# KNT-artificial neural network model for flux prediction of ultrafiltration membrane producing drinking water

**H.K. Oh*, M.J. Yu*, E.M. Gwon**, J.Y. Koo*, S.G. Kim* and A. Koizumi*****

* Department of Environmental Engineering, University of Seoul, Seoul, 130-743, Korea
(E-mail: *phoh9676@hanmail.net*; *myong@uos.ac.kr*; *jykoo@uos.ac.kr*; *sgkim-75@hanmail.net*)

** Department of Environmental Engineering and Biotechnology, Myongji Univ., Yongin, 449-728, Korea
(E-mail: *emgwon@mju.ac.kr*)

*** Department of Environmental Civil Engineering, Tokyo Metropolitan Univ., Hachioji, 192-0397, Japan
(E-mail: *akoiz@ecomp.metro-u.ac.jp*)

**Abstract** This paper describes the prediction of flux behavior in an ultrafiltration (UF) membrane system using a Kalman neuro training (KNT) network model. The experimental data was obtained from operating a pilot plant of hollow fiber UF membrane with groundwater for 7 months. The network was trained using operating conditions such as inlet pressure, filtration duration, and feed water quality parameters including turbidity, temperature and $UV_{254}$. Pre-processing of raw data allowed the normalized input data to be used in sigmoid activation functions. A neural network architecture was structured by modifying the number of hidden layers, neurons and learning iterations.

The structure of KNT-neural network with 3 layers and 5 neurons allowed a good prediction of permeate flux by 0.997 of correlation coefficient during the learning phase. Also the validity of the designed model was evaluated with other experimental data not used during the training phase and nonlinear flux behavior was accurately estimated with 0.999 of correlation coefficient and a lower error of prediction in the testing phase.

This good flux prediction can provide preliminary criteria in membrane design and set up the proper cleaning cycle in membrane operation. The KNT-artificial neural network is also expected to predict the variation of transmembrane pressure during filtration cycles and can be applied to automation and control of full scale treatment plants.

**Keywords** KNT-artifical neural network; modeling; permeate flux; transmembrane pressure; ultrafiltration

## Introduction

Ultrafiltration (UF) membranes have been accepted in the field of drinking water treatment over the past ten years, and have shown an excellent potential to remove pathogens through a physical barrier (EPA, 2001). UF membranes can remove particles, colloids and micro-organisms such as *Giardia, Cryptosporidium* and viruses easily from raw water and thus have been primarily applied to retrofit existing sand filtration in a conventional treatment process. Also UF with lower molecular weight cut-off (less than 10,000 daltons) or UF combined with pretreatment processes such as coagulation and powdered activated carbon adsorption can effectively remove dissolved materials in raw water, including dissolved organic matters, DBP precursors, synthetic organic compounds, taste and odor compounds (Lahoussine-Turcaud *et al*., 1990; Campos *et al*., 2000). Thus, the growing experience of UF technology in the water treatment industry has accumulated know-how about how to effectively design, operate and maintain membrane systems.

However, the cost of membrane application depends on how one manages permeate flux. Flux declines by membrane fouling mainly due to particle deposition on the membrane fibers and adsorption of organic matters on the surface or into the membrane pores. The irreversible and reversible resistances have total hydraulic resistances and subsequently result in a decrease of permeate flux during membrane filtration time. Therefore, it is first required to understand and control membrane foulings in UF

Water Science and Technology **Vol 50 No 8 pp 103–110** © IWA Publishing 2004

membrane operation and then develop a prediction model for membrane performance via long-term operation.

Much research has been done to examine membrane fouling mechanisms and to predict permeate flux and transmembrane pressure (TMP) during membrane operation. Jones *et al*. (1993) developed a hydraulic model based on the Hagen–Poiseuille equation. This model accurately predicted the flux and pressure for a continuous flow within a hollow fiber UF system under fouling and non-fouling conditions. Hong *et al*. (1997) systematically investigated the dynamic behavior of permeate flux in cross-flow membrane filtration of colloidal suspensions. They established the mechanistic model based on simplified particle mass balance for transient stages of filtration and Happel's cell model for the hydraulic resistance. Wang and Song (1999) demonstrated the theory of fouling dynamics by Carman–Kozeny could well predict the time-dependent flux and the arrival time to steady state under cross-flow membrane filtration experiments with a colloidal suspension of 0.12 μm. The steady-state flux of UF could be almost evaluated as the limiting flux of the process. However, since these deterministic models were developed for synthetic waters and restricted by boundary conditions and given assumptions, they were limited in applying to full-scale treatment plants with various water sources.

