Modelling of submerged membrane flocculation hybrid systems using statistical and artificial neural networks methods

L. Erdei, S. Vigneswaran and J. Kandasamy

ABSTRACT

Hybrid membrane filtration processes involve complex physical, chemical and biological phenomena, thus their mechanistic modelling is challenging. The chief advantages of statistical and artificial neural networks (ANN) models (data-driven models) are that they do not require assumptions and simplifications to establish relationships from data. This paper investigates the characteristics and performance of several data-driven methods to model a hybrid membrane system. The focus is on the application of regression analysis and artificial intelligence based methods to a steady-state system. Among empirically based approaches, ANN neural networks methods were found to be very useful to predict permeate quality and membrane fouling. In the past multivariate nonlinear regression had barely been investigated for process modelling in water and waste water treatment. In this study polynomial multivariate nonlinear regression showed a superior performance. Multivariate parametric nonlinear models could match the performance of the nonparametric ANN models in the empirical modelling of complex systems, especially when combined with advanced optimization methods. This paper gives the methodology of how one could optimize a membrane hybrid system using ANN, validating it with one set of data. The same procedure/methodology can be applied to similar systems.

Key words | artificial neural networks, flocculation, mathematical modelling, membrane hybrid systems, multivariate parametric nonlinear regression, organics

L. Erdei
S. Vigneswaran (corresponding author)
J. Kandasamy
Faculty of Engineering and IT,
University of Technology Sydney (UTS),
PO Box 123,
Broadway NSW 2007,
Australia
Tel.: +612-9514-2641
Fax: +612-9514-2633
E-mail: *s.vigneswaran@uts.edu.au*

INTRODUCTION

Membrane fouling depends on numerous compounds present in natural and waste waters, whose concentrations can significantly change. Theoretical models, where they exist, rely on carefully controlled bench-scale and pilot-scale experiments that enable the study of influencing factors in isolation. The chief advantages of statistical and artificial neural networks (ANN) models (data-driven modelling) are that it does not require assumptions and simplifications, and it establishes relationships from data.

The common approach in regression is trial and error, which involves the construction of charts from the available data, and their matches against various mathematical (basis or prototype) functions. Regression models incorporating higher-order polynomial functions were presented by previous research (Vaccari & Christodoulatos 1992), and their capabilities to model activated sludge waste water plants was demonstrated. In a multivariate polynomial function, the higher-order terms can approximate curvilinear patterns in the response, while the cross-product terms can account for interactions between predictor variables.

Artificial neural networks (ANN, or more commonly neural networks, also neural nets) are powerful data analysis methods that represent one of the biologically inspired paradigms in computing. There are many types of neural networks but the most commonly used is

doi: 10.2166/aqua.2010.064

the multilayer perceptron (MLP). An MLP consists of multiple simple processing units (called perceptrons, neurons or nodes) that communicate via connections or links (McCulloch & Pitts 1943).

Perhaps the first reported use of artificial neural networks for membrane separation modelling involved the microfiltration of raw cane sugar syrup feed stream (Dornier *et al.* 1995). This model enabled the prediction of fouling for both constant and variable transmembrane pressure and crossflow velocity.

The performance of MF/UF membrane plants depends on unsteady raw water characteristics and operating conditions, including backwash frequencies and durations. Delgrange *et al.* (1998) used neural networks to predict the incremental development of total hydraulic resistance in successive UF filtration cycles, which is a typical example of time series modelling (Box *et al.* 1994).

Accurate, though short-term (45 min), prediction of flux decline in a hollow fibre benchscale UF module was accomplished using MLP neural networks (Teodosiu *et al.* 2000). Recent research (Liu & Kim 2008) compared the prediction performances of the blocking laws and ANNs models in constant flowrate conditions using synthetic surface water. It was found that no single blocking law could fit membrane fouling over the entire experimental range, and the ANN models showed excellent agreement between experimental data and predictions.

The use of neural networks also was successful in the modelling and control of coagulation and flocculation processes (Gagnon *et al.* 1997; Maier *et al.* 2004). Stanley *et al.* (2000) concluded that artificial neural networks could successfully predict both turbidity and organic removal by enhanced coagulation. They also noted that historical plant data naturally incorporates the experience of operators, thus such models in this sense are knowledge-based. The models developed were suitable for predictive clarifier control and as a 'virtual laboratory' for process simulation and training.

While ANN applications are widely reported in the literature, the use of statistically based alternative methods is scarce. Multivariate polynomial regression is limited to experimental design using the surface response methodology (Box & Draper 1987), and is typically restricted to second-degree polynomials. Niemi & Palosaari (1993) briefly mentioned the use of this method in the calculation of flux decline and rejection for UF and RO membranes in bench-scale experiments. They found that theoretical models showed large errors of estimation, up to 25%. Multivariate polynomial regression required more experimental data but the resulting error range was only between 5% and 10%.

Hybrid membrane filtration processes involve very complex physical, chemical and biological phenomena, and their limited understanding hinders efficient application. Although traditional mechanistic models are not available at present, the application of data-driven modelling methods can provide valuable insights and effective tools for the design and operation of hybrid membrane systems.

This paper investigates the characteristics and performance of several data-driven methods to model a hybrid membrane system. The focus is on the application of regression analysis and artificial-intelligence-based methods to a steady-state system. The fundamental characteristic of the present approach is its data-driven nature, which allows its straightforward adaptation for many other complex processes, and can be extended to time-dependent and dynamic systems.

