Journal of Membrane Science xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

Journal of Membrane Science

journal homepage: www.elsevier.com/locate/memsci

# Neural network approach for modeling the performance of reverse osmosis membrane desalting

Dan Libotean [a], Jaume Giralt [a], Francesc Giralt [a], Robert Rallo [b], Tom Wolfe [c], Yoram Cohen [d,*]

[a] Fenomens de Transport, Departament d'Enginyeria Quimica, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Catalunya, Spain
[b] Fenomens de Transport, Departament d'Enginyeria Informatica i Matematiques, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Catalunya, Spain
[c] Toray Membrane USA, 12233 Thatcher Court, Poway, CA 92064, United States
[d] Water Technology Research Center, Chemical and Biomolecular Engineering Department, 5531 Boelter Hall, University of California, Los Angeles, CA 90095-1592, United States

ARTICLE INFO

*Article history:*
Received 1 December 2007
Received in revised form 8 October 2008
Accepted 18 October 2008
Available online xxx

*Keywords:*
Reverse osmosis
Artificial neural networks
Desalination
Flux decline
Salt passage
Process performance forecasting

ABSTRACT

A neural network-based modeling approach with back-propagation and support vector regression algorithms was investigated as a mean of developing data-driven models for forecasting reverse osmosis (RO) plant performance and for potential use for operational diagnostics. The concept of plant "short-term memory" time-interval was introduced to capture the time-variability of plant performance since both a state of the plant model and standard time-series analyses for both flux decline and salt passage did not result in realistic predictive horizons for practical purposes. Past information of normalized permeate flux and salt passage were introduced as unique input variables along with process operating parameters to capture short-term plant performance variability. Sequential models, where the time-variation within each forecasting time-interval was also taken as input information, and marching forecasting models, where target values were predicted at fixed future times from past plant information, were developed. Models were trained, with normalized permeate flux and salt passage, for various model architectures, memory time-intervals and forecasting times using both back-propagation and support vector regression approaches. State of the plant models (without forecasting) were able to describe the relatively small permeate flux variations but were unable to capture salt passage trends (for any present time condition) since unsteady state phenomena could not be properly described without plant memory information. Forecasting of plant performance, with both sequential and marching models, yielded good predictive accuracy for short-term memory time-intervals in the range of 8–24 h for permeate flux and salt passage for forecasting times up to 24 h. Current work is ongoing to extend the approach for longer time scales and to incorporate data-driven forecasting models of RO plant into control strategies and process diagnostics.


## 1. Introduction

The separations performance of reverse osmosis (RO) desalination (e.g., salt rejection and permeate flux) and membrane longevity are impacted by numerous factors including, but not limited to, membrane fouling and scaling, consistency of feed water quality and treatment effectiveness, and stability and reliability of plant process equipment. The development of advanced RO process control strategies would benefit from predictive models of plant operation that are capable of identifying deviations (as well as upsets) of process conditions due to fouling and mineral salt scaling [1]. There are, however, major obstacles to developing first principle deterministic models for predicting RO plant behavior including: (a) the complexity and potential variability of feed composition, (b) difficulty of real-time quantification of feed water fouling and scaling propensities, (c) lack of practical and accurately predictive fouling models that account for the interplay of various fouling mechanisms, and (d) lack of information regarding the precise role of membrane surface properties and membrane interactions with various foulants/scalants and fouling/scaling precursors [2].

In recent years, there have been various attempts to advance the use of artificial neural networks (ANN) as a viable approach to develop data-driven models to describe the performance of membrane processes [3–14]. The available ANN models have described flux decline and variation in rejection performance for a given time-invariant feed quality [4,5,10–14]. A limited number of studies have explored the use of ANN to describe dynamic ultrafiltration with backflow filtration [6–8] and in situations in which feed quality may be variable [3,9]. Previously developed ANN models of RO plant performance were based on the use of training data sets whereby the

* Corresponding author. Tel.: +1 310 825 8766.

0376-7388/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.memsci.2008.10.028

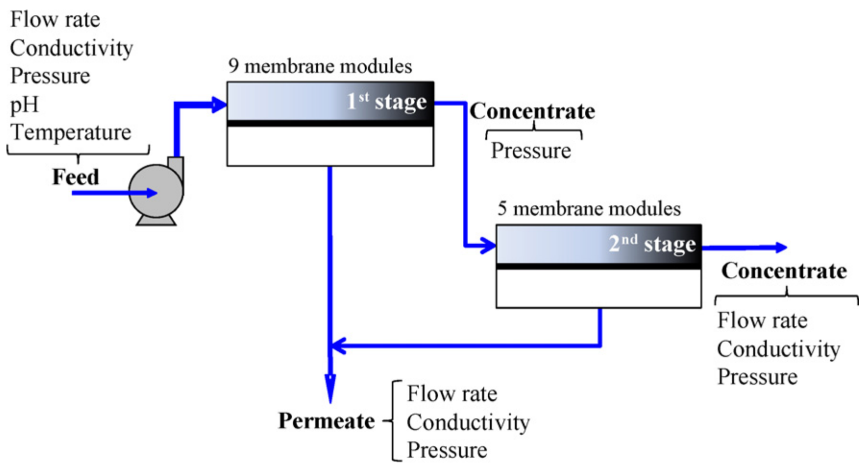


**Fig. 1.** Schematic of the two-stage RO plant indicating the various measured parameters. The first and second stages contained nine and five modules, respectively.