Therefore, recently statistical approaches began to emerge in permeate flux or TMP estimation (Oh *et al*., 2001). Among these modeling methods, the neural network has been used for a non-linear modeling tool to evaluate permeate flux and TMP in membrane filtration. Delgrange *et al*. (1998) used a neural network to model the time evolution of TMP for UF membranes. Input data were permeate flux, turbidity and pressure at the cycle start and at the end of the previous cycle. Thus, the elaborated networks were effectively able to model the effect of reversible and irreversible fouling on the pressure behavior. This methodology for the prediction of flux and pressure could be available for membrane productivity improvement, optimization, automation and design in real UF membrane processes. Teodosiu *et al*. (2000) constructed two neural network models to predict the flux at any time instant during filtration and after backwashing. The trained networks using a well-established experimental data set gave an accurate prediction of flux evolution in a limited flux and short filtration time.

In this study, a Kalman neuro training (KNT)-neural network was tried to model the permeate flux behavior for UF membranes. A pilot plant of hollow fiber UF membrane had been operated with groundwater for about 7 months. The KNT-network model was designed to predict the time evolution of flux from available data on raw water quality and operating conditions. Finally the model can be available in optimizing flux behavior in the filtration cycle, designing membrane systems and simulating real situations before running a pilot plant.

## Approach of KNT-artificial neural network

Neural network models are algorithms for cognitive work, such as learning, optimizing and predicting, which are based on concepts derived from the nature of the human brain (Müller *et al*., 1995). The multilayer neural network structure is illustrated in Figure 1. This network architecture is composed of nodes (neurons), links (synapses) and transfer functions. Neurons ($X_1 \cdots X_i$, $U_1 \cdots U_j$, $W_1 \cdots W_k$) process computation from the input data to produce the output. A weight ($W_{ij}$ and $W_{jk}$) is associated with each link between two nodes and represents the synaptic coupling strengths. The transfer function is defined for each processing element and determines the state of the node as a function of its bias. For this network, the non-linear step function was a sigmoid function.

The Kalman filter generally uses partial derivatives as linear approximations for nonlinear relations and is called *extended Kalman filtering* (Grewal and Andrews, 2001). The

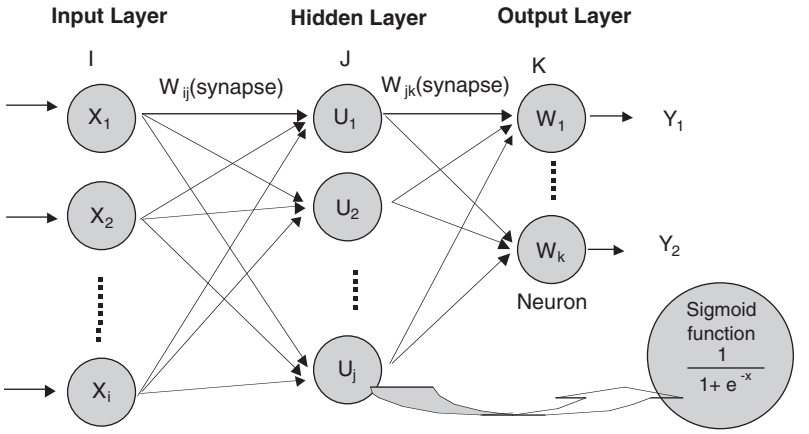


**Figure 1** Structure of a multilayer feed-forward network (→ : link, ○ : node)

algorithm of a *Kalman filter* through linear probability system is described as follows (Murase *et al*., 1994).

$$\text{The state equation } \{x\}_{t+1} = [F]_t \{x\}_t + [G]_t \{w\}_t \tag{1}$$

$$\text{The observation equation } \{y\}_{t+1} = [H]_t \{x\}_t + \{v\}_t \tag{2}$$

where $\{x\}_t$ is the state vector, $\{y\}_t$ is the observation vector, $\{w\}_t$ is Gaussian white noise, $\{v\}_t$ is measurement noise, $[F]_t$ is the system matrix, $[G]_t$ is the process matrix and $[H]_t$ is the observation matrix.