CONCEPTUAL MATHEMATICAL MODEL OF THE HYBRID MEMBRANE SYSTEM

The preliminary mathematical model of the hybrid membrane system is considered as a black box, which is defined by input and output observed data (Figure 1). These data describe the behaviour of a system and, through system identification, allow the development of a model that accounts for that behaviour. This system has three output parameters, namely permeate pH, DOC and TMP. The output parameters are controlled by the input parameters: coagulant dosage and permeate volume.

The pH of the hybrid system is naturally controlled by the coagulant dose and waste water alkalinity in the upstream process (in-line flocculation), and the downstream (direct membrane filtration) process does not influence this parameter significantly.

While permeate volume controls membrane fouling in dead-end filtration, this parameter can be replaced with

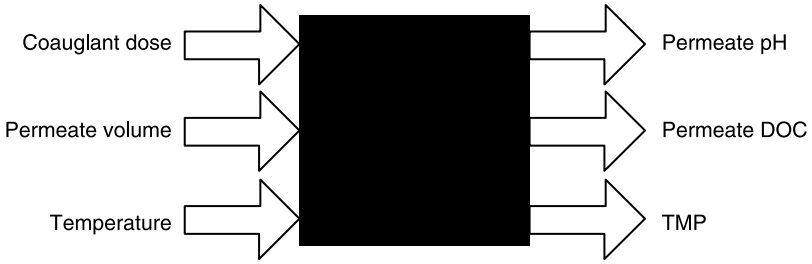

Figure 1 | Conceptual model of the hybrid membrane system.

filtration time, which is a proportional scalar quantity for constant flux filtration.

The effect of temperature on physicochemical and biological processes is usually significant and may affect both fouling and permeate quality. It was shown experimentally that a temperature change between 17 and 36°C did not influence significantly membrane material and cake properties (Boerlage *et al.* 2003). For hybrid membrane systems, temperature also may affect the in-line flocculation pre-treatment process. Kang & Cleasby (1995) examined the effect of temperature on flocculation. Particle destabilization and the zeta potential are slightly affected but particle agglomeration, enmeshment, floc size and strength vary significantly with temperature. They found that the physical phenomena (i.e. mixing) had only a minor role and the changes could be attributed to the effect of changed pH. The estimated change of pH is only 0.1 when the temperature varies between 20°C and 26°C, twice the range of actual changes that occurred during the experiments in this study. The small changes of pH result in negligible effects on solution chemistry (Kang & Cleasby 1995), on the removal of organic pollutants (Jiang *et al.* 1996), and on floc characteristics (Fitzpatrick *et al.* 2004). These results indicate that temperature can be treated as a latent variable whose effects are either known (fouling) or can be neglected (permeate quality) within the range of 20–26°C. These considerations allow the definition of the process model suitable for system identification (Figure 2).

Changes in permeate quality (in terms of pH and DOC) are controlled by the applied coagulant dosage, and do not vary significantly with filtered time. The development of TMP (fouling) depends on both filtration time T and coagulant dosage D. The exact details of the pH change, DOC reduction mechanism and fouling are not considered in the modelling process. Total DOC removal (TDR) is calculated as a percentage of the influent waste water.

EXPERIMENT, ANALYSIS AND MODELLING

Materials

This study was carried out with synthetic waste water. This synthetic waste water represents the biologically treated sewage effluent (BTSE) and its composition is given in Table 1 (Seo *et al.* 1997).

The chemicals used in experiments were of analytical grade, purchased from Sigma-Aldrich. For pH adjustment, hydrochloric acid, sodium hydroxide and hydrated lime were employed.

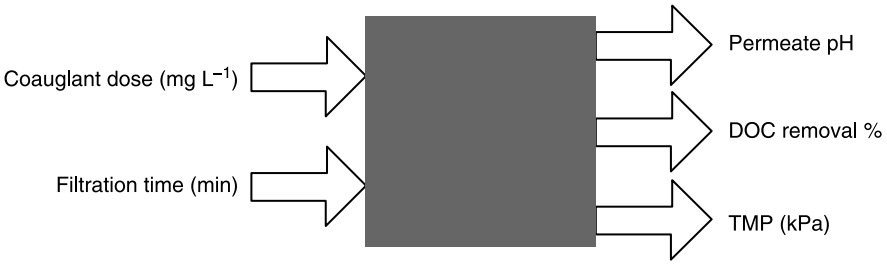

Figure 2 | Process model for system identification.

Table 1 | Constituents and characteristics of the synthetic waste water

Constituents	Concentration (mg L^{-1})	Fraction by DOC
Beef extract	1.8	6.5%
Peptone	2.7	13.8%
Humic acid	4.2	8.2%
Tannic acid	4.2	23.7%
Sodium lignin sulfonate	2.4	6.7%
Sodium lauryle sulphate	0.94	4.2%
Acacia gum powder	4.7	21.3%
Arabic acid (polysaccharide)	5.0	15.6%
$(NH_4)_2{\cdot}SO_4$	7.1	
K_2HPO_4	7.0	
NH_4HCO_3	19.8	
$MgSO_4{\cdot}7H_2O$	0.71	

Hybrid in-line flocculation–membrane filtration system

A schematic diagram of the laboratory-scale membrane filtration system used in conjunction with in-line flocculation ($FeCl_3$) pre-treatment is illustrated in Figure 3.

In each experiment the membrane tank was filled with waste water, and aeration (membrane bubbling) commenced at 90 L m^{-2} h^{-1} (volumetric) air flow rate. Compressed air was introduced for aeration via a soaker hose at the bottom of the tank, which also provided a hydrodynamic scouring effect to reduce membrane fouling. The characteristics of the Mitsubishi hollow fibre membrane module are shown in Table 2.