data points for training and testing were interdispersed throughout the complete data time-series. These models were reasonably successful for data interpolation (i.e., predictions for an input variable range for which the ANN model was trained) but lacked forecasting capability (i.e., performance predictions for time periods that were not covered by the training data set). The ability to forecast membrane plant performance, even for short forward time periods, would provide additional flexibility for developing an integrated process control strategy and as an early warning system to signal the need for remedial action (e.g., membrane cleaning, adjustment of process variables such as pressure and flow rates). Arguably, the ANN approach is data-driven and, therefore, results in a plant-specific performance model. However, such an approach would have the advantage of capturing the unique aspects of the plant under consideration, including specific operational behavior of plant equipment (e.g., pumps, valves, monitoring devices and control system), process elements (i.e., membrane modules, feed pretreatment modules), plant configuration, as well as feed quality variations.

Forecasting of process performance can be achieved by taking into account appropriate process input variables as well as a measure of the process memory. Indeed, forecasting over small predictive horizons was demonstrated in financial time-series problems [15] by the common procedure of using a fixed number of previous points (instants of time) of the target variable as model inputs. Neural algorithms and classifiers were also successfully applied in several multidimensional forecasting applications of engineering and fundamental interest concerning the inferential prediction of product quality from process variables [16] and the analysis of turbulent fluid flow phenomena [17], respectively. The application of the latter neural network-based multidimensional methods to the analysis of real RO plant performance is appropriate since plant data are typically noisy and the required predictive horizon is relatively large (e.g., 1 day).

In the present study, we demonstrate the feasibility of constructing ANN models of RO plant performance to describe temporal variations in permeate flux and salt passage with a unique capability for useful short-term forecasting of process performance degradation. Accordingly, the concept of process memory is introduced, whereby past-time (i.e., backward in time) process information of the target variable is utilized in a time-marching approach, along with operational process variables, to forecast plant performance. The present methodology serves to suggest the possibility of using such an ANN modeling approach for RO process control and process fault identification.

## 2. Experimental data and model development

### 2.1. RO pilot plant data

The experimental RO plant data used for building the ANN models were provided by the WaterEye Corporation (Grass Valley, CA). The 1 MGD (million gallons per day) RO brackish water desalination plant was of a 2:1 array configuration located at Port Hueneme, CA (Fig. 1), operated at 75% recovery. The first and second stages contained nine and five membrane modules, respectively. The monitored plant parameters for the feed stream included flow rate, conductivity, feed pressure, pH and temperature. Permeate and concentrate (i.e., brine) monitored parameters included flow rate, conductivity and pressure (including inter-stage pressure). Real time data of the above process parameters (Fig. 1) were collected every $\Delta\tau = 10$ min for a period of about 3 months. Feed salinity (Fig. 2) and operational parameters can vary during plant operation and this typically results in variations in permeate flow rate and salt passage as shown for the present plant data in Fig. 3.

Changes in permeate flow and salt passage can be the result of both variations of operating conditions and membrane fouling. Therefore, in order to effectively evaluate plant performance, it is necessary to compare permeate flow and salt passage rates at a standard reference condition. In the current study, the ASTM

