Desalination and Water Treatment
www.deswater.com
43 (2012) 281–290
April

doi: 10.1080/19443994.2012.672201
Taylor & Francis
Taylor & Francis Group

Analysis of reverse osmosis system performance using a genetic programming technique

Seung-Min Park[a], Jihee Han[a], Sangho Lee[a,*], Jinsik Sohn[a], Young-Min Kim[b], June-Seok Choi[b], Suhan Kim[c]

[a]*School of Civil and Environmental Engineering, Kookmin University, Jeongneung-Dong, Seongbuk-Gu, Seoul 136-702, Korea*
Tel. +82 2 910 4529; Fax: +82 2 910 4939; email: sanghlee@kookmin.ac.kr
[b]*Department of Construction Environmental Research, Korea Institute of Construction Technology, 2311 Daehwa-Dong, Ilsan-gu, Kyonggi-do, Korea*
[c]*Department of Civil Engineering, Pukyong National University, Yongdang-Dong, Nam-Gu, Busan, Korea*

Received 25 December 2011; Accepted 10 February 2012

ABSTRACT

Reverse osmosis (RO) membrane process has been considered a promising technology for water treatment and desalination. However, it is difficult to predict the performance of pilot- or full-scale RO systems because numerous factors are involved in RO performance, including membrane scaling, fouling, and deterioration. This study was intended to develop a practical model for the analysis of pilot-scale RO processes. A genetic programming (GP) technique was applied to correlate key operating parameters and RO permeability statistically. The GP model was trained using a set of experimental data from a RO pilot plant with a capacity of 1,000 m^3/day and then used to predict its performance. The comparison of the GP model calculations with the experiment results revealed that the GP model was a useful tool for predicting the efficiency of pilot-scale RO systems. The GP model also allowed the in-depth analysis of RO system performance even under unsteady conditions.

Keywords: Reverse osmosis; Desalination; Genetic programming; Prediction; Model

1. Introduction

Reverse osmosis (RO) processes have been the technology of choice for seawater desalination and wastewater reclamation [1,2] due to their many advantages including low energy requirements, small footprint, modular design, and low water production costs. However, the loss of permeability caused by membrane fouling is still a serious problem in designing and operating RO processes. In addition, membrane deterioration due to irreversible damage of the active layer of the membrane is also a critical problem. Therefore, understanding the causes of RO membrane fouling/deterioration and developing strategies for fouling control are of paramount importance for successful application of RO technology.

Many theoretical techniques have been attempted to analyze RO fouling and found to be successful to investigate fouling mechanisms in laboratory-scale RO systems [3–5]. Nevertheless, they seem to be less useful to predict RO performance in pilot- and full-scale plants. Unlike fouling phenomena in bench-scale RO systems, RO fouling and deterioration in large-scale

*Corresponding author.

plants are complicated and hard to be understood based on mechanistic models.

Accordingly, this research was intended to develop a model for the analysis of the performance of RO processes in pilot-scale systems. A genetic programming (GP) technique was applied to correlate key operating parameters and RO permeability statistically. The GP model was trained using a set of experimental data from a RO pilot plant with a capacity of 1,000 m^3/day and then used to predict its performance. The GP model calculations were compared with the experiment results to examine the usefulness of the GP model for predicting the efficiency of pilot-scale RO systems.

2. Theoretical approach

2.1. Mathematical model for RO performance prediction

Effective operation and maintenance of RO plants requires the analysis of current data and the prediction of future data. Although many rigorous models for RO simulation have been developed, they are not very useful to apply to pilot- or full-scale plants due to the restriction of key parameters for the model calculation. Accordingly, semi-empirical or empirical models based on simple mathematical equations are preferred in pilot or full plants.

In this work, a simple form of the mathematical model was applied to fit the pilot plant data [3]:

$$J = \frac{L_p}{1+\alpha V}(\Delta P - \Delta\pi) \quad (1)$$

where J is the permeate flux; L_p is the water permeability of the membrane; ΔP is the applied pressure; $\Delta\pi$ is the osmotic pressure; α is the empirical constant for fouling or membrane degradation; and V is the total permeate volume before membrane cleaning. The model parameters, α and L_p, were obtained by fitting the model calculations to the experimental data using the following equation:

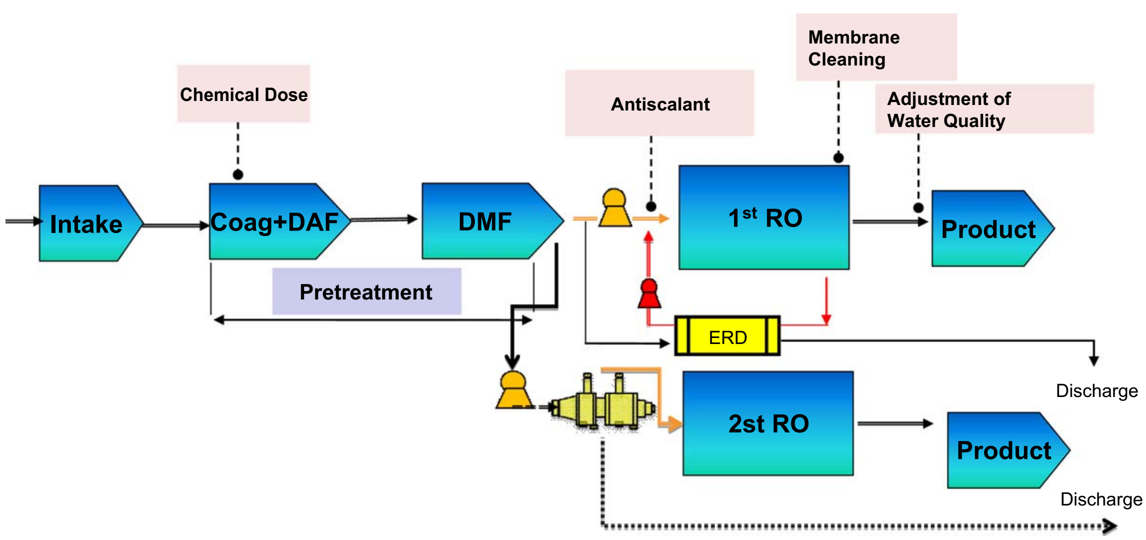

Fig. 1. Schematic diagram of the SWRO desalination pilot plant in Gijang, Korea.

Table 1
Summary of the SWRO desalination pilot plant

Intake	Pretreatment		RO	
Open intake	Coagulation + DAF	DMF	First RO	Second RO
Maximum: 3,000 m^3/day TDS: 35,000 ppm Temperature: 4–25°C	Capacity: 2,500 m^3/day	Capacity: 2,400 m^3/day Turbidity: 0.1NTU SDI: less than three	Capacity: 500 m^3/day Recovery: 40–50% RO: 16 inch ERD: pressure exchanger	Capacity: 500 m^3/day Recovery: 40–50% RO: 16 inch ERD: turbo charger

$$s = \frac{(\Delta P - \Delta\pi)}{J} = \frac{1 + \alpha V}{L_p} \quad (2)$$

Thus, the linear regression of V and s gives the slope ($=\alpha/L_p$) and the intersection ($=L_p$).

The effect of temperature (T) was considered by introducing the concept of temperature correction factor (TCF):

$$\mathrm{TCF} = e^{\left(2640\left(\frac{1}{298.15} - \frac{1}{T+273.15}\right)\right)} \quad (3)$$

$$L_p = L_{p,0}\mathrm{TCF} \quad (4)$$

Where $L_{p,0}$ is the water permeability of the membrane at 273.15 K.

Clearly, mathematical models such as Eqs. (1)–(4) allow better understanding of RO systems. However, it is difficult to consider complex phenomena in pilot- or full-scale systems, leading to poor prediction of operation data using these models. Moreover, the model parameters, which are empirically determined, may be changed with time. This is why data-driven models are required to predict the operation data in pilot or full RO plants.

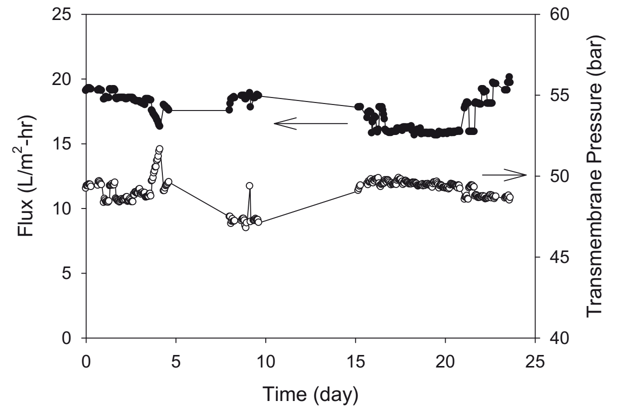

Fig. 3. Flux and transmembrane pressure for the SWRO pilot plant.

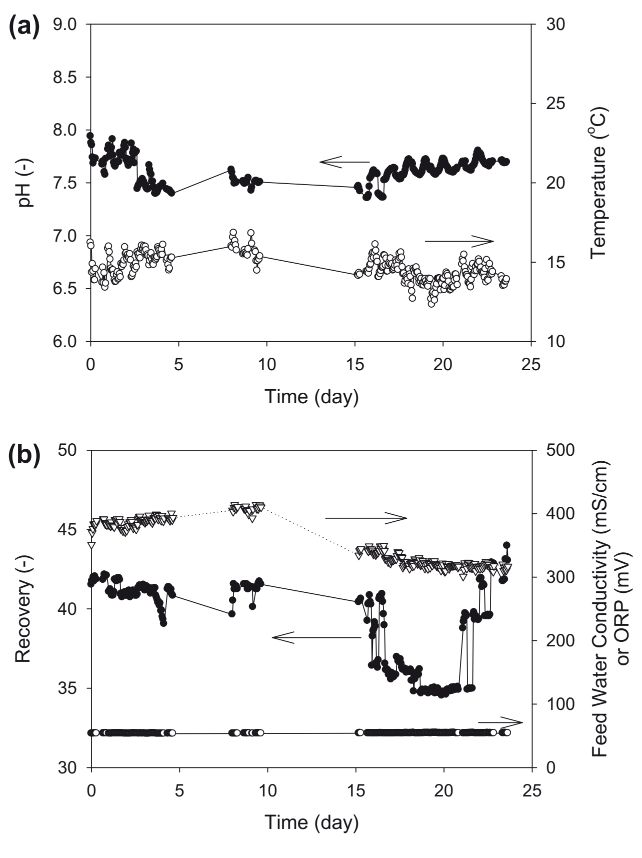

Fig. 2. Characteristics of feed water for the SWRO pilot plant. (a) pH and temperature and (b) conductivity, ORP, and recovery.