Synthetic waste water was pumped into the membrane tank at constant flow rates using a peristaltic pump (Easy-FLOW-VS model, TPS, Australia). The concentration of the stock coagulant solution varied according to the required dosage and the delivery rate of the coagulant pump (101UI model, Watson Marlow, USA). A tee fitting connected the feed and the coagulant ($FeCl_3$) dosing tubes, and small-bore tubes (3 mm internal diameter) provided 10–60 s flash mixing time for coagulation. The feed was forced through the Mitsubishi-Rayon membrane module under suction, as provided by a peristaltic pump (EasyFLOW-VS model, TPS, Australia) and permeate was collected in a container for the backwash supply. The plant was controlled by a SCADA system (Smith *et al.* 2005; Smith *et al.* 2006).

The backwash regime consisted of 30 min filtration and 0.5 min of backflush with the backwash rate set at twice the rate of filtration. This regime was chosen to limit the length of experimental runs to manageable lengths of time. The latter was related to manual flow control, which required continual attendance to pump adjustment to maintain constant fluxes. The backwash flow was provided by a peristaltic pump (EasyFLOW-VS model, TPS, Australia).

In this study the available data sets consist of records of TMP with time steps and temperature, and spot measurements of pH and DOC for raw waste water and membrane permeate. Operational parameters (DO, temperature and TMP) were monitored and logged by a computer. Experimental data were corrected to a temperature of 20°C.

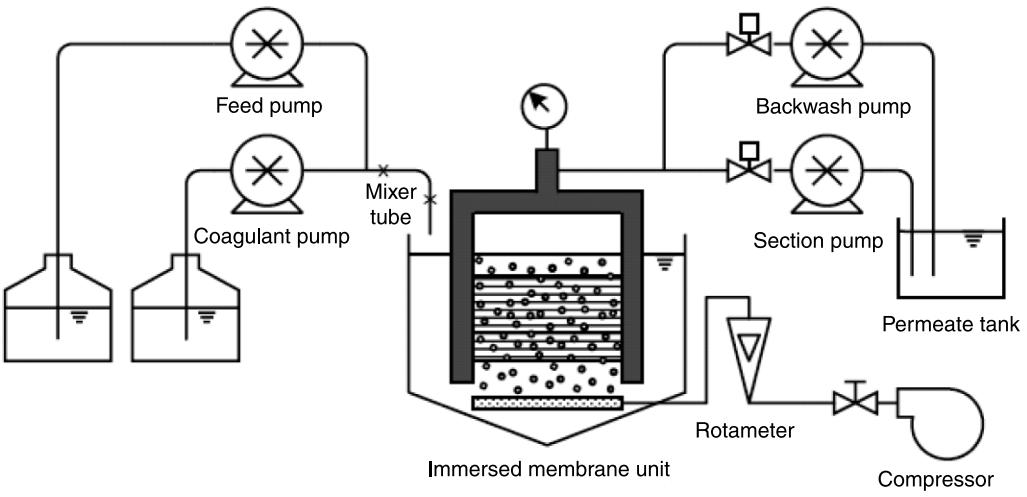

Figure 3 | Schematic diagram of the experimental bench-scale hybrid in-line flocculation–membrane system.

Table 2 | Characteristics of the Mitsubishi hollow fibre membrane module

Material	Polyethylene
Pore size (μm)	0.1
Fibre inner diameter (mm)	0.27
Fibre outer diameter (mm)	0.41
Number of fibres	320
Fibre length (mm)	120
Surface area (m^2)	0.05
Packing density ($m^2\,m^{-3}$)	9,858
Manufacturer	Mitsubishi-Rayon

The DOC content of grab waste water samples was determined using a Phoenix 8000 UV-Persulfonate TOC analyser. Since all samples were filtered with 0.45 mm Millex syringe membranes, the reported values represent DOC. pHs were measured with Hach 2000 and WTW pH-330 meters, respectively.

Membrane fouling data

The data set was obtained from constant flux experiments at $36\,L\,m^{-2}\,h^{-1}$ that determined the development of TMP (membrane fouling) for various ferric chloride coagulant doses. Time and corresponding TMP data were monitored under PLC control every 2 s, and the 30 s averages were displayed and logged by a SCADA system. The data is shown in Figure 4.

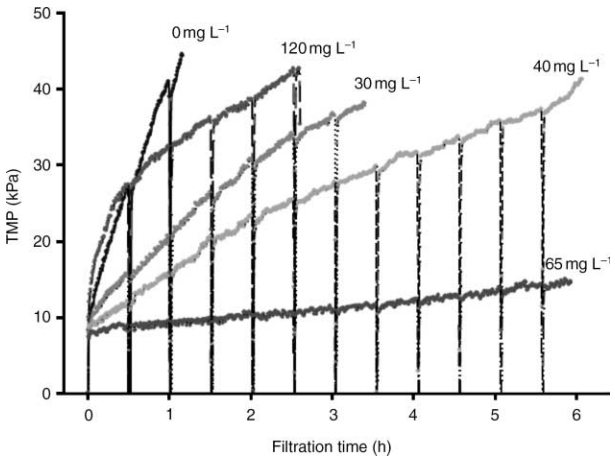

Figure 4 | Membrane fouling at a constant flux of $36\,L\,m^{-2}\,h^{-1}$ (Mitsubishi 0.1 μm pore size membrane) as a function of $FeCl_3$ coagulant dosage used for in-line flocculation.