Both $\{w\}_t$ and $\{v\}_t$ are zero-mean white Gaussian noises and can be described in terms of covariance matrices:

$$E\left[\begin{pmatrix}\{x\}_t \\ \{v\}_t\end{pmatrix}\left(\{x\}_s^T \{v\}_s^T\right)\right] = \begin{pmatrix}[Q]_t & 0 \\ 0 & [R]_t\end{pmatrix}\delta_{ts}[R]_t > 0 \tag{3}$$

Based on this system, the essential equations defining the continuous form of the extended Kalman filter (EKF) are given for the iteration algorithm.

$$\text{EKF equation } \{\hat{X}\}_{t/t} = \{\hat{X}\}_{t/t-1} + [K]_t[\{y\}_t - [H]_t \{\hat{X}\}_{t/t-1}] \tag{4}$$

$$\text{Kalman gain } [K]_t = \{P\}_{t/t-1}[H]_t^T[[H]_t \{P\}_{t/t-1}[H]_t^T + [R]_t]^{-1} \tag{5}$$

$$\text{Covariance matrix for estimate error } [P]_{t/t} = [P]_{t/t-1} - [K]_t[H]_t[P]_{t/t-1}\ t = 0,1, \ldots \tag{6}$$

$$\text{Initial condition } \{\hat{X}\}_{0/-1} = \hat{X}_0,\ [P]_{0/-1} = \Sigma_0 \tag{7}$$

where both $\{X\}_{t/t}$ and $\{P\}_{t/t}$ at any instant on the trajectory are calculated by the previous $\{X\}_{t/t-1}$ and $\{P\}_{t/t-1}$. Given initial value, $x_0$ and $\Sigma_0$, the values for $[K]$ and $[P]$ can be computed from the known $[H]_t$ and the EKF equation calculates to obtain a new estimate $\{X\}$.

The observation equation is the nonlinear expression by Eq. (8) and Taylor series expansion of $f\{h(\{x\}_t)_t\}$ with respect to $\{\hat{X}\}_{t/t-1}$ is linearized as follows.

$$\text{The observation equation } \{y\}_t = \{h(\{x\}_t)\}_t + \{v\}_t \tag{8}$$

$$\{h(\{x\}_t)\}_t = \{h(\{X\}_t)\}_t + [H]_t(\{x\}_t - \{\hat{X}\}_{t/t-1}) + \ldots \tag{9}$$

where $[H]_t = \partial(h)_t/\partial\{x\}_{t/x-\overline{X_{t/t-1}}}$ and if the higher order terms can be neglected, then this method transforms the nonlinear measurement to a linear system.

## Materials and methods

### Experimental equipment and operation mode of UF

The schematic of the experimental equipment is shown in Figure 2. The hollow fiber membrane module with a molecular weight cut-off of 100,000 daltons, has an effective membrane area of 0.58 $m^2$. The material of the membrane is polysulfone and permeate flux is specified by 6–8 L/min with pure water at 0.7 bar. The UF membrane system was operated with groundwater in a cross-flow mode and set by applied pressure as described in Table 1. The initial condition was recovered by chemical cleaning at the end of each operation mode. Turbidity, UV absorbance at 254 nm and temperature were measured daily and also permeate flux and TMP were continuously monitored.

### Data processing and generalization

The network inputs were chosen among water quality and operation parameters. Applied pressure, filtration duration and temperature have stronger relationships with permeate flux than turbidity and $UV_{254}$. Because the groundwater had low turbidity and $UV_{254}$, the effect of contaminants on permeate flux did not appear significant. The whole data were 8 sets and normalized between 0 and 1, as described in Table 2.