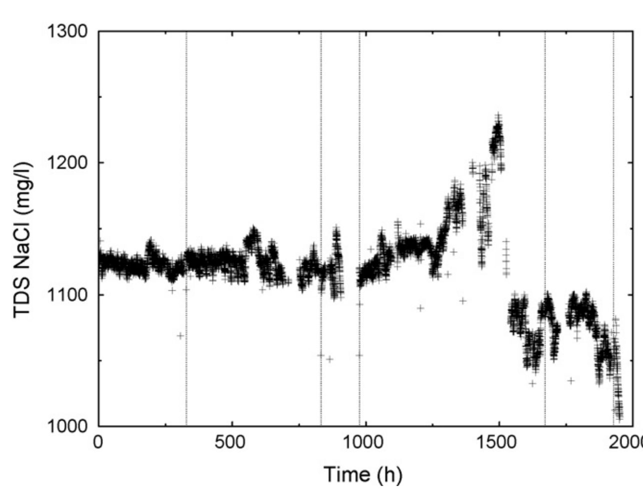


**Fig. 2.** Variability of feed water salinity (expressed as mg/L, TDS) during the RO plant evaluation period.

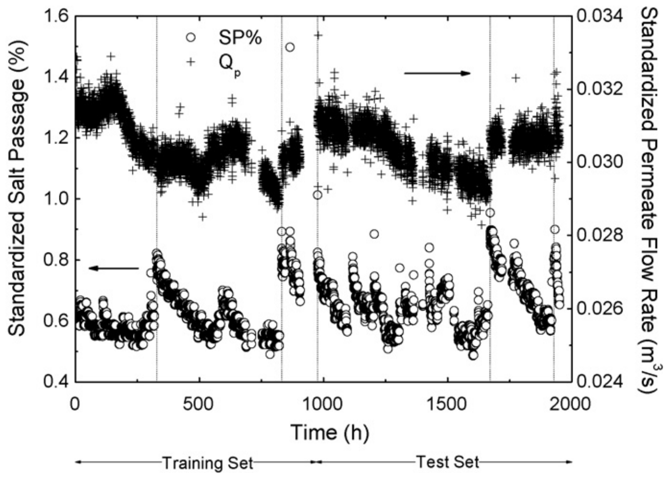


**Fig. 3.** Time-evolution of the standardized salt passage and standardized permeate flow. The vertical dotted lines represent process startup after plant shut down.

4516-00 [18] method was used to standardize the permeate flow rate with respect to temperature, pressure and osmotic pressure (which is related to the solution salinity) at 25 °C, and with respect to the measured process parameters for the first monitoring point for each plant startup period. Briefly, according to the above ASTM procedure, the standardized permeate flow was calculated as

$$Q_{p,s} = \frac{\left[P_{f,s} - \left(\left(P_{f,s} - P_{c,s}\right)/2\right) - P_{p,s} - \pi_{b,s} + \pi_{p,s}\right] \mathrm{TCF}_s}{\left[P_{f,a} - \left(\left(P_{fa} - P_{c,a}\right)/2\right) - P_{p,a} - \pi_{b,a} + \pi_{p,a}\right] \mathrm{TCF}_a} Q_{p,a} \qquad (1)$$

where $Q_p$ is the permeate flow rate, $P_f$, $P_c$ and $P_p$ are the feed, concentrate and permeate pressures (kPa), respectively, $\pi_b$ and $\pi_p$ are the brine and permeate osmotic pressures (kPa), respectively, and where the subscripts 'a' and 's' refer to the actual and the standardized measurement values, respectively. The temperature correction factor, TCF was calculated as TFC = exp[3020 (1/298.15 − 1/$T$)], in which $T$ is the absolute temperature in degrees K [19]. It is noted that, in the preceding analysis, the RO plant performance with respect to the permeate flow was expressed in terms of the permeate flux non-dimensionalized with respect to the initial standardized flux for the performance period under consideration.

The standardized percent salt passage (%SP) for the RO process was calculated as

$$\%\mathrm{SP}_s = \frac{\mathrm{EPF}_a}{\mathrm{EPF}_s}\frac{\mathrm{TCF}_a}{\mathrm{TCF}_s}\frac{C_{b,s}}{C_{b,a}}\frac{C_{f,a}}{C_{f,s}}\%\mathrm{SP}_a \qquad (2)$$

where EPF is the average RO element permeate flow rate, and $C_b$ and $C_f$ are the brine and feed concentrations, respectively, with subscripts 'a' and 's' as defined previously. The brine concentration was expressed as a log-mean average and calculated in terms of the recovery $Y$ (i.e., permeate to feed flow rates ratios), according to $C_b = C_f \ln[(1/(1-Y))/Y]$. In the current analysis, concentrations were all expressed in terms of mg/L of total dissolved solids (TDS) based on the conversion from the measured conductivity to equivalent NaCl concentration in mg/L. The conversion from measured conductivities for the brine and permeate streams was accomplished by constructing conductivity (μS/cm) – TDS (mg/L) correlations derived from multi-electrolyte calculations using the OLI Analyzer software [20]. The correlations were developed based on the average ionic composition of the feed (for the period of 14/03/00–25/03/00; Table 1), by calculating the conductivity that would result from various concentration levels for the feed water and correspondingly produced permeate for different levels of salt passage. Based on the resulting conductivity for the equivalent NaCl

**Table 1**
Average feed composition.

| Ion/Species | mg/L |
|---|---|
| $SiO_2$ | 28.0 |
| $CO_2$ | 8.7 |
| $Na^+$ | 103.5 |
| $K^+$ | 4.6 |
| $Ca^{2+}$ | 142.6 |
| $Mg^{2+}$ | 39.6 |
| $Fe^{3+}$ | 0.1 |
| $Ba^{2+}$ | 0.03 |
| $HCO_3^-$ | 261 |
| $Cl^-$ | 46.2 |
| $F^-$ | 0.4 |
| $SO_4^{2-}$ | 445 |
| $NO_3^-$ | 1.46 |
| $CO_3^{2-}$ | 0.83 |
| TDS (mg/L) | 1082.0 |
| pH | 7.68 |

TDS (mg/L) the following correlations were obtained,

$$\text{Permeate}: \quad \mathrm{TDS_{NaCl}} = 0.346(\mathrm{Cn})^{1.017} \qquad (3a)$$

$$\text{Feed}-\text{brine}: \quad \mathrm{TDS_{NaCl}} = 0.141(\mathrm{Cn})^{1.157} \qquad (3b)$$

in which Cn is the conductivity (μS/cm). The linear correlation coefficients were about unity for both correlations, with average relative errors of $8.1 \times 10^{-3}$ and $1.7 \times 10^{-2}$% for Eqs. (3a) and (3b), respectively.