Table 3 | Permeate water quality data for the system model

Coagulant dose (mg/L)	pH	TDR %
0	7.65	23.3
20	6.44	45.2
30	6.04	60.4
40	5.72	64.4
50	5.56	69.1
65	5.25	71.4
90	4.30	73.7
120	3.14	74.5

The evolution of TMP with time shows large abrupt changes at 30-min intervals in the form of drop lines. These transient patterns were caused by periodic membrane backflush that lasted for 0.5 min. These transient data are regarded as outliers for the overall (long-term, irreversible) development of TMP on a macro-scale. These outliers were removed by an interpolation technique described in detail in Erdei (2009).

Permeate quality data

The permeate water quality data for modelling is summarized in Table 3, as a function of the dose of $FeCl_3$ coagulant used for in-line flocculation.

Modelling methods and verifications

Computations and simulations were carried out mainly in the MATLAB R2007a (MathWorks, USA) environment. In two instances, other (freely available) software was used for computations, which are referenced in the relevant sections. These applications ran on a personal computer.

RESULTS AND DISCUSSION

Multivariate polynomial regression analysis of membrane fouling

A single global (aggregate) model is developed that shares the complete data set and provides common parameters. A second degree bivariate polynomial function (Equation (1)) was used as a baseline nonlinear model for the

Table 4 | Summary statistics of bivariate polynomial models

Equation	2nd	3rd	4th	6th	2nd SR*	6th SR*
Least-squares method, SSE	105,918	2,819	1,156	414	2,331	788
Mean squared error, MSE	23.1	1.2	0.5	0.2	1	0.3
Root mean squared error, RMSE	4.8	1.1	0.7	0.4	1	0.6
Maximum of absolute errors, ME	24.2	11.9	8.5	2.2	8.2	6.3
Coefficient of regression, R^2	0.8240	0.9953	0.9981	0.9986	0.9922	0.9970
No. of terms	6	10	15	27	9	15

*Stepwise regression model.

prediction of TMP. Fitting this model by solving the related system of equations provides a mathematical model with coefficients given in the Appendix.

$$\mathrm{TMP} = a + bT + cD + dT^2 + \mathrm{e}TD + fD^2 \quad (1)$$

where T is filtration time (min) and D is the coagulant dose (mg/L).

The fit of this model is demonstrated in Table 4. Third-, fourth- and sixth-degree multivariate polynomial functional models of TMP development with time, respectively, with coefficients are given in the Appendix, and their fits are shown in Table 4. The sixth-degree model could achieve excellent fit (Figure 5). The fourth-degree model reasonably follows the observed values, but it still requires improvement to eliminate the prediction errors for 120 mg/L coagulant dose (Figure 6(a)).

The use of higher-order polynomials increased model performance but also significantly increased the size of the resulting mathematical equations. Large models can be reduced using stepwise regression (Hocking 1976), which is an important method to ensure model parsimony. Stepwise regression involves adding terms to and removing from the model until further change does not reduce, or increase, the prediction error significantly. In this way, individual variables and/or terms that contribute little to the overall performance can be effectively identified and removed.

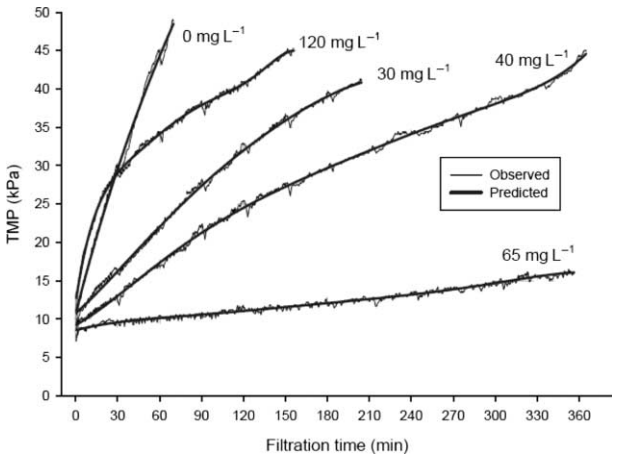

Figure 5 | Sixth-degree bivariate polynomial fit of membrane fouling.

Stepwise regression was applied to the second-degree bivariate polynomial regression model by adding three additional terms (Equation (2)) with coefficients provided in the Appendix. This dramatically improved the fit of model (Equation (2)). The increased performance is clearly demonstrated in Table 4.

$$\mathrm{TMP} = a + bT + cD + dT^2 + eTD + fD^2 + gT^3D + hT^2D^2 + iTD^2 \quad (2)$$

The sixth-degree polynomial model was trimmed by the use of manual stepwise regression from 49 terms to 15 terms, which allows direct comparison with the fourth-degree model that also has 15 terms, Table 4. Observation of Figure 6(b) indicates a reasonable fit, which is clearly better than is shown by the fourth-degree polynomial model of comparable complexity.

Experiments with various models also revealed that the inclusion of higher-order terms in equations is desirable. Higher-order polynomial terms are sensitive to small changes in curve trends, and can follow the variations caused by different coagulant doses in each experiment as illustrated by Figure 6a and b.

