Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2019.23382

143 (2019) 1–6
March

Prediction of fouling in a pilot-scale microfiltration plant using model tree for drinking water treatment

Yongjun Choi, Jaehyun Ju, Youngkyu Park, Sangho Lee*

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 (S. Lee)

Received 31 August 2018; Accepted 15 October 2018

ABSTRACT

A major obstacle to further incorporation of microfiltration (MF) membrane processes in water treatment plants is trans-membrane pressure (TMP) increase due to membrane fouling. Modelling and simulation of changes in TMP may be useful to describe fouling through the identification of the most relevant operating conditions. In this study, M5P model tree was applied to predict the fouling of the pilot scale MF system for drinking water. A 500 m^3/d MF pilot plant was operated for 1 year to analyze the performance of MF process under various operating conditions. The effects of operating parameters on membrane performance were evaluated based on the comparison of TMP as a function of operating time. The M5P model tree used five input variables including turbidity (NTU), temperature (°C), total organic carbon (mg/L), total operating period (d) and operating period after clean in place (d). The results of application of the M5P model tree indicated high correlation coefficients between the measured and predicted output variables. Therefore, it appears that the M5P model tree is applicable in the long-term prediction of the membrane performance in the pilot-scale MF systems.

Keywords: Microfiltration; Fouling; Model tree; M5P model; Drinking water

1. Introduction

In the past decades, microfiltration (MF) has been applied as an advanced water treatment process for drinking water production for the efficient removal of particulate pollutants, turbidity and microorganisms under low operating pressures [1,2]. Although there has been a great deal of advancement in the development of MF membrane processes, the main factor that limits the membrane performance is fouling by contaminants in feed water [3,4]. Membrane fouling which is an unavoidable phenomenon in the membrane process causes a decrease of the permeability in time as a consequence of the deposit of the solutes on or in the membrane pores [5–7]. Moreover, understanding and predicting membrane fouling is even challenging especially when the foulants in feed water vary with time [8,9]. Although prediction of fouling in MF membrane processes using mechanistic and statistical models has been widely developed, it is still difficult to gain a quantitative prediction results with various operational conditions due to the complexity of fouling phenomenon [10,11].

This study focused on investigation of the fouling characteristics of pilot-scale MF membrane process and prediction of the fouling using the M5P model tree from a long-term pilot-scale operation data. Factors affecting the extent of membrane fouling were examined in connection with the feed water quality and operational parameters such as temperature, turbidity, total organic carbon (TOC), total operating period and operating period after clean in place (CIP).

2. Materials and methods

2.1. Model tree

In this study, M5P model tree was used to predict the membrane filtration performance from the MF pilot plant

* Corresponding author.

data. The M5P model tree was created in WEKA software. It includes a wide variety of learning algorithms and preprocessing tools [12]. The M5P model tree is a predictive technique that has become increasingly noticed since Quinlan introduced it in 1992 [13]. M5P model tree is a combination of data classification and regression which works shown as Fig. 1 [14–16]. M5P model tree has several advantages over other machine learning algorithms [15,16]. It is simple to understand and interpret the models because the relation rules can be explicitly observed from the tree. This method does not take many preprocessing steps. M5P model trees can be used for multiple output problems. It is possible to validate a model using statistical test. The M5P model tree is a structured tree that depicts graphical if-then-else rules of the hidden or implicit knowledge inferred from the dataset in a top–down way [15–17]. At each step, a decision is made whether to partition the training set or to introduce a regression function as a leaf node.

2.2. MF pilot plant

A schematic diagram of the MF pilot plant used is provided in Fig. 2. A pilot-scale submerged filtration system that used hollow fiber modules with a nominal pore size of 0.1 μm (Lotte Chemical, Korea) was operated for 1 year to examine the effect of seasonal variation in feed water qualities and to analyze the performance of MF process under various operating conditions. The pilot plant was designed to produce 500 m^3/d of product water with over 90% recovery. The system was automatically operated and the data was collected using a computer. The results were analyzed in terms of the trans-membrane pressure (TMP). Operating conditions are as follows: 40-min filtration, and 2-min backwash with permeate and pressurized air. The raw water collected from Han River in Korea was used as the feed water after the coagulation pretreatment using polyaluminum chloride.