### Evaluation of KNT-neural network model

For diagnostic checking for fitted model, error analysis was performed.

The correlation coefficient ($R$) was evaluated for model fit and calculated by actual values ($y_t$) and fitted values ($\hat{y}_t$) using Eq. (10).

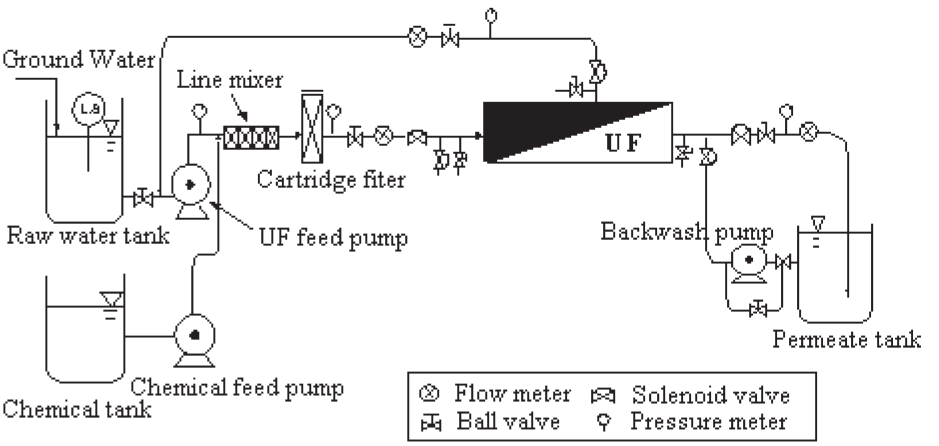


**Figure 2** Schematic diagram of UF membrane pilot plant

**Table 1** Initial operating conditions for each operation step

| Parameters / Operating step | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Applied pressure (bar) | 1.00 | 0.51 | 0.58 | 0.8 | 0.87 | 0.64 | 0.91 | 0.58 |
| Transmembrane pressure (bar) | 0.69 | 0.37 | 0.43 | 0.57 | 0.62 | 0.48 | 0.48 | 0.42 |
| Inlet flow (L/min) | 6.00 | 5.00 | 5.40 | 6.20 | 6.50 | 5.70 | 6.60 | 5.40 |
| Linear velocity (m/s) | 0.40 | 0.35 | 0.36 | 0.42 | 0.43 | 0.38 | 0.44 | 0.38 |

**Table 2** Normalization methodology of input factors

| Input factors | Generalization method | Range of input data | |
|---|---|---|---|
| | | Untreated | Normalized |
| Permeate flux ($y_t$) | $y_t = Y_t/(\max. Y_t + \min. Y_t)$ | 0.41–1.10 | 0.27–0.73 |
| Temperature ($x_{1t}$) | $X_{1t} = X_{1t}/(\max. X_{1t} + \min. X_{1t})$ | 17.0–31.9 | 0.35–0.65 |
| Turbidity ($x_{2t}$) | $X_{2t} = X_{2t}/(\max. X_{2t} + \min. X_{2t})$ | 0.13–1.10 | 0.10–0.89 |
| $UV_{254}$ ($x_{3t}$) | $X_{3t} = X_{3t}/(\max. X_{3t} + \min. X_{3t})$ | 0.003–0.068 | 0.090–0.980 |
| Applied pressure ($x_{4t}$) | $X_{4t} = X_{4t}/(\max. X_{4t} + \min. X_{4t})$ | 0.51–1.0 | 0.33–0.67 |
| Filtration duration ($x_{5t}$) | $X_{5t} = X_{5t}/30$ | 10.0–23.0 | 0.33–0.78 |

$$R = \frac{\sum_{t-1}^{n}\{(y_t - m_y)(\hat{y}_t - \hat{m}_y)\}}{\sqrt{\sum_{t-1}^{n}(y_t - m_y)^2 \sum_{t-1}^{n}(\hat{y}_t - \hat{m}_y)^2}} \tag{10}$$

where $m_y$ and $\hat{m}_y$ are the mean measured value and mean fitted value, respectively. If $R$ approaches 1, the model fit is appropriate for forecasting the fluctuation of real data.

