
- [10] T.-M. Lee, H. Oh, Y.-K. Choung, S. Oh, M. Jeon, J.H. Kim, S.H. Nam and S. Lee, Prediction of membrane fouling in the pilot-scale microfiltration system using genetic programming. Desalination, 247 (2009) 285–294.
- [11] F.D. Francone, Discipulus: Owner's Manual. 2001: Machine Learning Technologies, Inc.
- [12] APHA, AWWA and WEF, Standard Methods for the Examination of Water and Wastewater. 21th ed., Am. Public Health Assoc., Washington, D.C., 2005.

f[0]*=f[0]; f[0] += f[0];f[0]+=f[0]; tmp=f[1]; f[1]=f[0]; f[0]=tmp; f[0]-=f[1]; tmp=f[3]; f[3]=f[0]; f[0]=tmp; tmp=f[0]; f[0]=f[0]; f[0]=tmp; f[1]*=f[0]; f[0]=Math.sqrt(f[0]); f[0]*=f[0]; if (!cflag) f[0] = f[2]; cflag=(Double.isNaN(f[0]) || Double.isNaN(f[3])) ? true : (f[0] < f[3]);f[0]+=f[2]; cflag=(Double.isNaN(f[0]) || Double.isNaN(f[0])) ? true : (f[0] < f[0]);f[0]/=-1.238061666488648f; f[0]-=f[2]; f[0]/=f[0]; f[0]/=f[1]; cflag=(Double.isNaN(f[0]) || Double.isNaN(f[1])) ? true : (f[0] < f[1]);cflag=(Double.isNaN(f[0]) || Double.isNaN(f[2])) ? true : (f[0] < f[2]);f[0]*=f[0]; f[0]/=f[2]; f[0]=Math.cos(f[0]); f[0]=-f[0];f[0]-=-1.364008665084839f; if (cflag) f[0] = f[1]; f[3]+=f[0]; f[1]*=f[0]; f[0]+=-0.6102392673492432f; tmp=f[3]; f[3]=f[0]; f[0]=tmp; tmp=f[0]; f[0]=f[0]; f[0]=tmp; f[0]-=-1.364777803421021f;

f[0]/=v[4];f[0]*=1.086833715438843f; f[0]+=1.366016626358032f; cflag=(Double.isNaN(f[0]) || Double.isNaN(f[2])) ? true : (f[0] < f[2]);cflag=(Double.isNaN(f[0]) || Double.isNaN(f[3])) ? true : (f[0] < f[3]);f[0]=Math.sin(f[0]); f[3]*=f[0]; f[0]=Math.abs(f[0]); f[0]-=-1.063283443450928f; f[0]/=v[4]; f[0]*=v[0]; tmp=f[3]; f[3]=f[0]; f[0]=tmp; f[2]/=f[0]; f[0]+=1.366016626358032f; tmp=f[3]; f[3]=f[0]; f[0]=tmp; cflag=(Double.isNaN(f[0]) || Double.isNaN(f[3])) ? true : (f[0] < f[3]); f[1]-=f[0]; f[0]=-f[0]; f[0]-=0.1756083965301514f; f[0]*=-0.494312047958374f; if (cflag) f[0] = f[2]; tmp=f[2]; f[2]=f[0]; f[0]=tmp; f[0]+=f[2]; f[0]/=f[0]; f[0]*=Math.pow(2,trunc(f[1])); f[0]*=-1.063283443450928f; if (cflag) f[0] = f[2]; f[0]-=f[1]; if (cflag) f[0] = f[2]; f[0]-=-1.549970149993897f; f[0] + = f[3];f[0]/=v[4]; f[1]/=f[0]; if (!cflag) f[0] = f[0]; f[0]=Math.sin(f[0]); f[0]*=Math.pow(2,trunc(f[1]));

f[0]-=f[1]; f[1]-=f[0]; f[0]-=-1.549970149993897f; f[0]-=-1.364777803421021f; f[0]/=v[4]; f[1]/=f[0]; cflag=(Double.isNaN(f[0]) || Double.isNaN(f[3])) ? true : (f[0] < f[3]); f[0]=Math.sin(f[0]); f[0]=Math.abs(f[0]); f[0]*=f[0]; tmp=f[1]; f[1]=f[0]; f[0]=tmp; f[0]=Math.abs(f[0]); f[1]*=f[0]; f[0]-=f[0]; f[0]=Math.cos(f[0]); f[0]/=f[0]; f[0]/=f[3]; f[0]/=f[3]; f[0]+=f[1]; f[0]-=-1.549970149993897f; f[0]+=0.7790718078613281f; f[0]/=v[4]; if (cflag) f[0] = f[2]; if (cflag) f[0] = f[2]; f[1]+=f[0]; f[0]/=f[3]; f[0]+=1.366016626358032f; f[0]*=v[1]; if (Double.isInfinite(f[0]) || Double.isNaN(f[0])) { f[0]=0; }; return (float) f[0]; }

128