3. Results and discussion

3.1. Feed water quality

Fig. 3 shows the variations in turbidity, temperature and TOC of feed water used for membrane filtration. Online monitoring equipments (HACH, USA) were used to measure turbidity, water temperature and TOC. The data from the long-term membrane filtration system were collected for 1 year, during which chemical washing was conducted one time. Turbidity, temperature and TOC in feed water significantly changed with time; turbidity: 1.5–81.5 NTU, temperature: 2.1°C–26.9°C and TOC: 1.1–7.1 mg/L. In summer (from mid-July to mid-August), the feed water turbidity increased rapidly due to the frequent rains. TOC was high in spring (from March to April) and fall (from mid-September to mid-October) due to algae. The average temperature of the water in winter (from December to February) was only 3°C, and in summer, 24°C. Three distinct periods were observed over a year; high turbidity concentration period, high TOC concentration period and low temperature period.

3.2. Changes in TMP of pilot-scale submerged MF plant

Fig. 4 shows the variations in the TMP in whole operating period. It seems that the extent of fouling was significantly affected by the feed water quality. A rapid increase in TMP between mid-July and mid-August resulted from increased turbidity. From October to February, the TMP dramatically increased because of the decreased water temperature.

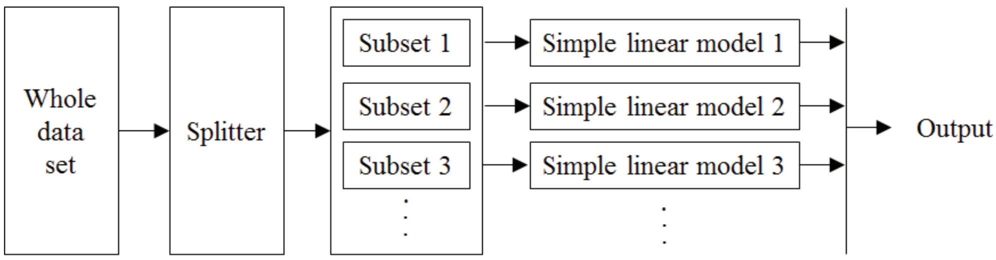

Fig. 1. Induction of a M5P model tree as a modular model [15,16].

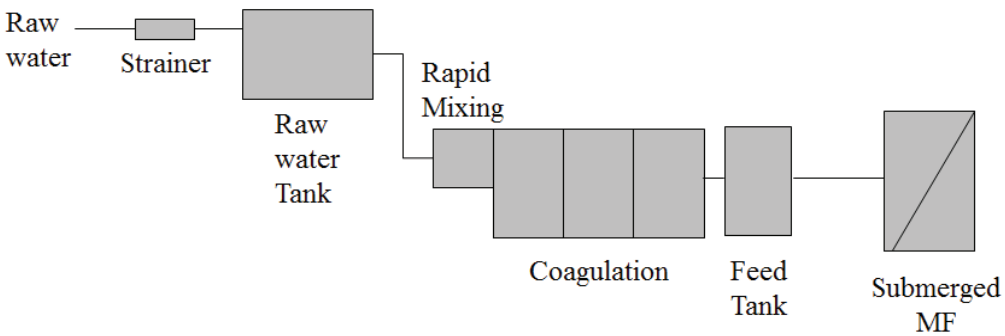

Fig. 2. Schematic diagram of the submerged MF pilot plant.