### 2.2. Data preprocessing and analysis

The standardized percent salt passage and permeate fluxes (Fig. 3) and pressure data (Fig. 4) revealed gaps in data acquisition and missing data as a result of plant shut down periods. Process interruptions (e.g., membrane cleaning and/or replacement) were identified via operational logs and verified by inspecting discontinuities in the time-series data for the process variables. The startup times after interruptions are indicated by the dotted vertical lines (at $t$ = 328, 832, 976, 1671, and 1928 h) in Figs. 3 and 4 for the percent salt passage, standardized permeate flux and pressure, respectively; the plant was shut down for various periods just prior to these times (see Figs. 3 and 4). There were other periods of interruptions in data acquisition as is evident in these two figures; however, the RO plant continued its operation during those

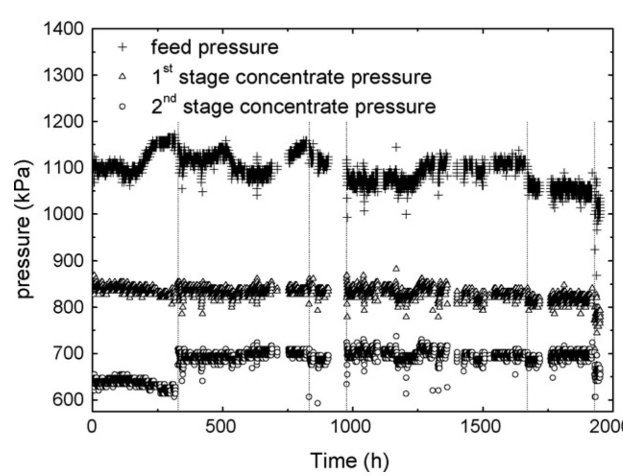


**Fig. 4.** Time-evolution of feed and concentrate pressures in the two RO stages. The vertical dotted lines represent process startup after plant shut downs. In the period studied, the range of permeate pressure variation was 62–117 kPa (9–17 psi). All pressures are given as relative gauge values.

periods. The standardized permeate flux and percent salt passage (Fig. 3) were normalized relative to the corresponding standardized values, $Q_{po}$ and $SP_o$, at the beginning of the overall operational period (14/01/01 in Fig. 3).

Actual state-of-the-plant (ASP) models for the standardized permeate flux and percent salt passage performance at a given time $t$, were first developed based on plant process parameters measured at the same given time for which model predictions were sought. The target dimensionless variables $Q_p(t)/Q_{po}$ and $SP(t)/SP_o$ were predicted from the selected plant parameters measured at the target time (i.e., $Q_p(t)/Q_{po} = f$[Plant parameters(t)] and $SP(t)/SP_o = f$[Plant parameters$(t)$]). It is important to note that the forecasting model does not include direct input information regarding the target variable. For example, permeate and concentrate conductivities were not used as an input variable since it is a process parameter that infers the measure of salt passage. Subsequent to the ASP models, a series of forecasting models were enabled through the inclusion of time-history information of the target variables (e.g., permeate flux and salt passage). The present approach makes use of a simple linear "short-term memory" (STM) time-interval $\Delta t$ that includes past and current instantaneous target variable information to capture plant variations in operating conditions and to enable a forecasting capability in a simple manner and over a sufficiently large time horizon from the plant operation point of view. Specifically, forecasting is achieved with a model that is provided with the selected plant process parameters and target variable values at the present time $t$ and the target variable value at $t - \Delta t$, with forecasting predictions made at future times beyond $t$. The above modeling approach is truly forecasting since the effects of time-dependent plant operating conditions on the target output variables (e.g., permeate flux and salt passage) are predicted at future times.

Three different forecasting approaches were evaluated in the present study. The first was the standard time-series correlation (STSC) method [21] that was investigated for its suitability to forecast the standardized permeate flux and percent salt passage, from their respective values measured at previous time-instants. The second approach was based on sequential forecasting (SF) models in which the target variables were forecasted sequentially at all future times within a predictive interval $\Delta t_p \le \Delta t$ by using the same plant parameters, as in the ASP models, measured at the current time, together with current ($t$) and past ($t - \Delta t$) values of the target variable. The STM interval $\Delta t$ was incorporated into the above model by dividing the overall process operational period into equal $N_{\Delta t_p}$ time-intervals $\Delta t_p$. Each interval, $\Delta t_p$, began at an overall operational time $t_N^o$, where $N = 1,\ldots, N_{\Delta t_p}$, with the model incorporating the target variable and plant information at this time $t_N^o$, as well as past-time target information at $(t_N^o - \Delta t)$. To facilitate model building and sequential forecasting of all time-instances within each $\Delta t_p$, independently of the operational time period being examined, each interval $\Delta t_p$ was in turn divided into $N_{\Delta\tau} = \Delta t_p/\Delta\tau$ data intervals of $\Delta\tau = 10$ min, in accordance with the frequency at which plant data were recorded. Time ($\tau$) information was also provided into the model after resetting the time counter to zero ($\tau_O = 0$) at the beginning of each interval $\Delta t_p$ ($t = t_N^o; \tau = \tau_o = 0$). Thus, the actual process time within each interval $\Delta t_p$, $t_N^0 < t \le (t_N^0 + \Delta t)$, was transformed into a time domain within the forecasting interval (i.e., $0 < \tau \le \Delta t_p$). Accordingly, the SF models were of the form,