ANN modelling of membrane fouling

Artificial neural networks (ANNs, or more commonly neural networks, also neural nets) are powerful data

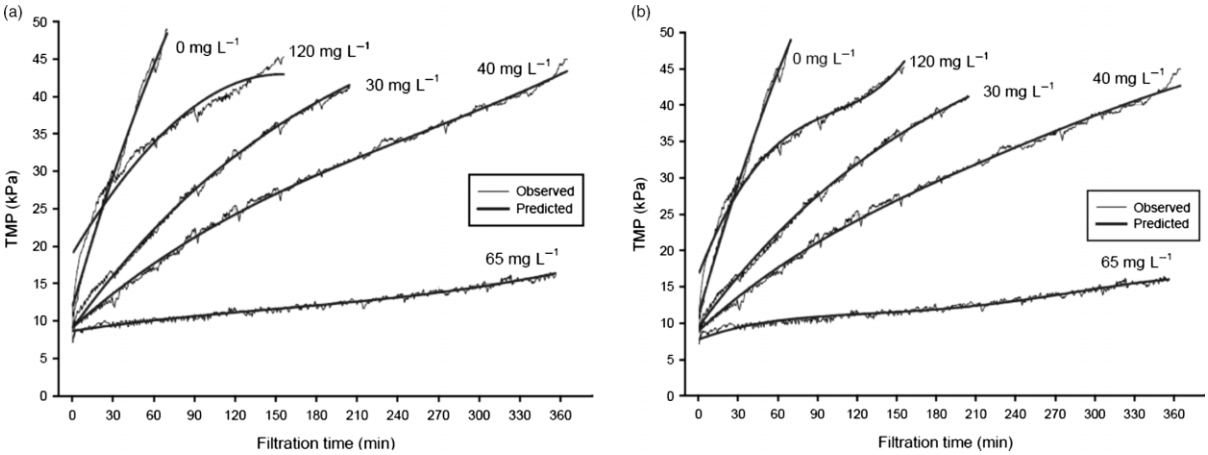

Figure 6 | Comparison of (a) second-degree bivariate polynomial fit, (b) sixth-degree bivariate polynomial fit of membrane fouling using stepwise regression.

analysis methods that represent one of the biologically inspired paradigms in computing. An ANN consists of multiple simple processing units (called perceptron, neuron or node) that communicate via connections or links (McCulloch & Pitts 1943). Figure 7 shows a multiple-input neuron, which is the standard building block of these networks.

A neuron usually has multiple inputs (independent variables) $x_1, x_2, \ldots, x_M$, as shown in Figure 7. The neuron multiplies each of its inputs with the corresponding weights $w_{1,1}, w_{1,2}, \ldots, w_{1,M}$ of the weight matrix. There is a special connection called bias b, whose value is set to 1, and summed with the weighted inputs to obtain the net input n.

$$n = w_{1,1}p_1 + w_{1,2}p_2 + \ldots + w_{1,M}p_M + b$$

or more concisely in matrix form

$$n = \boldsymbol{W}\boldsymbol{p} + b$$

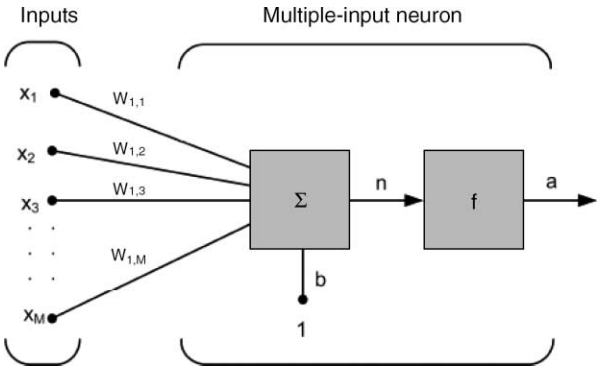

Figure 7 | Multiple-input neuron.

The neuron applies a transfer function to input n to obtain output a i.e.

$$a = f(\boldsymbol{W}\boldsymbol{p} + b)$$

There are many types of neural networks but the most commonly used is the multilayer perceptron (MLP). In MLP, the transfer function can be linear (especially for the output neuron) but typically is a nonlinear, S-shaped, smooth and easily derivable function, such as the log-sigmoid or tanh.

The power of neural networks results from their ability to map input and output data. This ability is often 'learned' under supervision that involves the repeated presentation of examples (input and desired output data) to the network. During training, the network calculates a predicted value using the actual input, and compares it with the desired value to obtain the error, for each example, after each other. The learning algorithm iteratively reduces the overall error (usually the SSE) in each training cycle by adjusting the weights of all links, starting from the last layer and progressing backward from layer to layer towards the input.

Other ANNs include Radial Base Function (RBF) Neural Network and General Regression Neural Network (GRNN). In RBF the transfer function is based on the Gaussian (bell-shaped) function. GRNN uses transfer functions approximated from statistical techniques. Details of these ANN models are given elsewhere (Box & Draper 1987). Adaptive logic networks (ALNs) represent an alternative computing paradigm to neural networks, and approximate transfer functions with piecewise straight lines and simple shapes.

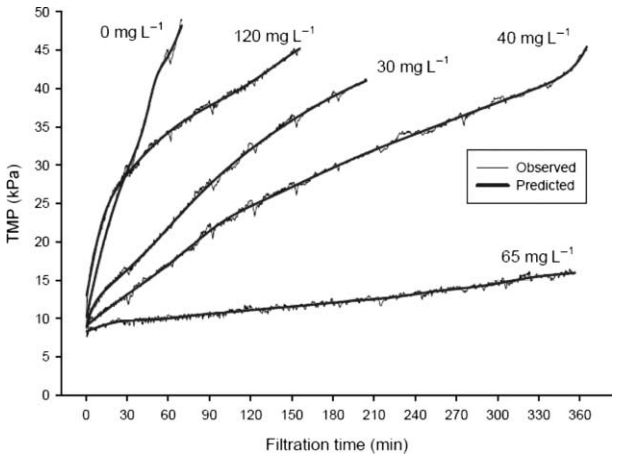

Figure 8 | MLP model fit of membrane fouling using 15 nodes in the first, and 6 nodes in the second hidden layer.

MLP neural network

The applied MLP network with two hidden layers could accommodate local changes of variables providing reasonable fits at initial and final (extreme) values of TMP, Figure 8. This network has 15 and 6 hidden nodes in the first and second layer, respectively, with sigmoid transfer functions. The data was randomly split into 68% training, 16% validation and 16% testing sets. Using small learning rates of 0.1 with the QuickProp training algorithm (Fahlman 1988) over 5,000 training cycles, this model consistently achieved very good fits.