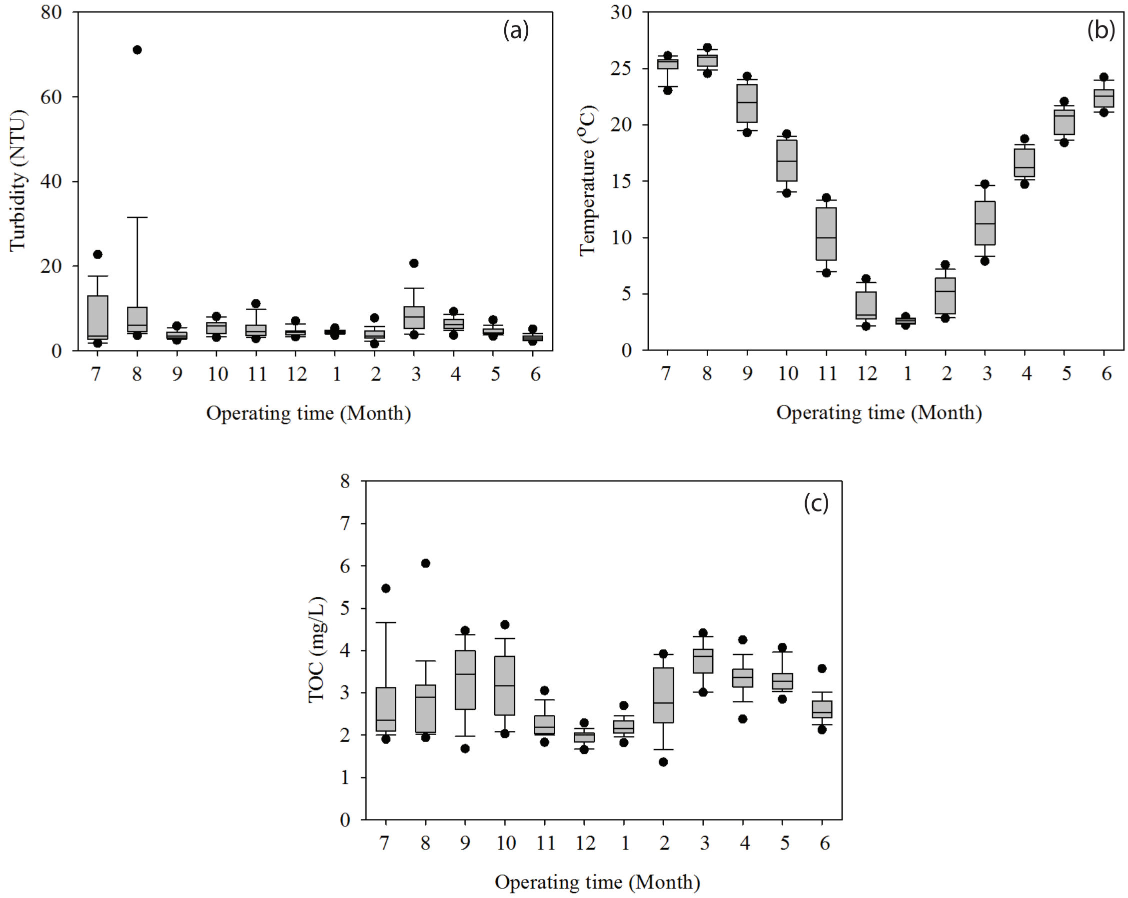

Fig. 3. Seasonal variations in the feed water quality: (a) turbidity (NTU), (b) temperature (°C) and (c) TOC (mg/L).

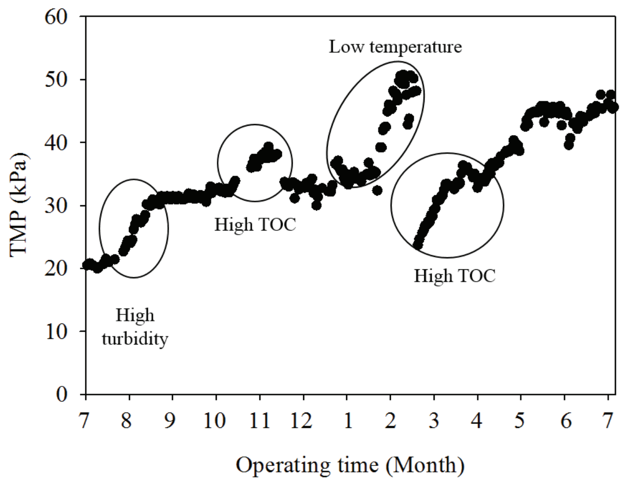

Fig. 4. Changes in TMP for pilot-scale submerged MF plant.