$$\frac{Q_p\left(t_N^o + \tau_j\right)}{Q_{po}} = f\left[\frac{Q_p\left(t_N^o - \Delta t\right)}{Q_{po}}; \frac{Q_p\left(t_N^o\right)}{Q_{po}}; \text{Plant parameters}\left(t_N^o\right); \tau_j\right];$$
$$j = 1, \ldots, N_{\Delta\tau};\ N = 1, \ldots, N_{\Delta t_p} \tag{4a}$$

$$\frac{SP\left(t_N^o + \tau_j\right)}{SP_o} = f\left[\frac{SP\left(t_N^o - \Delta t\right)}{SP_o}; \frac{SP\left(t_N^o\right)}{SP_o}; \text{Plant parameters}\left(t_N^o\right); \tau_j\right];$$
$$j = 1, \ldots, N_{\Delta\tau}; N = 1, \ldots, N_{\Delta t_p} \tag{4b}$$

with $\tau_j = j\Delta\tau$, $j = 1,\ldots, N_{\Delta\tau}$, included as input variable in appropriate time-units. It is noted that the first time-derivatives of the target variables, at the beginning of operational periods after maintenance, were taken equal to zero (i.e., $Q_p(t_N^o - \Delta t)/Q_{po} = Q_p(t_N^o)/Q_{po}$; $SP(t_N^o - \Delta t)/SP_o = SP(t_N^o)/SP_o$) since each represented a new operational period without immediate past history. Only predictive intervals smaller than the STM interval (i.e., $\Delta t_p \le \Delta t$) were explored in the SF models given the linear nature of the memory term and the inclusion of time ($\tau_j$) as an input variable (Eqs. (4a) and (4b)).

The third forecasting modeling approach was based on a marching forecast (MF). In the MF models, the target variables are continuously predicted at future times $(t + \delta)$ according to,

$$\frac{Q_p\left(t + \delta\right)}{Q_{po}} = f\left[\frac{Q_p(t - \Delta t)}{Q_{po}}; \frac{Q_p(t)}{Q_{po}}; \text{Plant parameters}(t)\right] \tag{5a}$$

$$\frac{SP\left(t + \delta\right)}{SP_o} = g\left[\frac{SP(t - \Delta t)}{SP_o}; \frac{SP(t)}{SP_o}; \text{Plant parameters}(t)\right] \tag{5b}$$

where the time-derivatives for the target variable were also set to be zero at the beginning of each operational period commencing after maintenance, as in the case of the SF approach. In building the SF and MF models, STM intervals ($\Delta t$ value) providing past information, and forecasting times $\Delta t_p$ and $\delta$ (Eqs. (4) and (5)), were assessed over the ranges of $2\,\text{h} \le \Delta t \le 48\,\text{h}$, $2\,\text{h} \le \Delta t_p \le 24\,\text{h}$ and $2\,\text{h} \le \delta \le 24\,\text{h}$.

It is important to note that, the impact of membrane fouling and/or scaling on membrane performance is a cumulative effect over time with flux decline becoming more severe with fouling progression. However, the time scales associated with plant readjustment to changes in operational conditions (e.g., pressure and flow rate changes) can be much shorter than the fouling time scale, especially relative to flux decline of less than about 5–10% (typically marking the recommended level that triggers the need for membrane cleaning). Therefore, plant process time scales are expected to be unique to the configuration and hardware of the plant, the process set points and quality of the feed. In the present case study, permeate flux and salt passage fluctuations and temporal pattern (Fig. 3) were of durations of the order of hours and days. Therefore, it is desirable for forecasting time-intervals to be at least within the range of the expected time scales for the plant.

### 2.3. Algorithms

ANN models were built, separately, for the standardized permeate flux and percent salt passage (i.e., target variables) using both a back-propagation (BP) algorithm [22] and support vector machines (SVM) [23–25] to establish the relationships between the selected model inputs and the target variables (permeate flux and salt passage). Details on the subject of ANN model building by BP and SVM can be found elsewhere [22–25]. Briefly, the current BP architectures consisted of one input layer (with the number of inputs required by each model tested), one output layer (one output target variable) and a hidden layer in which a different number of neurons were used for different models to evaluate the performance of various model architectures. The linear transfer function was utilized for the input and output layers and a hyperbolic tangent transfer function was used for the hidden layer [26]. A Levenberg–Marquardt technique [27,28] was used during the learning phase (training and validation) for adjusting the weights by backward propagation [26] of the error between the ANN output of an input pattern and the corresponding target process variable (i.e., percent salt passage or permeate flux). The SVM's are kernel-based supervised learning methods that can be applied to classification and regression [23–25]. SVM methods were originally developed to classify linearly separable datasets by means of a hyperplane defined by support vectors. In non-linear classification problems the coordinates of the input data are mapped into a high-dimensional feature space where the hyperplane is computed with non-linear kernel functions [23–25]. Kernels operate in the input space where classification can be readily attained as the weighted sum of kernel functions evaluated at the support vectors. In the present approach, SVM were applied to develop support vector regression (SVR) models [23] utilizing linear (dot) kernels.