2.2. Application of GP

GP is an evolutionary algorithm-based methodology inspired by biological evolution to find computer programs that perform a user-defined task [6]. 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 are 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:

1. *Initialize the population*: A GP system (a software tool to make GP models) creates a population of programs randomly.
2. *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 mutation to the winners and produces two "Children" or "Offspring."
4. *Replace the losers*: After the search operators have been applied to the copies of the winners (the off-

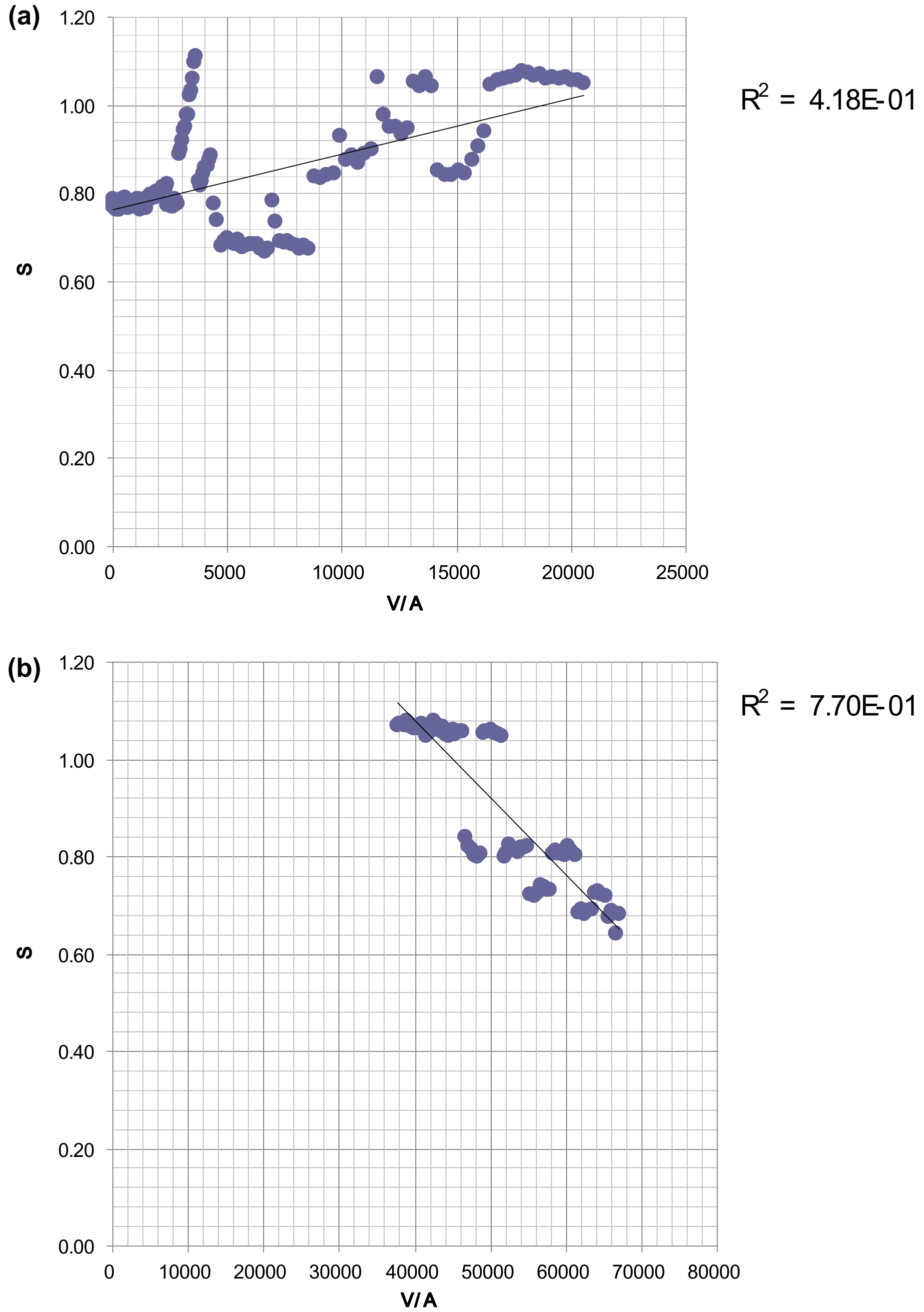

Fig. 4. Relations between V/A and S. (a) early stage and (b) late stage.

spring), 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 2–4 until the run is terminated.