Also mean absolute percentage error (MAPE) was used for model diagnosis using Eq. (11). This value means residual between predicted values and measured values. As MAPE is near 0, the fitness of model is good.

$$\text{MAPE} = \frac{1}{n}\sum_{t-1}^{n}\left|\frac{(y_t - \hat{y}_t)}{y_t}\right| \times 100 \tag{11}$$

where $n$ is data number, $y_t$ is measured and $\hat{y}_t$ is estimated.

Finally mean-squared error (MSE) was applied for model validity using Eq. (12). This value can coincidently evaluate expectancy and variance. As this value is lower, the model is more significant.

$$\text{MSE} = \frac{1}{n}\sum_{t-1}^{n}(y_t - \hat{y}_t)^2 \tag{12}$$

## Results and discussion

Figure 3 shows variation of raw water quality during UF membrane operation. Turbidity and pH remained almost unchanged but temperature and UV absorbance at 254 nm showed high variation.

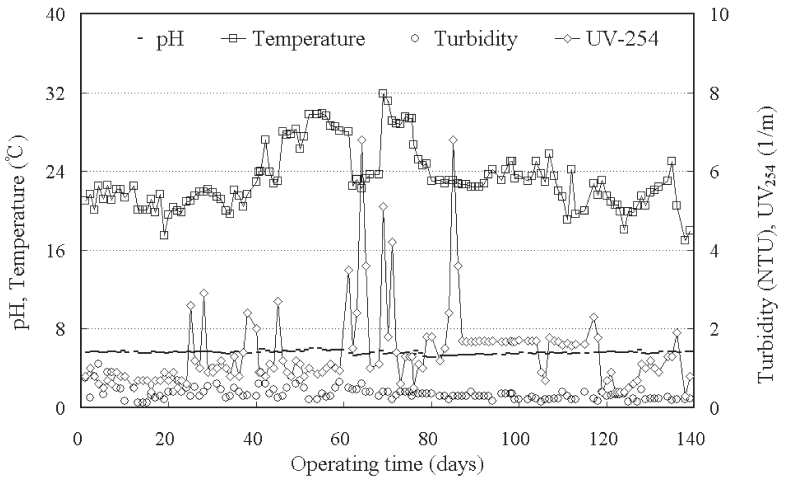


**Figure 3** Water quality of raw water during UF membrane operation

Because hidden layers of the neural network structure are arbitrary, optimum architecture should be designed based on error analysis. The number of training iterations was 500 times and the error goal was $10^{-5}$ of relative error. Table 3 summarizes the correlation coefficient (R) for training periods. As the processing element increased, the correlation coefficient between measured and predicted values increased.

Also error analysis between measurements and predictions was performed to evaluate the accuracy of different network structures (Figure 4). As layers and neurons increased, the errors of prediction lowered. In this study, optimum network structure was investigated based on both more than 0.95 of correlation coefficient and less than 0.02 of maximum mean squared error. Network structures satisfying the criteria were 1 layer with more than 5 neurons, 2 layers with more than 4 neurons, 3 layers with more than 4 neurons, 4 layers with more than 5 neurons, and 5 layers with more than 4 neurons. Among those different network structures, the neural network with 3 layers of 5 neurons (3*L*-5*N*) was selected as an appropriate model structure due to less complicatedness and better stability still within the goal.

Concerning the learning phase, the 3*L*-5*N* network structure was well established for flux behavior in each filtration mode, and resulted in good prediction of the permeate flux as shown in Figure 5. We could evaluate that the network was so well trained by the input parameters in predicting the flux decline in this period. The validity of the 3*L*-5*N* network was evaluated using the data after 98 days of operation in the testing phase, and results of testing were also plotted in Figure 5. A minor difference between predicted and experimental fluxes was observable in each initial and end filtration cycle for both training and testing periods. However, the whole evolution of flux decline in the filtration cycle was correctly