The trend lines smoothly predict the experimental data, and no overfitting can be observed. The fitting statistics are also excellent (Table 5).

Radial base function (RBF) neural network

Figure 9 shows the performance of a radial base function (RBF) network that had 50 neurons in the hidden layer with a spread of 40, and was trained using the complete data set.

This network was considered as an adequate candidate for modelling, since it has a performance largely comparable to that of the MLP model, except for the twice as large maximum error (Table 5). The increase of node numbers in this model reduced the SSE, but this had an undesirable effect in the form of local overfitting. This effect was especially noticeable for the 40 mg/L dosage curve, where the network started to fit short-range changes of TMP observable at backwash events.

It is fair to state that RBF networks are neither superior (contrary to claims by Chen & Kim (2006)) nor inferior to MLP networks, rather that they have particular characteristics to suit different requirements. RBF models train fast, are more suitable for localized approximations, but do not generalise well for regions where there is little or no input-output data available (Moody & Darken 1989). For the same reasons, they have very poor extrapolation capabilities. In contrast, MLP networks learn relatively slowly, but their approximation alters the entire input-output response domain to make them more suitable for global models.

General regression neural network (GRNN)

GRNN networks are primarily intended for function approximation and are very easy to design. Figure 10 shows the fit of a GRNN network model of fouling.

This model used a spread value of one, and has almost as good a fit as the RBF model (Table 5). The design of GRNN networks is faster than of RBF networks, and involves only the estimation of the spread value. Larger spread values resulted in smoother curves but increased the error. It is worthy to note that GRNN networks cannot extrapolate at all. Thus, if they are trained on subsets of data, care must be taken to include the extreme values of input variables in the data set.

Table 5 | Summary statistics of ANN models

Equation	MLP	RBF	GRNN	ALN*
Least-squares method, SSE	316	337	353	551
Mean squared error, MSE	0.15	0.146	0.15	0.23
Root mean squared error, RMSE	0.37	0.38	0.39	0.49
Maximum of absolute errors, ME	2.28	3.6	3.23	4.77
Coefficient of regression, R^2	0.999	0.999	0.999	0.998

*Not classed as an ANN model but included for comparison.

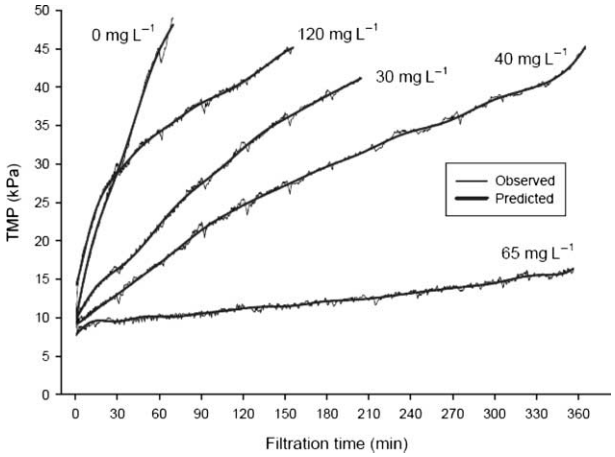

Figure 9 | RBF model fit of membrane fouling using 50 nodes and a spread value of 40.

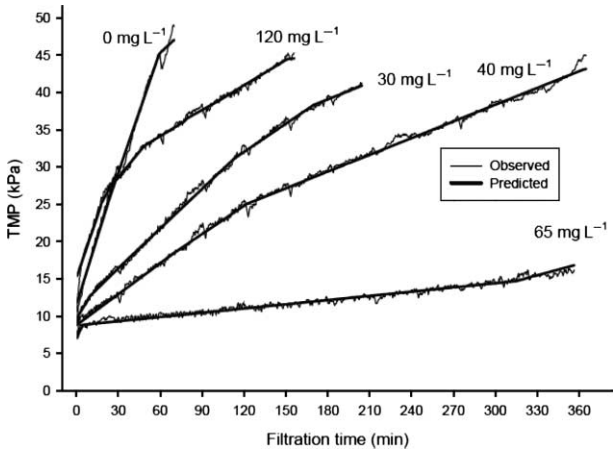

Figure 11 | Adaptive logic network model fit of membrane fouling with constrained training and single tree.

Adaptive logic network modelling of membrane fouling

Figure 11 illustrates an ALN model of TMP evolution, produced with the ALNFit Pro software (Dendronic Decisions).

This simple model was specifically produced to show the use of piecewise line sections for curve approximation by this method. Nevertheless, even this simple model has reasonable statistics of fit (Table 5). Like neural networks, ALNs can learn and approximate any continuous function to arbitrary accuracy. For the illustrated model, the data set was split into three portions, 30%, 25% and 35%, for testing, training and validation, respectively. Adaptive logic networks train relatively fast and their training times are comparable to those of RBF networks. From a functional (i.e. user) point of view, ALN is a non-parametric modelling tool, a lesser known but very powerful alternative to traditional neural networks.