In this study, the GPLAB (ECOS, Portugal) was used to make models for simulating membrane permeability in RO systems [7].

3. Materials and methods

3.1. Pilot plant

A schematic diagram of the pilot plant used for the field tests is shown in Fig. 1. The system consists of an intake facility, pretreatment process, and RO units. Coagulation combined with dissolved air floatation and dual media filter was used as the pretreatment to RO. Suspended solids and colloidal substances are agglomerated by dosing coagulant (Alum) upstream of

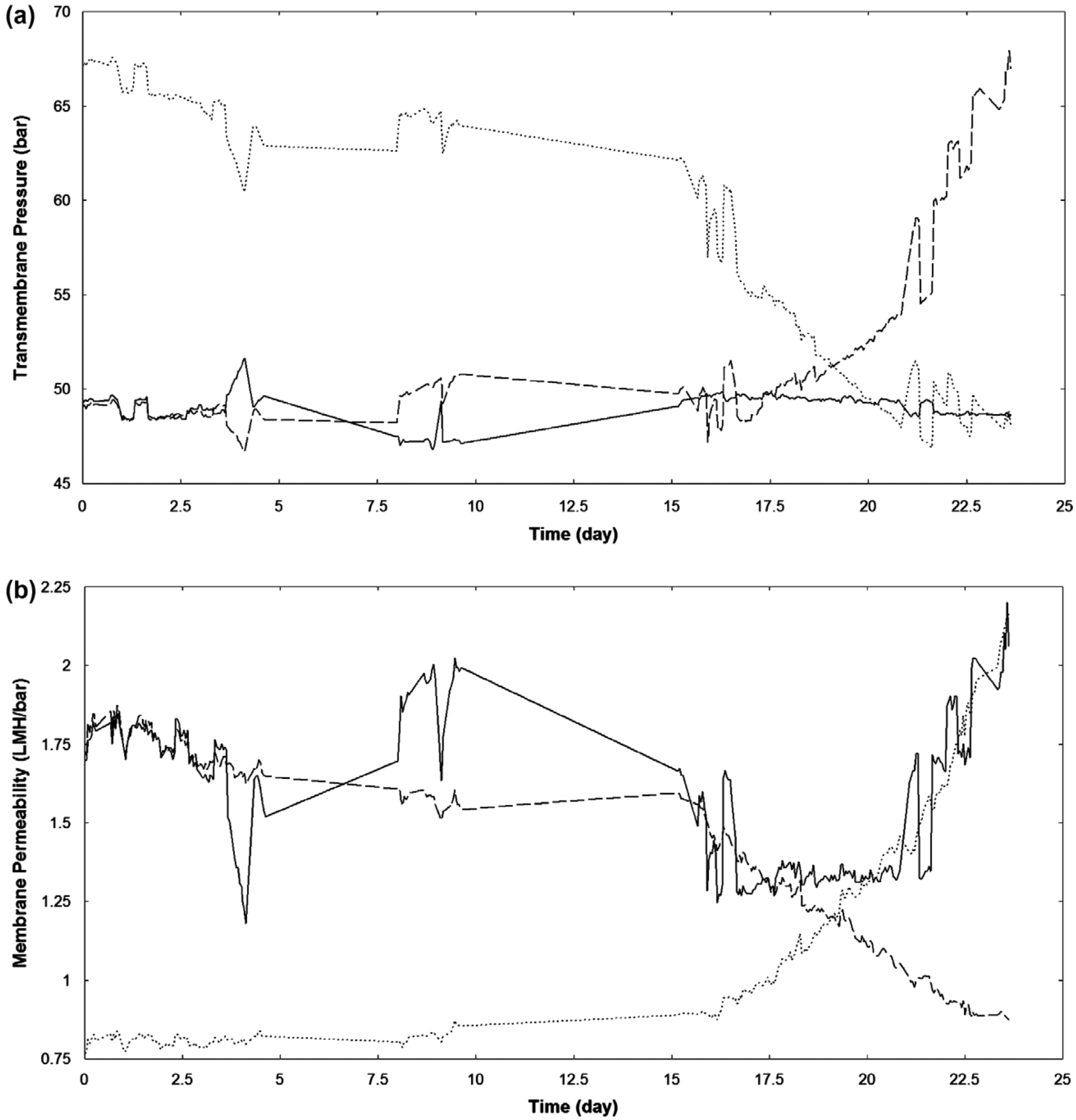

Fig. 5. Comparison of mathematical model fits with pilot plant data. (a) Transmembrane pressure and (b) Membrane permeability (—: pilot plant data; –: mathematical model fitted to the early-stage data; ···: mathematical model fitted to the late-stage data).

the dissolved air flotation (DAF). Turbidity, algae, flocculated colloidal matters, and suspended solids were removed in DAF and dual media filter (DMF), with targets below 0.1 nephelometric turbidity units (NTU) and silt density index (SDI) of 4.

The RO units were comprised of two pressure vessels containing 16 inch RO elements. Each vessel had eight RO elements, allowing 1,000 m^3/day of maximum permeate flow from the pilot plant. The first and second RO units used different RO membranes from different manufacturers and the energy recovery devices were different. Table 1 summarised the key design parameters for the pilot plant.