**Table 3** Evaluation of neural network structure by correlation coefficient in learning period

| Number of hidden layers | Number of hidden neurons | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | 0.7633 | 0.8985 | 0.9285 | 0.9492 | **0.9615** | **0.9714** | **0.9785** | **0.9865** |
| 2 | 0.6784 | 0.9081 | 0.9497 | **0.9777** | **0.9543** | **0.9885** | **0.9964** | – |
| 3 | 0.6851 | 0.9085 | 0.9499 | **0.9760** | **0.9970** | **0.9989** | **0.9989** | – |
| 4 | 0.6499 | 0.9178 | 0.9489 | **0.9881** | **0.9879** | **0.9915** | – | – |

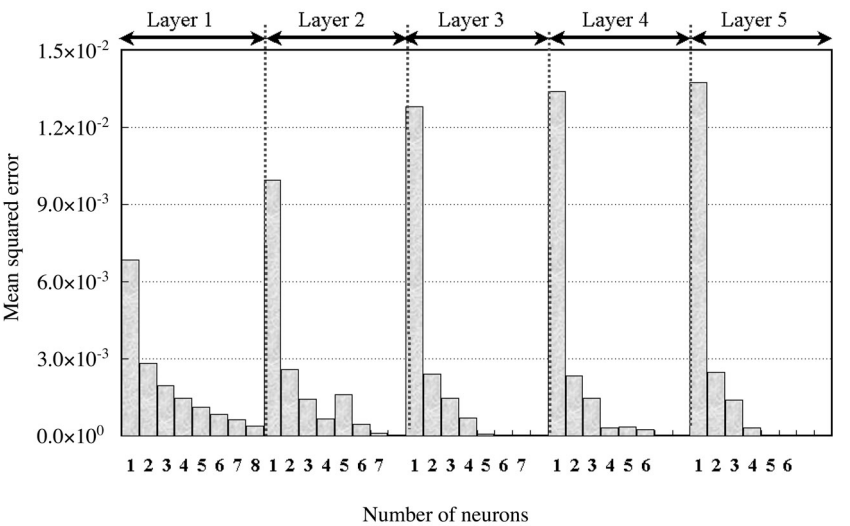


**Figure 4** Results of mean squared error in the learning phase

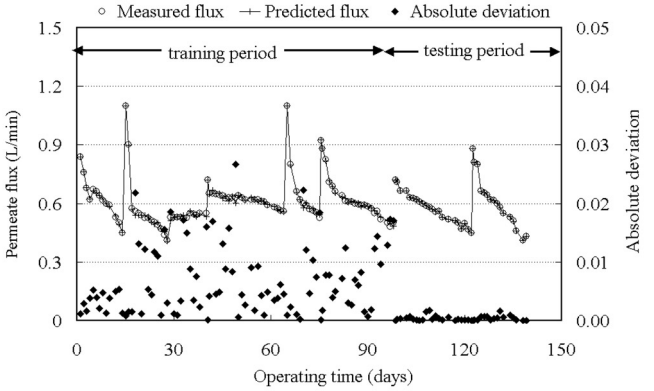


**Figure 5** Prediction of flux behavior and prediction

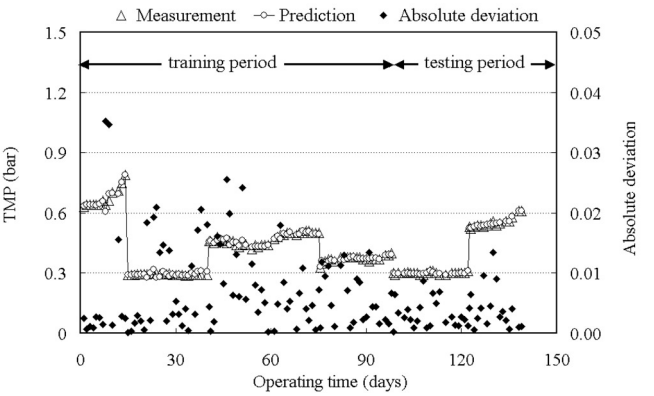


**Figure 6** Comparison between TMP measurement evaluation of absolute errors

predicted. Figure 5 also describes the model fit with absolute errors in the learning and testing phases. Mean absolute deviation of 0.011 and 0.004 in each learning and testing allowed good training and predicting for flux behavior but these deviations increased slightly in both initial flux and final flux in a filtration duration. Maximum deviation of 0.0017 in the testing period corresponded to 0.102 mL/hr of flux.