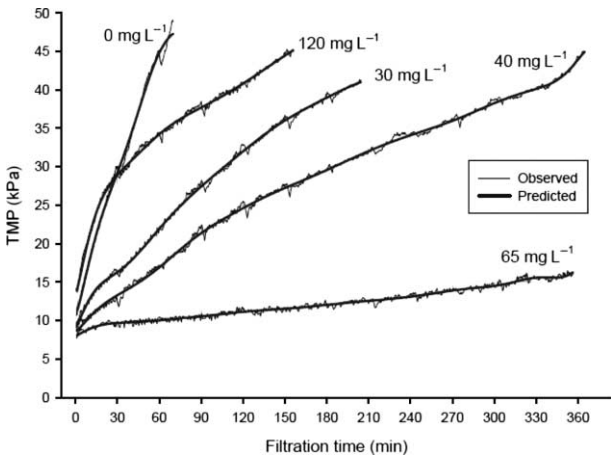

Figure 10 | GRNN model fit of membrane fouling using a spread value of 1.

CONCLUSIONS

Hybrid membrane filtration processes involve complex physical, chemical and biological phenomena, thus their mechanistic modelling is challenging. Among empirically based approaches, data-driven modelling using neural networks historically has been found to be very useful to predict permeate quality and membrane fouling, while multivariate nonlinear regression still barely has been investigated for process modelling in water and waste water treatment.

This paper gives the methodology of how one could optimize the membrane hybrid system using ANN, validating it with one set of data. The same procedure/methodology can be applied to similar systems.

Polynomial multivariate nonlinear regression showed a superior performance. The accuracy of the model was easily improved by using higher-order polynomials, and a sixth-degree polynomial equation showed excellent statistics with an SSE of 414 and R^2 of 0.9986. Nevertheless, the use of full polynomials results in relatively large, non-parsimonious models. The use of stepwise regression yielded more compact models with better fits than full polynomial models of similar complexity.

A comprehensive evaluation of MLP networks indicated that two hidden layers are essential to provide small, parsimonious networks. It was shown that the performance

of a neural network using 15 and 6 neurons in the hidden layers was excellent.

GRNN neural networks were extremely simple to construct, and yielded models with fitting quality comparable to RBF neural networks. ALNs represent a very powerful alternative to traditional neural networks. The obtained models in all aspects could match those obtained using neural networks.

A major finding of this study was that the hitherto unutilised multivariate parametric nonlinear models could match the performance of the non-parametric ANN models in the empirical modelling of complex systems, especially when combined with advanced optimization methods. Such models are easy to inspect, to audit and to use in engineering practice.

ACKNOWLEDGEMENTS

The research is funded by an Australian Research Council Discovery Project.

REFERENCES

Boerlage, S. F. E., Kennedy, M., Aniye, M. P. & Schippers, J. C. 2003 Applications of the MFIi-UF to measure and predict particulate fouling in RO systems. *J. Membr. Sci.* **220**(1–2), 97–116.

Box, E. P. G. & Draper, R. N. 1987 *Empirical Model-building and Response Surfaces*, 1st edition. Wiley, New York.

Box, G., Jenkins, M. G. & Reinsel, C. G. 1994 *Time Series Analysis: Forecasting and Control*, 3rd edition. Prentice Hall, Upper Saddle River, NJ.

Chen, H. & Kim, A. S. 2006 Prediction of permeate flux decline in crossflow membrane filtration of colloidal suspension: a radial basis function neural network approach. *Desalination* **192**(1–3), 415–428.

Delgrange, N., Cabassud, C., Cabassud, M., Durand-Bourlier, L. & Lainé, J. M. 1998 Modelling of ultrafiltration fouling by neural network. *Desalination* **118**(1–3), 213–227.

Dornier, M., Decloux, M., Trystram, G. & Lebert, A. 1995 Dynamic modeling of crossflow microfiltration using neural networks. *J. Membr. Sci.* **98**(3), 263–273.

Erdei, L. 2009 *Development of low-pressure hybrid membrane filtration systems for advanced wastewater treatment*. PhD Thesis, Faculty of Engineering and IT, University of Technology, Sydney, Australia.

Fahlman, S. E. 1988 Faster-learning variations on back-propagation: An empirical study. *Proceedings, 1988 Connectionist Models Summer School*, Morgan-Kaufmann, Pittsburgh, Pennsylvania, 17–26 June.

Fitzpatrick, C. S., Fradin, E. & Gregory, J. 2004 Temperature effects on flocculation, using different coagulants. *Water Sci. Technol.* **50**(50), 171–175.

Gagnon, C., Grandjean, B. P. A. & Thibault, J. 1997 Modelling of coagulant dosage in a water treatment plant. *Artif. Intell. Eng.* **11**(4), 401–404.

Hocking, R. R. 1976 The analysis and selection of variables in linear regression. *Biometrics* **32**, 1–49.

Jiang, J.-Q., Graham, N. J. D. & Harward, C. 1996 Coagulation of upland coloured water with polyferric sulphate compared to conventional coagulants. *JWSRT-AQUA* **45**(3), 143–154.

Kang, L.-S. & Cleasby, J. L. 1995 Temperature effects on flocculation kinetics using Fe(III) coagulant. *J. Environ. Eng.* **121**(12), 893–901.

Liu, Q.-F. & Kim, S.-H. 2008 Evaluation of membrane fouling models based on bench-scale experiments: a comparison between constant flowrate blocking laws and artificial neural network (ANNs) model. *J. Membr. Sci.* **310**(1–2), 393–401.

Maier, H. R., Morgan, N. & Chow, C. W. K. 2004 Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters. *Environ. Model. Softw.* **19**(5), 485–494.

McCulloch, S. W. & Pitts, W. 1943 A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **5**, 115–133.

Moody, J. E. & Darken, C. J. 1989 Fast learning in networks of locally-tuned processing units. *Neural Comput.* **1**, 281–294.

Niemi, H. & Palosaari, S. 1993 Calculation of permeate flux and rejection in simulation of ultrafiltration and reverse osmosis processes. *J. Membr. Sci.* **84**(1–2), 123–137.