3.2. Data acquisition and analysis

All the data from the instruments in the pilot plant were automatically stored in a database. Water quality parameters such as conductivity, pH, oxidation reduction potential (ORP) and temperature were obtained from these online sensors. Flux and transmembrane pressure were determined from the readings of pressure sensors and flow meters. The data were converted as an Excel file for convenient handing. The analysis was carried out using the operation data during 24 days.

4. Results and discussion

4.1. Feed water characteristics and operation data

The information on the feed water quality is summarized in Fig. 2. The pH ranges from 7.3 to 8.0 and the temperature ranges from 12 to 16°C. There were no significant variations in the feed water conductivity and ORP. At the same period, the permeate flux ranges from 16 L/m^2-h to 21 L/m^2-h due to unstable operations by the worker (Fig. 3). This led to variations in the permeate recovery. Transmembrane pressure was changed from 47 to 52 bar. It should be noted that the flux and transmembrane pressure showed different trends. For example, on the fourth day, the flux decreased but the transmembrane pressure increased. On the 21st day, however, the flux increased but the transmembrane pressure decreased. This suggests that the membrane permeability was not constant during the operation.

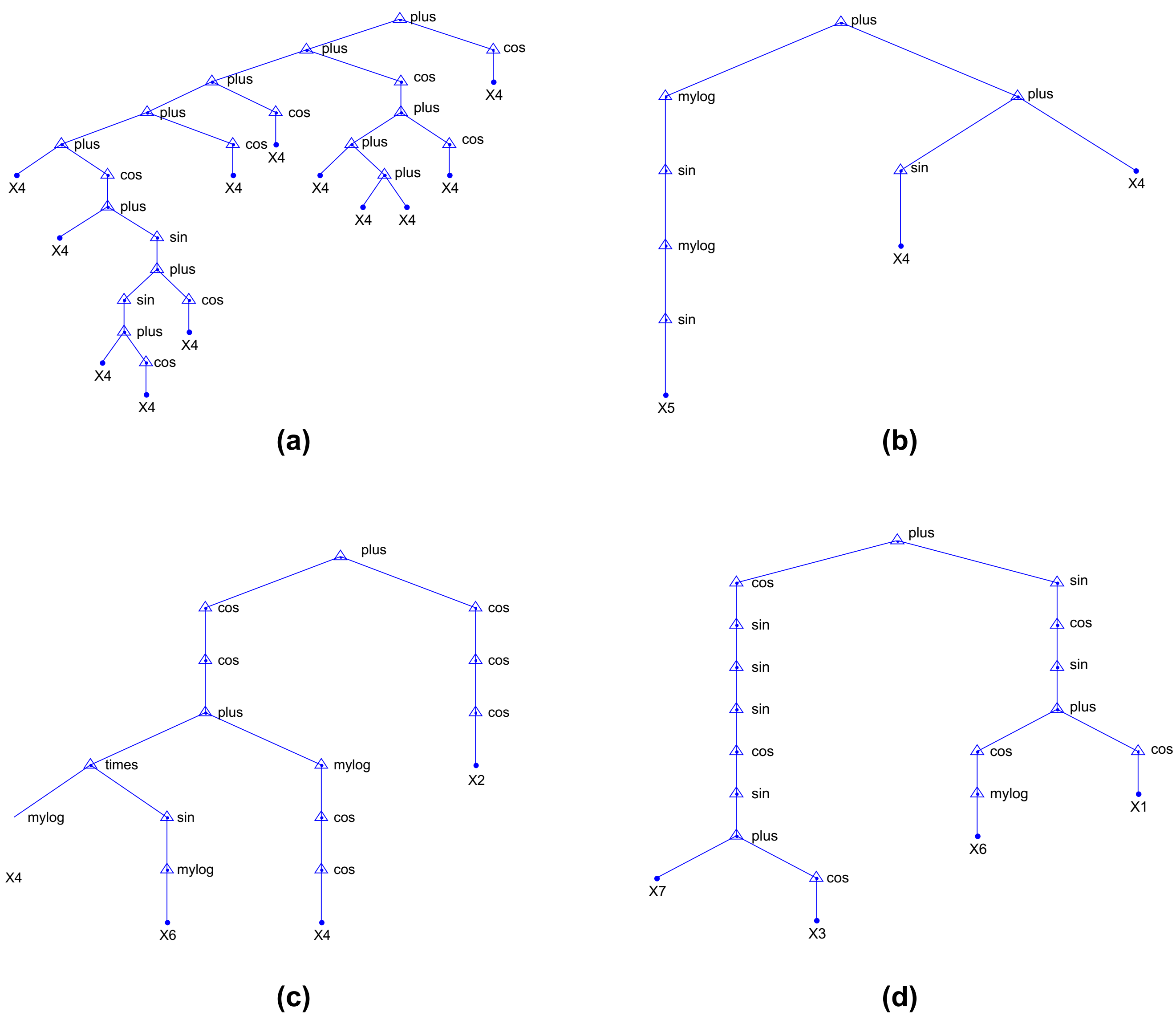

Fig. 6. Structures of GP models. (a) GP model fitted for transmembrane pressure data in the early stage, (b) GP model fitted for transmembrane pressure data in the late stage, (c) GP model fitted for membrane permeability data in the early stage, and (d) GP model fitted for membrane permeability data in the late stage.