From March to April, the TMP increase again due to high TOC while feed water temperature gradually increased. Accordingly, the TMP of pilot-scale submerged MF plant was found to be sensitive to the seasonal variation in feed water quality. Therefore, it can be concluded that the TMP of pilot-scale submerged MF plant properly reflects the seasonal differences in the feed water quality.

3.3. Model fit to the experimental data using the model tree

As described, M5P model tree was developed to simulate the performance of the pilot-scale submerged membrane system. The M5P model tree, unlike many commonly used data-driven methods, can partly explain the system and reveal hidden patterns from the data [18].

The M5P model tree developed from collected data over 1-year period which can cover all probable seasonal variations in the studied variables was primarily used for finding functional dependencies between the operating parameters. All on-line daily data, temperature, TOC, turbidity, total operating period (TOP) and operating period after CIP (OPAC) were used to induce the model tree.

Tree model obtained from M5P tree model is given in Fig. 5. The model tree was composed of 25 multivariable linear equations (or linear models, LM) shown as Table 1, each of which is valid under specific operating conditions in the system given by specific values of the operation conditions (on-line daily data). The 10-fold cross-validations of the developed model tree led to a correlation of 0.97 and root mean square error of 22.87 conditions until reaching each LM in the leaves. The operating parameters at the top of the tree were the most discriminating [17]. In this study, total operating period was the most influential parameter affecting the TMP in the M5P model tree. And the second most discriminating variable was the operating period after CIP.