Seo, G. T., Ohgaki, S. & Suzuki, Y. 1997 Sorption characteristics of biological powdered activated carbon in BPAC-MF (biological activated carbon microfiltration) system for refractory organic removal. *Water Sci. Technol.* **35**(7), 163–170.

Smith, P. J., Vigneswaran, S., Ngo, H. H., Nguyen, H. & Aim, R. B. 2005 Design of a generic control system for optimizing back flush durations in a submerged membrane hybrid reactor. *J. Membr. Sci.* **255**(1–2), 99–106.

Smith, P. J., Vigneswaran, S., Ngo, H. H., Nguyen, H. & Aim, R. B. 2006 Application of an automation system and a supervisory control and data acquisition (SCADA) system for the optimal operation of a membrane adsorption hybrid syste. *Water Sci. Technol.* **53**(4–5), 179–184.

Stanley, S. J., Baxter, C. W., Zhang, Q. & Shariff, R. 2000 *Process Modelling and Control of Enhanced Coagulation*. AWWARF, Denver.

Teodosiu, C., Pastravanu, O. & Macoveanu, M. 2000 Neural network models for ultrafiltration and backwashing. *Water Res.* **34**(18), 4371–4380.

Vaccari, D. A. & Christodoulatos, C. 1992 Generalized multiple regression techniques with interaction and nonlinearity for system identification in biological treatment process. *Instrum. Soc. Am. Trans.* **31**(1), 97–102.

First received 8 June 2009; accepted in revised form 7 December 2009

APPENDIX

Appendix | Coefficients of polynomial regression equations

Parameter	2nd	3rd	4th	6th	2nd SR*	6th SR*
a	26.7561	1.43×10^{1}	1.17×10^{1}	1.07×10^{1}	1.15×10^{1}	1.12×10^{1}
b	0.2951	5.07×10^{-1}	6.34×10^{-1}	9.04×10^{-13}	6.34×10^{-1}	1.39×10^{-13}
c	− 0.9049	-1.42×10^{-1}	-2.94×10^{-1}	2.84×10^{-1}	3.40×10^{-5}	-5.84×10^{-2}
d	− 0.0003	-5.66×10^{-4}	-1.63×10^{-3}	-2.75×10^{-11}	1.40×10^{-3}	4.18×10^{-12}
e	− 0.0023	-1.11×10^{-2}	-1.51×10^{-2}	7.10×10^{-1}	-1.08×10^{-1}	6.41×10^{-1}
f	0.0081	-1.11×10^{-3}	1.07×10^{-2}	-3.16×10^{-2}	-9.95×10^{-4}	-1.75×10^{-2}
g		6.95×10^{-7}	9.36×10^{-7}	5.54×10^{-4}	-2.83×10^{-7}	1.63×10^{-4}
h		2.69×10^{-6}	4.44×10^{-5}	-7.39×10^{-15}	-1.78×10^{-2}	-3.10×10^{-16}
i		6.98×10^{-5}	7.19×10^{-5}	-2.91×10^{-3}	1.26×10^{-4}	-1.45×10^{-3}
j		2.38×10^{-5}	-1.56×10^{-4}	-3.69×10^{-6}		-4.82×10^{-7}
k			-2.70×10^{-11}	-6.87×10^{-7}		3.10×10^{-8}
l			-1.14×10^{-8}	-1.08×10^{-18}		1.48×10^{-20}
m			-3.18×10^{-7}	2.10×10^{-4}		3.90×10^{-5}
n			2.56×10^{-7}	-2.58×10^{-6}		-3.79×10^{-7}.
o			7.59×10^{-7}	5.05×10^{-11}		-8.35×10^{-13}
p				-3.89×10^{-17}		
q				1.36×10^{-20}		
r				1.83×10^{-14}		
s				-1.36×10^{-2}		
t				1.42×10^{-4}		
u				-2.91×10^{-23}		
v				1.19×10^{-19}		
w				2.55×10^{-8}		
x				-3.96×10^{-12}		
y				5.16×10^{-6}		
z				1.38×10^{-12}		
a1				-1.63×10^{-16}		

*Stepwise regression.

2nd order $\mathrm{TMP} = a + bT + cD + dT^2 + eTD + fD^2$

3rd order $\mathrm{TMP} = a + bT + cD + dT^2 + eTD + fD^2 + gT^3 + hTD^2 + iTD^2 + jD^3$

4th order $\mathrm{TMP} = a + bT + cD + dT^2 + eTD + fD^2 + gT^3 + hTD^2 + iTD^2 + jD^3 + kT^4 + lT^3D^2 + mT^2D^2 + nTD^3 + oD^4$

6th order $\mathrm{TMP} = a + bT + cD + dT^2 + eTD + fD^2 + gT^3 + hTD^2 + iTD^2 + jD^3 + kT^4 + lT^3D^2 + mT^2D^2 + nTD^3 + oD^4 + oT4D2 + pT5D4 + qT5D6 + rT^3D^5 + sD^2 + tD^5 + uT^6D^6 + vT^6D^4 + wT^4 + xT^5D + yT^3 + zT^3D^4 + alT^4D^5$

Stepwise regression

2nd order $\mathrm{TMP} = a + bT + cD + dT^2 + eTD + fD^2 + gT^3D + hT^2D^2 + iTD^2$

6th order $\mathrm{TMP} = a + bTD6 + cD + dD6 + eT + fTD + gTD2 + hT^2D^6 + iT^2 + jTD^3 + kT^3D + lT^4D^6 + mT^2D + nT2D^2 + oT^4D2$

where T is filtration time (min) and D is the coagulant dose (mg/L).