4.2. Application of mathematical model

To begin, the empirical model based on the mathematical model for RO transport was applied. Since the early-stage and the late-stage showed different trends of the flux and transmembrane pressure, the model fits were obtained using the early-stage data (0–5 days) and the late-stage data (20–24 days). Fig. 4 shows the model fits to determine the empirical parameters for the mathematical models. For the first fit, L_p and α were determined as $1.31\,L/m^2$-h-bar and 1.66×10^{-5}, respectively. For the second fit, L_p and α were determined as $5.85 \times 10^{-1}\,L/m^2$-h-bar and -9.24×10^{-6}, respectively.

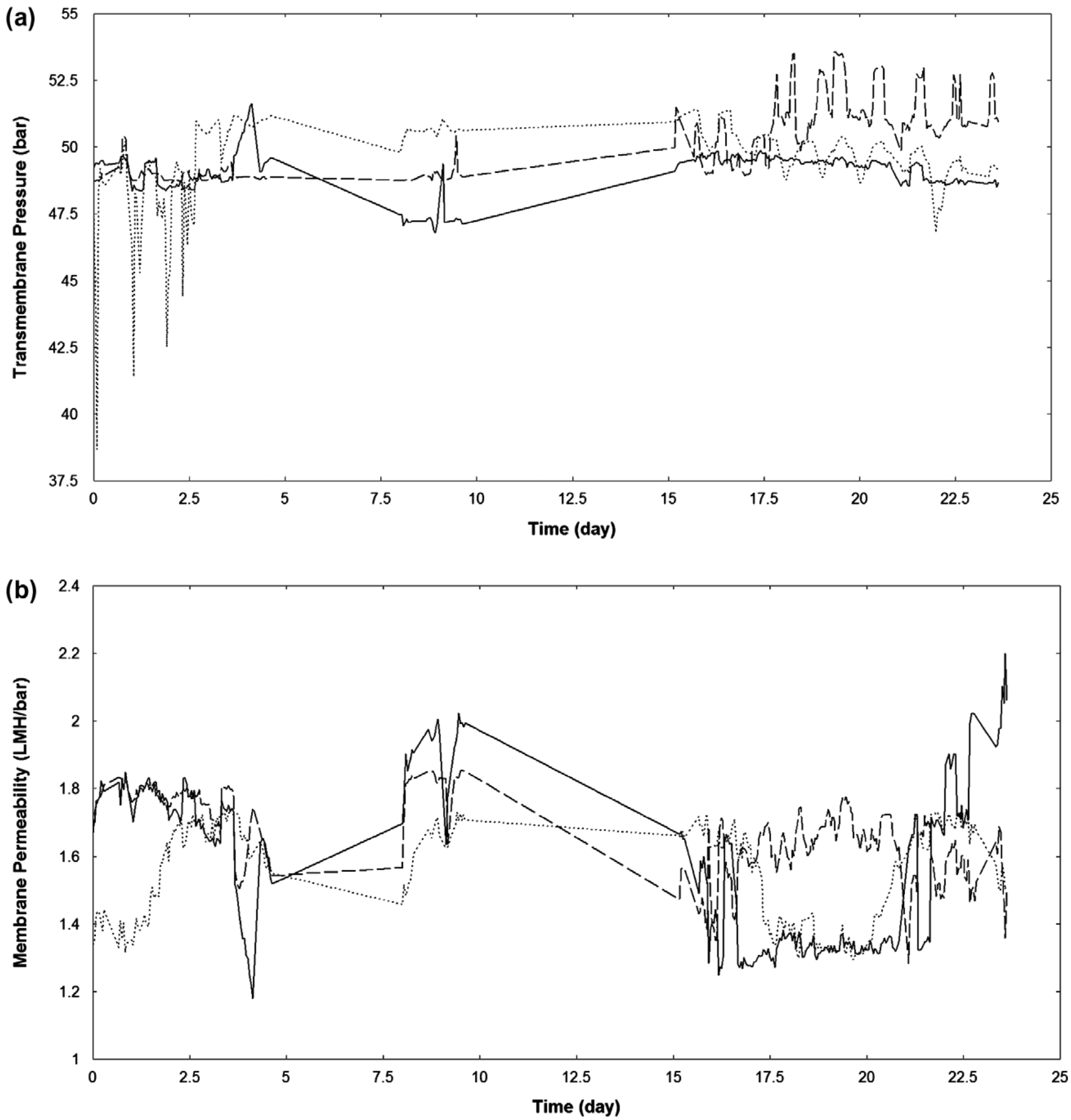

Fig. 7. Comparison of GP model fits with pilot plant data. (a) transmembrane pressure and (b) membrane permeability. (—: pilot plant data; –: mathematical model fitted to the early-stage data; ···: mathematical model fitted to the late-stage data).