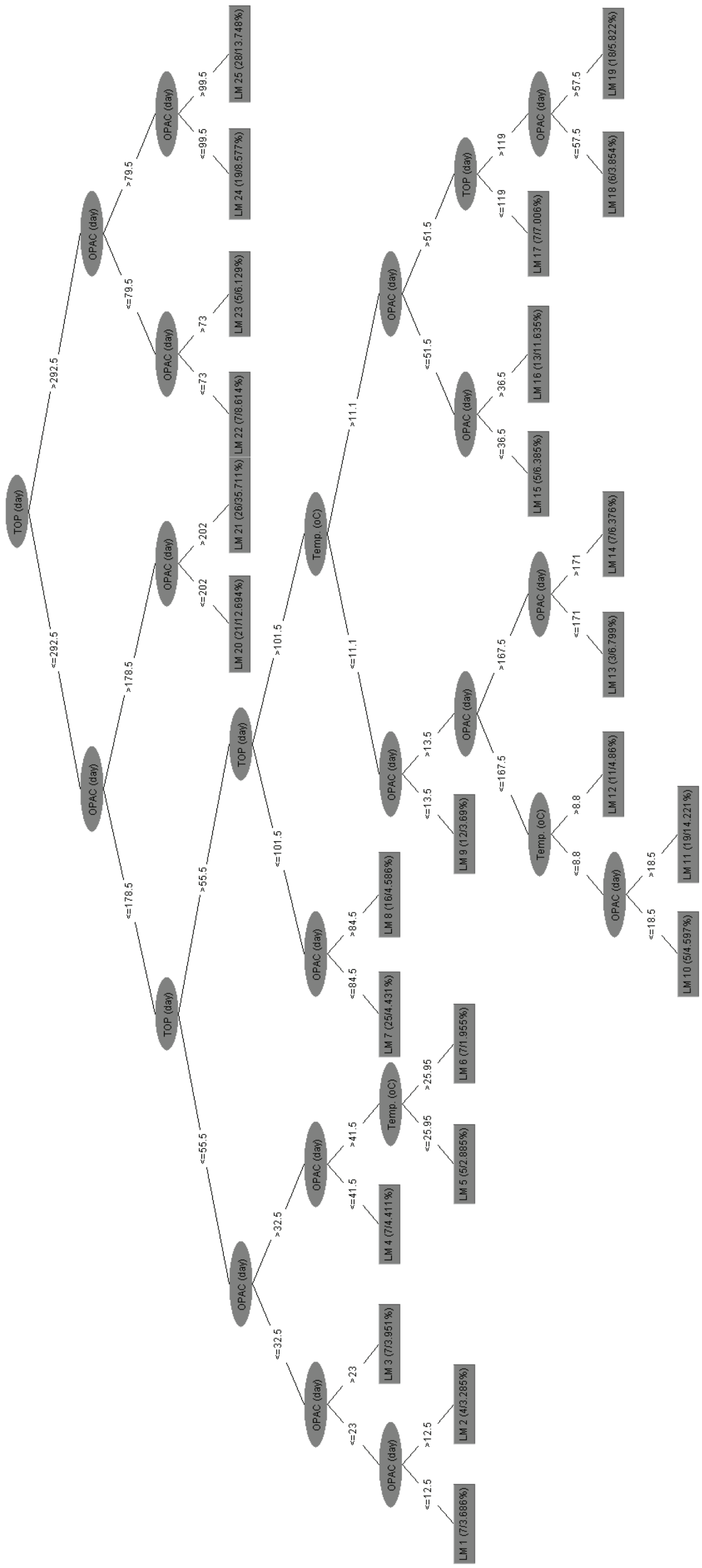

Fig. 5. M5P tree for TMP prediction of pilot-scale submerged MF plant.

Table 1
Linear models from M5P tree model for TMP prediction

No	Linear models (LM)
1	TMP(kPa)=0.1386×OPAC(d)+0.0211×TOP(d)+0.0286×temperature(°C)-0.003×turbidity(NTU)+0.1917×TOC(mg/L)+18.8136
2	TMP(kPa)=0.1394×OPAC(d)+0.0211×TOP(d)+0.0286×temperature(°C)-0.003×turbidity(NTU)+0.1917×TOC(mg/L)+18.8335
3	TMP(kPa)=0.1689×OPAC(d)+0.0211×TOP(d)+0.0286×temperature(°C)-0.003×turbidity(NTU)+0.1917×TOC(mg/L)+18.2848
4	TMP(kPa)=0.1549×OPAC(d)+0.0211×TOP(d)+0.0328×temperature(°C)-0.003×turbidity(NTU)+0.1412×TOC(mg/L)+19.7686
5	TMP(kPa)=0.1455×OPAC(d)+0.0211×TOP(d)-0.0495×temperature(°C)-0.003×turbidity(NTU)+0.1917×TOC(mg/L)+22.5281
6	TMP(kPa)=0.1455×OPAC(d)+0.0211×TOP(d)-0.042×temperature(°C)-0.003×turbidity(NTU)+0.1917×TOC(mg/L)+22.2946
7	TMP(kPa)=0.0413×OPAC(d)+0.0195×TOP(d)+0.1695×temperature(°C)+0.093×turbidty(NTU)+0.2923×TOC(mg/L)+21.9295
8	TMP(kPa)=0.0437×OPAC(d)+0.0195×TOP(d)+0.1695×temperature(°C)+0.093×turbidity(NTU)+0.2923×TOC(mg/L)+22.131
9	TMP(kPa)=0.1606×OPAC(d)+0.0178×TOP(d)+0.1991×temperature(°C)+0.1892×turbidity(NTU)+1.0401×TOC(mg/L)+17.0525
10	TMP(kPa)=0.0424×OPAC(d)+0.0178×TOP(d)+0.2128×temperature(°C)+0.1669×turbidity(NTU)+0.6397×TOC(mg/L)+20.9027
11	TMP(kPa)=0.0424×OPAC(d)+0.0178×TOP(d)+0.2128×temperature(°C)+0.1669×turbidity(NTU)+0.661×TOC(mg/L)+20.9327
12	TMP(kPa)=0.0424×OPAC(d)+0.0178×TOP(d)+0.229×temperature(°C)+0.1736×turbidity(NTU)+0.6604×TOC(mg/L)+20.8734
13	TMP(kPa)=0.0882×OPAC(d)+0.0178×TOP(d)+0.1616×temperature(°C)+0.1736×turbidity(NTU)+0.8356×TOC(mg/L)+13.3852
14	TMP(kPa)=0.0621×OPAC(d)+0.0178×TOP(d)+0.1616×temperature(°C)+0.1736×turbidity(NTU)+0.8356×TOC (mg/L)+18.0456
15	TMP(kPa)=0.0242×OPAC(d)+0.0181×TOP(d)+0.2078×temperature(°C)+0.0926×turbidity(NTU)+0.2729×TOC(mg/L)+25.17
16	TMP(kPa)=0.0272×OPAC(d)+0.0181×TOP(d)+0.2078×temperature (°C)+0.0926×turbidity(NTU)+0.2285×TOC(mg/L)+25.0932
17	TMP(kPa)=0.0597×OPAC(d)+0.0202×TOP(d)+0.2078×temperature(°C)+0.0926×turbidity(NTU)-0.1103×TOC(mg/L)+23.3292
18	TMP(kPa)=0.0348×OPAC(d)+0.0187×TOP(d)+0.2078×temperature(°C)+0.0926×turbidity(NTU)+0.0227×TOC(mg/L)+26.1933
19	TMP(kPa)=0.0348×OPAC(d)+0.0189×TOP(d)+0.2078×temperature(°C)+0.0926×turbidity(NTU)+0.0899×TOC(mg/L)+26.0845
20	TMP(kPa)=0.1441×OPAC(d)+0.0172×TOP(d)+0.1467×temperature(°C)-0.003×turbidity(NTU)+0.176×TOC(mg/L)+4.7236
21	TMP(kPa)=0.2661×OPAC(d)+0.0172×TOP(d)+0.1467×temperature(°C)-0.003×turbidity(NTU)+0.176×TOC(mg/L)-16.6498
22	TMP(kPa)=0.1547×OPAC(d)+0.009×TOP(d)+0.0658×temperature(°C)-0.01×turbidity(NTU)+0.9987×TOC (mg/L)+22.9951
23	TMP(kPa)=0.2119×OPAC(d)+0.009×TOP(d)+0.0658×temperature(°C)-0.01×turbidity(NTU)+0.9987×TOC (mg/L)+19.0658
24	TMP(kPa)=0.0419×OPAC(d)+0.009×TOP(d)+0.0658× temperature(°C)+0.192×turbidity(NTU)+0.5789×TOC(mg/L)+33.5527
25	TMP(kPa)=0.1013×OPAC (d)+0.009×TOP(d)+0.0658×temperature(°C)+0.6571×turbidity(NTU)