Figure 6 shows the comparison between the measured and estimated TMP during the learning and testing periods. When the network was trained with temperature, turbidity, $UV_{254}$, linear velocity and filtration time, the 2L-6N structure satisfactorily predicted the TMP in a filtration cycle with mean absolute deviation of 0.019 and 0.008 in each learning and testing period. Maximum deviation of 0.013 in the testing period corresponded to 0.013 bar of TMP.

## Conclusions

A KNT-neural network model was designed to predict the time evolution of flux from available data on raw water quality and operating conditions of a pilot plant of an ultrafiltration membrane. Five inlet data with temperature, turbidity, $UV_{254}$, inlet pressure and filtration duration were suitable to predict the permeate flux in each filtration cycle. Viscosity effect by temperature, fouling effect by both turbidity and $UV_{254}$, and foulant loads with applied pressure and filtration duration were sufficiently trained for the neural network. Consequently, a KNT-neural network with 3 layers and 5 neurons was used to good accuracy for predicting the flux behavior. The correlation coefficients were 0.997 and 0.999, respectively, in the learning and testing periods. However, if a neural network is more

trained by the flux recovery rate from backwashing and chemical cleaning, and fouling dynamics from supplementary experiments, the validity of this network model can be rather improved. With this good prediction, we can evaluate the productivity by the predicted flux in a given filtration duration. Then we can optimize the membrane module for given water demand and use the model in monitoring the membrane operation by flux behavior estimation. This model can also provide criteria for the time to prepare backwashing and chemical cleaning.



## References


Campos, C., Schimmoller, L., Marinãs, B.J., Snoeyink, V.L., Baudin, I. and Laîné, J.M. (2000). Adding PAC to remove DOC. *J. AWWA*, **92**(8), 69–83.

Delgrange, N., Cabassud, C., Cabassud, M., Durand-Bourlier, L. and Laîné, J.M. (1998). Neural networks for prediction of ultrafiltration transmembrane pressure-application to drinking water production. *J. Membr. Sci.*, **150**, 111–123.

EPA (2001). Low-pressure membrane filtration for pathogen removal: Application, implementation, and regulatory issues. EPA **815-C-01-001**, Office of Water, USA.

Grewal, M.S. and Andrews, A.P. (2001). *Kalman filtering: Theory and practice using Matlab*. 2nd edn, John Wiley & Sons, New York, pp. 1–24.

Hong, S.K., Faibish, R.S. and Elimelech, M. (1997). Kinetics of permeate flux decline in crossflow membrane filtration of colloidal suspensions. *J. Colloid Interface Sci.*, **166**, 267–277.

Jones, K.L., Odderstol, E.S., Wetterau, G.E and Clark, M.M. (1993). Using a hydraulic model to predict hollow-fiber UF performance. *J. AWWA*, **85**, 87–97.

Lahoussine-Turcaud, V., Wiesner, M.R., Bottero, T. and Mallevialle, J. (1990). Coagulation in pretreatment for ultrafiltration of a surface water. *J. AWWA*, **82**(12), 76–81.

Murase, H., Koyama, S. and Ishida, R. (1994). *Kalman neuro computing by personal computers*. Morikita Publication, Tokyo, pp. 24–83.

Müller, B., Reinhardt, J. and Strickland, M.T. (1995). *Neural networks/An introduction*. Springer, Berlin, pp. 3–71.

Oh, H.K., Yu, M.J. and Gwon, E.M. (2001). Prediction of permeate flux through UF membrane with low molecular weight cutoffs. *J. Korean Society of Environ. Eng.*, **23**, 1515–1526.

Teodosiu, C., Pastravanu, O. and Macoveanu, M. (2000). Neural network models for ultrafiltration and backwashing. *Wat. Res.*, **34**, 4371–4380.

Wang, L. and Song, L. (1999). Flux decline in crossflow microfiltration and ultrafiltration: experimental verification of fouling dynamics. *J. AWWA*, **160**, 41–50.