As demonstrated in Fig. 5, the mathematical model cannot fit the pilot plant data. In case of the first model fit, the early-stage operation data can be explained but the late-stage operation data could not be matched with the model calculations. The second model fit failed to interpret the early-stage operation data and exhibited large deviations from the pilot plant data. Both transmembrane pressure and membrane permeability showed similar trends for these mathematical model fits.

4.3. Application of GP model

Instead of the mathematical modeling approach, GP model was applied to match the complicated pilot plant data. The models were trained using two different operation data, including the early-stage data (0–5 days) and the late-stage data (20–24 days). Thus, two kinds of GP models were obtained for predicting transmembrane pressure and membrane permeability. Although the data are not shown, the fitness functions for the GP model fits reached stable values after 30 generations. Fig. 6 illustrates the structures of equations determined by the GP model fits. As a result of this procedure, four equations were obtained.

The comparisons of the GP model fits with pilot plant data are illustrated in Fig. 7. The GP models match the pilot plant data well. Although there are deviations between the model calculations and experimental data, the overall trends in the transmembrane pressure and membrane permeability were successfully predicted by the GP models. Moreover, the GP model showed better fits to the membrane permeability than the transmembrane pressure.

Fig. 8 compares the GP model fits with the mathematical model fits. The deviations between the mathematical models and the pilot plant data were substantial. On the other hand, the GP models showed closed fits to the pilot plant data. This implies that the GP model is more useful than the simple mathematical models for the prediction and analysis of pilot or full plant operation data.

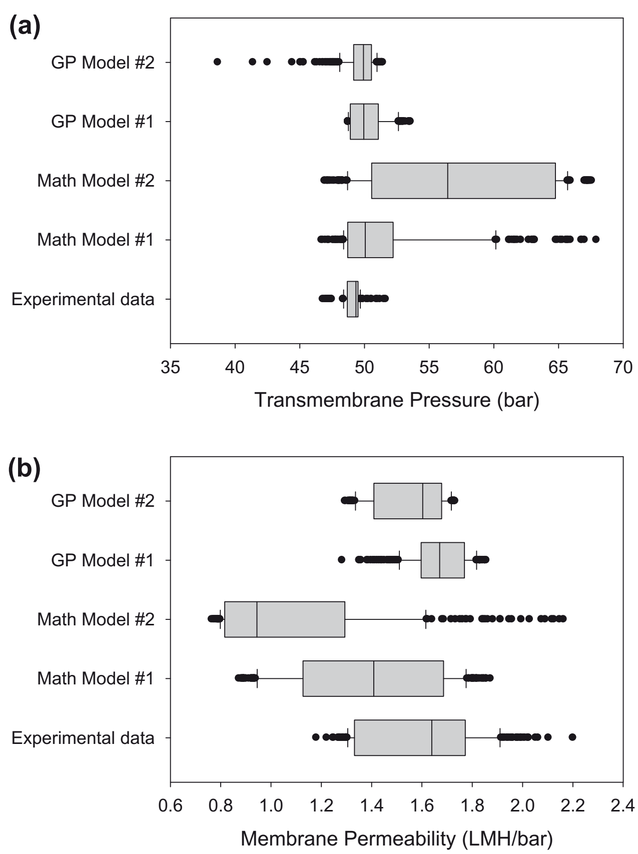

Fig. 8. Comparison of GP models with mathematical models. (a) Transmembrane pressure and (b) membrane permeability. Math models #1 and #2 were fitted to the early-and late-stage pilot plant data, respectively. GP models #1 and #2 were fitted to the early-and late-stage pilot plant data, respectively.

4.4. Sensitivity analysis for GP model

Fig. 9 shows the sensitivity of the GP model on input variables. The first GP model, which was fitted to the early-stage data, has large sensitivity to conductivity, flux, and ORP. On the contrary, the second GP model, which was fitted to the late-stage data, is sensitive to temperature, recovery, time, and ORP. Although it is not clear how the sensitivity results are correlated with the importance of input variables at this point, further works would allow to access information on the plant operation using the GP modeling approach.

5. Conclusions

In this work, a GP-based model was suggested to analyze the operation data of a pilot plant for seawater desalination. It was demonstrated that the GP model is useful to understand the behavior of pilot-scale systems. Accordingly, the GP model appears to be a better tool for RO systems under fluctuating conditions than mathematical or mechanistic models. Moreover, the GP model fit to the membrane permeability was better than that to the transmembrane pressure, suggesting that the combination of GP model and mathemati-