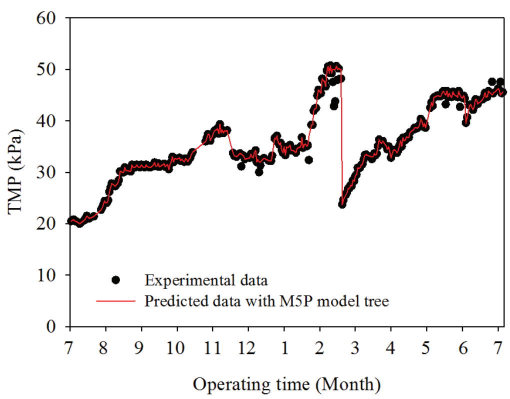

Fig. 6. Comparison of experimental data with prediction results obtained from M5P model tree.

Each LM enables the estimation of the TMP as a linear regression of multiple operating parameters. LM 1, for example, is valid when the TOP is below 292.5 d, the OPAC is under 178.5 d, the TOP is under 55.5, the OPAC is under 32.5 d, the OPAC is under 23 d and, finally, the OPAC is below 12.5 d. All of the LMs are listed in Table 1. From the LM equations, it is possible to know which parameters are the most influential because each parameter is multiplied by a weighting factor [18].

The comparison between measured and predicted TMP values is shown in Fig. 6. It is observed that the model prediction results track the observed data very well (*R* value = 0.97) as shown in Fig. 6. The M5P model tree showed high strength and a linear relationship direction between the predicted data and experimental data. It is observed that the output tracks the targets very well. This suggests that the M5P model tree has the potential for long-term (order of month) prediction of the membrane performance in pilot-scale systems in the presence of seasonal variations of feed water quality.

4. Conclusions

In this study, M5P tree model was used to describe fouling phenomena as measured by TMP over 1 year at a submerged MF pilot plant operated under different conditions. Results show that the seasonal differences in the feed water quality such as turbidity, temperature and TOC significantly affected the MF membrane fouling in pilot-scale plant.