The neural network and kernel-based models were trained using data from the first three periods of operation (0–906 h) which contained a total of 4569 data points (Fig. 3) representing about 50% of the overall dataset. Internal model validation was carried out with about 20% of the data set for the first 906 h of plant operation (selected randomly; see Figs. 3 and 4). The internal validation step served as a stopping criterion for the ANN learning algorithm. Model testing was performed by using the data in the last three operational periods, starting at $t = 976$ h and forward in time (Figs. 3 and 4), as an external validation data set. These data were not used in the learning phase, neither for model training, nor for internal validation; in other words, model testing was accomplished with an external data set. Thus, model evaluation was carried out in a marching time-series approach that avoided interpolation of test data within the training set. Overall, approximately 40% of the entire plant data set was used for ANN model training, 10% of the data were used for internal validation within the test set, and the remaining 50% of the time-series sequence (i.e., the last three operational periods in Figs. 3 and 4) were used for external model testing.

The adequacy of both the ANN architecture and the "short-term memory" time-intervals for the various models were assessed using a model quality index, $q^2$, calculated separately for the training (tr) and test (te) data sets [29] according to,

$$q^2_{\mathrm{tr}} = 1 - \frac{\sum_{i=1}^{n_{\mathrm{tr}}} \left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n_{\mathrm{tr}}} (y_i - \bar{y}_{\mathrm{tr}})^2} \qquad (6)$$

$$q^2_{\mathrm{te}} = 1 - \frac{\sum_{i=1}^{n_{\mathrm{te}}} \left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n_{\mathrm{te}}} (y_i - \bar{y}_{\mathrm{tr}})^2} \qquad (7)$$

in which $y_i$, $\hat{y}_i$ are respectively the experimental and predicted values of the dependent variable while $\bar{y}_{\mathrm{tr}}$ is the averaged values of the experimental dependent variable for the training set. The quality index, $q^2$, varies between 0 and 1, with unity being the maximum attainable quality. Note that $q^2_{\mathrm{te}}$ in Eqs. (6) and (7) has been normalized with respect to the averaged value of the experimental dependent variable for the training set, $\bar{y}_{\mathrm{tr}}$, to penalize (decrease) the quality of the test set when over-fitting occurs during the training phase. It is noted that models with $q^2 > 0.5$ are considered to have predictive capabilities [30–32]. In addition, model performance was also evaluated by means of the mean absolute error (MAE), percent error, and linear correlation coefficient.

## 3. Results and discussion

### *3.1. Actual state-of-the-plant models*

ASP models (i.e., $\Delta t = 0$) were first developed to assess the consistency of training and testing data sets, and the influence of process variables and algorithms in model building. Several BP architectures, with 2–11 neurons in the hidden layers, together with SVR models, were evaluated for modeling the permeate flux and salt passage. Feature selection analyses and expert criteria on the performance of reverse osmosis membrane desalting revealed that the most suitable set of input variables for the target variables (permeate flow and salt passage), for the present plant, were the feed flow rate, feed conductivity, overall pressure drop, pressure drop across stage 2, and feed pH,

$$\frac{Q_{\mathrm{p}}(t)}{Q_{\mathrm{po}}} = f\left[Q_{\mathrm{feed}}(t),\ \ \mathrm{Cn}_{\mathrm{feed}}(t),\ \ \mathrm{pH}_{\mathrm{feed}}(t),\ \ \Delta p_{\mathrm{overall}}(t),\ \ \Delta p_{\mathrm{stage\,2}}(t)\right] \qquad (8)$$

$$\frac{\mathrm{SP}(t)}{\mathrm{SP_o}} = f\left[Q_{\mathrm{feed}}(t),\ \ \mathrm{Cn}_{\mathrm{feed}}(t),\ \ \mathrm{pH}_{\mathrm{feed}}(t),\ \ \Delta p_{\mathrm{overall}}(t),\ \ \Delta p_{\mathrm{stage\,2}}(t)\right] \qquad (9)$$

It is noted that the feed, concentrate and permeate pressures were not explicitly incorporated into the ASP and forecasting models since these variables are considered in the normalization of the permeate flux (Eq. (1)). Although the different process pressures are included in the standardized permeate flux (Eq. (1)), the approximate ASTM 4516-00 accepted averaging of the transmembrane pressure driving force [18] does not fully account for the non-linear relation between the pressure and osmotic pressure (at the membrane surface) along the membrane modules and the observed permeate flux. Therefore, the use of the pressure drop across the 2nd RO stage and the overall pressure drop provides additional important information not captured in the permeate flux standardization procedure. It is also noted that it is also possible to use the pressure drops across the 1st and 2nd RO stages; however, the choice of a pressure drop across one stage and the overall pressure drop provided a greater difference in magnitude between the two pressure variables. Temperature was not included as a model input variable for the ASP and forecasting models since temperature was included in the standardized calculations for permeate flux and salt passage target variables. Although in the present data pH variation was only in the range of 6.4–7.6, it was included as an input variable since pH is known to affect mineral scaling as well as colloidal and biofouling.