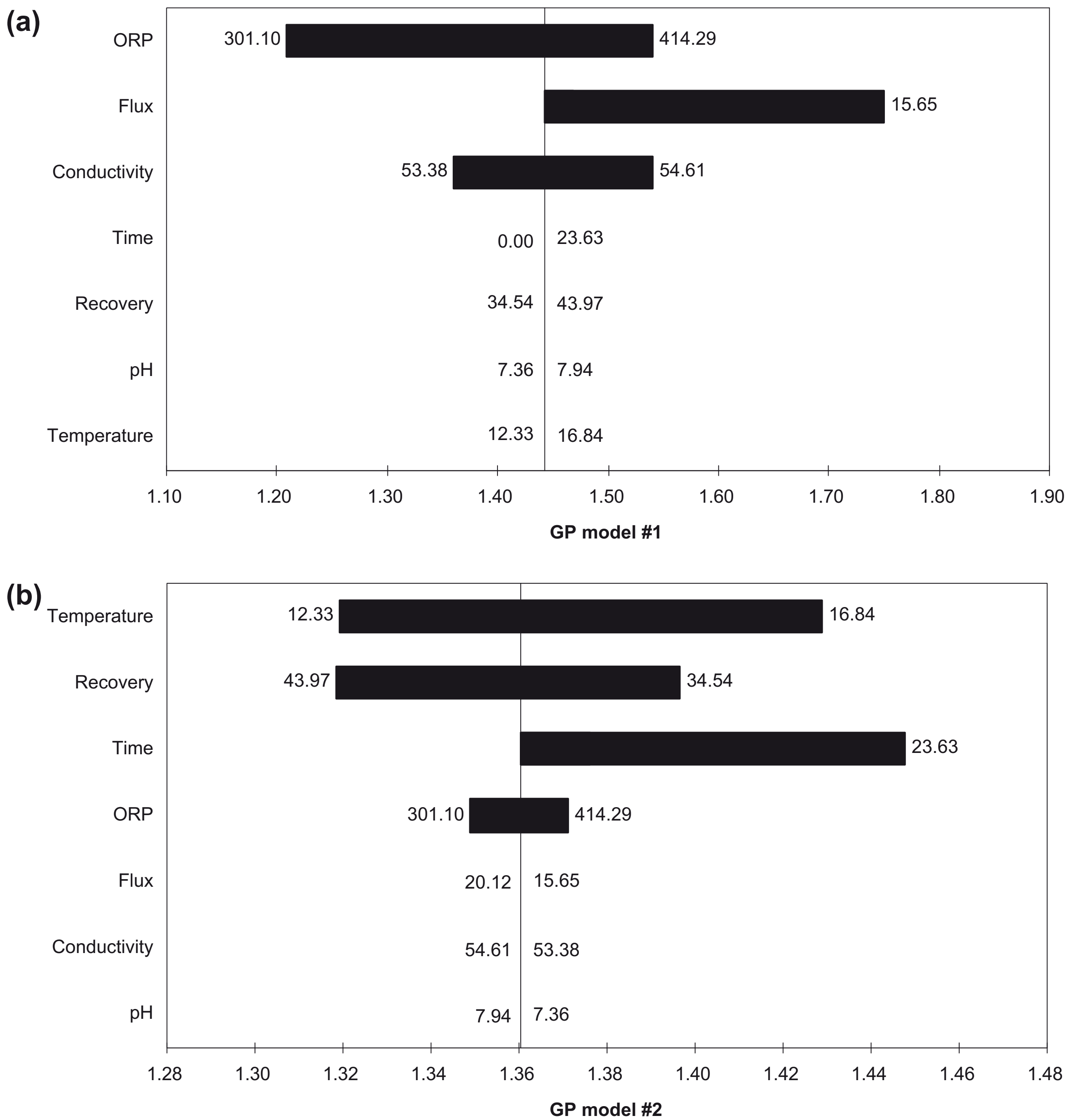

Fig. 9. Sensitivity analysis of the GP models. (a) GP model fitted to the early-stage data and (b) GP model fitted to the late-stage data.

cal model can improve the accuracy of the model prediction.

Acknowledgment

This research was supported by a grant (code# C106A152000106A085700220) from Plant Technology Advancement Program funded by Ministry of Construction and Transportation of Korean government.

References

[1] H. Hyung, J.H. Kim, A mechanistic study on boron rejection by sea water reverse osmosis membranes, J. Membr. Sci. 286 (1–2) (2006) 269–278.
[2] S. Atkinson, Japan's largest sea-water desalination plant uses Nitto Denko membranes, Membr. Technol. 2005(4) (2005) 10–11.
[3] H.-J. Oh, T.-M. Hwang, S. Lee, A simplified simulation model of RO systems for seawater desalination, Desalination 238(1–3) (2009) 128–139.

[4] M. Verhuelsdonk, T. Attenborough, O. Lex, T. Altmann, Design and optimization of seawater reverse osmosis desalination plants using special simulation software, Desalination 250 (2) (2010) 729–733.

[5] C.S. Slater, C.A. Brooks, III, Development of a simulation model predicting performance of reverse osmosis batch systems, Sep. Sci. Tech. 27 (11) (1992) 1361–1388.

[6] T.-M. Lee, H. Oh, Y.-K. Choung, S. Oh, M. Jeon, J.H. Kim, S.H. Nam, S. Lee, Prediction of membrane fouling in the pilot-scale microfiltration system using genetic programming, Desalination 247 (2009) 285–294.

[7] S. Silva, GPLAB: A genetic programming toolbox for MATLAB (2007) http://gplab.sourceforge.net.