The results of application of M5P tree model indicated high correlation coefficient (*R* value) between the measured and predicted output variables reaching up to 0.97. This means that the M5P tree model has great potential to the long-term (order of months) prediction of the membrane performance at different operating conditions of the pilot-scale system.

Acknowledgements

This subject is supported by Korea Ministry of Environment as "Global Top Project (2017002100001)" and by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 18IFIP-B116951-03).

References

[1] J.M. Laine, D. Vial, P. Moulart, Status after 10 years of operation-overview of UF technology today, Desalination, 131 (2000) 17–25.
[2] B. Van der Bruggen, C. Vandecasteele, T.V. Gestel, W. Doyen, R. Leysen, A review of pressure-driven membrane processes in wastewater treatment and drinking water production, Environ. Prog., 22 (2003) 46–56.
[3] S. Chellam, J.G. Jacangelo, T.P. Bonacquisti, Modeling and experimental verification of pilot-scale hollow fiber, direct flow microfiltration with periodic backwashing, Environ. Sci. Technol., 32 (1998) 75–81.
[4] K.J. Howe, M.M. Clark, Fouling of Microfiltration and Ultrafiltration Membranes by Natural Waters, Environ. Sci. Technol., 36 (2002) 3571–3576.
[5] R.H. Peiris, C. Halle, H. Budman, C. Moresoli, S. Peldszus, P.M. Huck, R.L. Legge, Identifying fouling events in a membrane-based drinking water treatment process using principal component analysis of fluorescence excitation–emission matrices, Water Res., 44 (2010) 185–194.
[6] L. Wang, X. Wang, K. Fukushi, Effects of operational conditions on ultrafiltration membrane fouling, Desalination, 229 (2008) 181–191.
[7] N. Lee, G. Amy, J. Croue, Low-pressure membrane (MF/UF) fouling associated with allochthonous versus autochthonous natural organic matter, Water Res., 40 (2006) 2357–2368.
[8] W. Gao, H. Liang, J. Ma, M. Han, Z.-l. Chen, Z.-s. Han, G.-b. Li, Membrane fouling control in ultrafiltration technology for drinking water production: a review, Desalination, 272 (2011) 1–8.
[9] S. Babel, S. Takizawa, H. Ozaki, Factors affecting seasonal variation of membrane filtration resistance caused by Chlorella algae, Water Res., 36 (2002) 1193–1202.
[10] Q.F. Liu, S.H. Kim, S. Lee, Prediction of microfiltration membrane fouling using artificial neural network models, Sep. Purif. Technol., 70 (2009) 96–102.
[11] F. Schmitt, R. Banuc, I.T. Yeom, K.U. Do, Development of artificial neural networks to predict membrane fouling in an anoxic-aerobic membrane bioreactor treating domestic wastewater, Biochem. Eng. J., 133 (2018) 47–58.
[12] H.I. Witten, E. Frank, M.A. Hall, Data Mining, Principle Machine Learning Tools and Techniques, Burlington, MA, 2011.
[13] J.R. Quinlan, Learning with Continuous Classes, in: Proc. 5th Australian Joint Conference on Artificial Intelligence, World Scientific, Singapore, 1992, pp. 343–348.
[14] Y. Wang, I.H. Witten, Inducing Model Trees for Continuous Classes, in Proc. 9th European Conference on Machine Learning, 1997.
[15] B. Bhattacharya, D.P. Solomatine, Neural networks and M5P model trees in modelling water level-discharge relationship, Neurocomputing, 63 (2005) 381–396.
[16] D.P. Solomatine, Optimisation of hierarchical modular models and M5 trees, in Proc. of International Joint Conference on Neural Networks, Budapest, Hungary, 2004.
[17] E.K. Onyari, F.M. Ilunga, Application of MLP neural network and M5P model tree in predicting streamflow: a case study of Luvuvhu Catchment, South Africa, Int. J. Innov. Manage. Technol., 4 (2013) 1–15.
[18] M. Dalmau, N. Atanasova, S. Gabarrón, I.R. Roda, J. Comas, Comparison of a deterministic and a data driven model to describe MBR fouling, Chem. Eng. J., 260 (2015) 300–308.