The performance of the ASP models is summarized in Table 2 in terms MAE and the quality index $q^2_{\mathrm{te}}$ obtained for the test data sets of $Q_{\mathrm{p}}/Q_{\mathrm{po}}$ and $\mathrm{SP}/\mathrm{SP_o}$ with the optimal BP 5:3:1 and SVR with a radial basis function kernel ($C = 1$; $\varepsilon = 1 \times 10^{-4}$; $\sigma = 1$) architectures. Both ANN models are consistent in terms of architectures since the small number of neurons in the hidden layer of the BP algorithm (5:3:1) is matched by the relatively small number (1069 vectors, 23% of training data) of support vectors needed to build the SVR-based model. Table 2 indicates that while the performances of both BP and SVR models are reasonable for the normalized permeate flux ($\bar{\varepsilon} \approx 0.9\%$; $q^2_{\mathrm{te}} \approx 0.6$), with the latter yielding slightly better predictions, both BP and SVR-based models were unable to satisfactorily predict the salt passage ($\bar{\varepsilon} \geq 9.2\%$; $q^2_{\mathrm{te}} \approx 0.0$).

The time-sequences of the normalized permeate flux and salt passage values, predicted with the ASP models for the respective test data sets, are shown in Figs. 5 and 6, respectively. It is clear from these figures and Table 2 that, while performances of ASP models for permeate flow are reasonable ($q^2_{\mathrm{te}} \approx 0.6$) and with consistent trends, in the case of $\mathrm{SP}/\mathrm{SP_o}$, both BP and SVR-based models tend to predict a nearly constant salt passage value close to the average of the complete test set. These salt passage results indicate that the time-variation of the five selected plant parameters (feed flow rate, feed conductivity, feed pH, overall pressure drop, 2nd stage pressure drop), is insufficient to capture plant performance at the same current time without using past salt passage information.

Please cite this article in press as: D. Libotean, et al., Neural network approach for modeling the performance of reverse osmosis membrane desalting, J. Membr. Sci. (2008), doi:10.1016/j.memsci.2008.10.028

**Table 2**
Summary of the performance of the ASP and SF models for the prediction of permeate flux and salt passage for the test data sets with the best BP and SVR architectures. The SF model results are reported for three different STM values ($\Delta t = 8$, 12 and 24 h) and several forecasting intervals ($\Delta t_p$).

| MODELS | Time (h) | | Actual state | Sequential forecasting (SF) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\Delta t = 0$ | $\Delta t = 8$ | | | $\Delta t = 12$ | | | | $\Delta t = 24$ | | | | |
| | | | | $\Delta t_p = 2$ | $\Delta t_p = 4$ | $\Delta t_p = 8$ | $\Delta t_p = 2$ | $\Delta t_p = 4$ | $\Delta t_p = 6$ | $\Delta t_p = 12$ | $\Delta t_p = 2$ | $\Delta t_p = 4$ | $\Delta t_p = 8$ | $\Delta t_p = 12$ | $\Delta t_p = 24$ |
| Permeate flux ($Q_p/Q_{po}$) | 5:3:1 | $\bar{\varepsilon}$ | 0.93 | – | – | – | – | – | – | – | – | – | – | – | – |
| | | $q^2_{te}$ | 0.56 | – | – | – | – | – | – | – | – | – | – | – | – |
| | 8:3:1 | $\bar{\varepsilon}$ | – | 0.83 | 0.89 | 0.92 | 0.82 | 0.82 | 0.87 | 0.89 | 0.82 | 0.87 | 0.93 | 1.00 | 0.99 |
| | | $q^2_{te}$ | – | 0.68 | 0.64 | 0.61 | 0.69 | 0.70 | 0.64 | 0.64 | 0.64 | 0.61 | 0.55 | 0.47 | 0.49 |
| | SVR | $\bar{\varepsilon}$ | 0.91 | 0.83 | 0.87 | 0.91 | 0.81 | 0.83 | 0.85 | 0.91 | 0.88 | 0.89 | 0.90 | 1.05 | 1.18 |
| | | $q^2_{te}$ | 0.61 | 0.68 | 0.65 | 0.62 | 0.70 | 0.68 | 0.67 | 0.62 | 0.59 | 0.59 | 0.58 | 0.42 | 0.31 |
| Salt passage ($SP/SP_o$) | 5:3:1 | $\bar{\varepsilon}$ | 13.05 | – | – | – | – | – | – | – | – | – | – | – | – |
| | | $q^2_{te}$ | −0.40 | – | – | – | – | – | – | – | – | – | – | – | – |
| | 8:3:1 | $\bar{\varepsilon}$ | – | 4.13 | 4.67 | 5.92 | 3.80 | 5.13 | 4.67 | 5.70 | 3.80 | 4.16 | 5.62 | 5.78 | 7.97 |
| | | $q^2_{te}$ | – | 0.83 | 0.80 | 0.66 | 0.84 | 0.77 | 0.79 | 0.69 | 0.88 | 0.86 | 0.76 | 0.73 | 0.53 |
| | SVR | $\bar{\varepsilon}$ | 9.18 | 4.39 | 4.42 | 5.71 | 4.21 | 4.60 | 4.89 | 5.63 | 3.96 | 3.89 | 4.65 | 4.96 | 5.76 |
| | | $q^2_{te}$ | 0.02 | 0.81 | 0.80 | 0.64 | 0.82 | 0.79 | 0.77 | 0.65 | 0.85 | 0.85 | 0.81 | 0.78 | 0.71 |

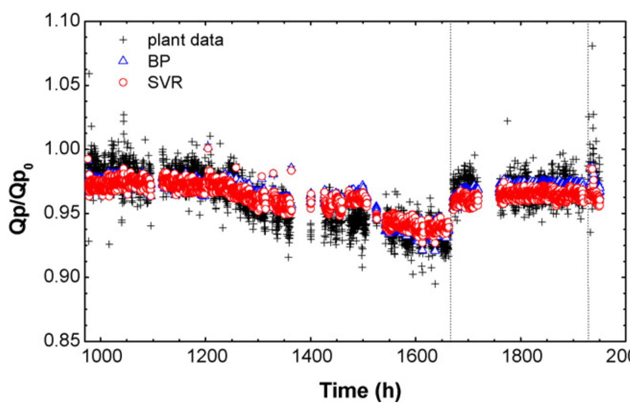


**Fig. 5.** Comparison between experimental permeate fluxes and those predicted with a BP (5:3:1 architecture) and SVR state-of-the-plant (ASP) models ($\Delta t = 0$) over the complete test data set (see the training and testing periods in Fig. 3).

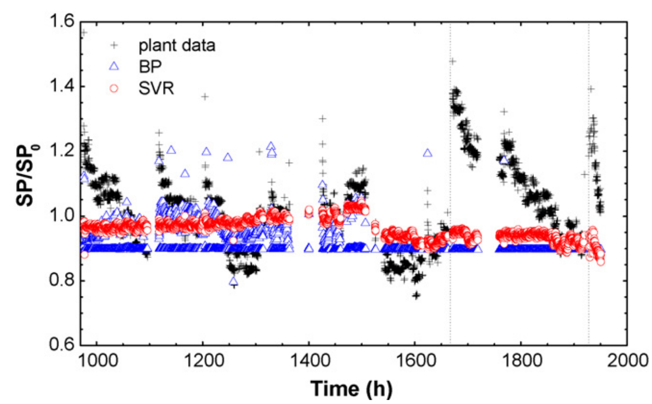


**Fig. 6.** Comparison between experimental salt passage values and those predicted with a BP (5:3:1 architecture) and SVR ASP models ($\Delta t = 0$) over the complete test data set (see the training and testing periods in Fig. 3).

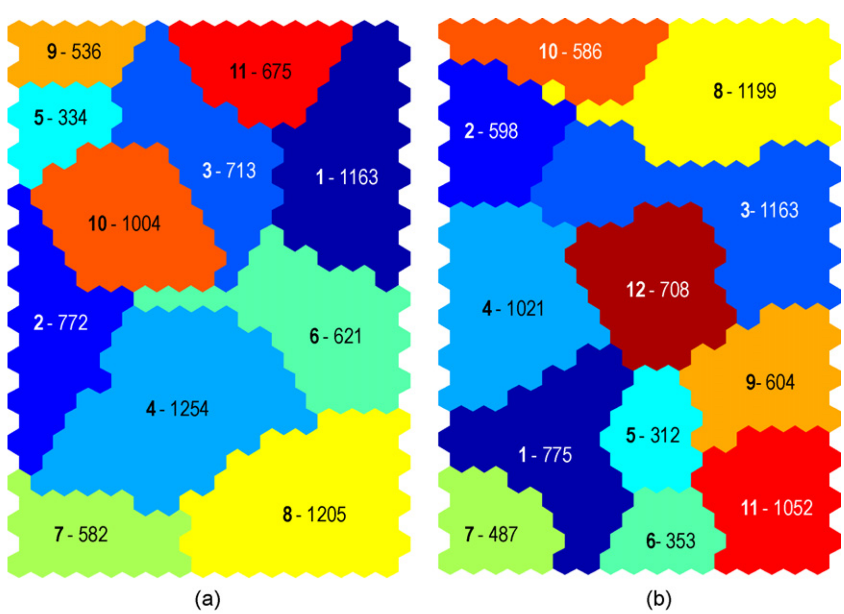


**Fig. 7.** SOM clustering of RO plant operation patterns (training and test) into families: (a) normalized permeate flux and (b) percent salt passage. The first number in each cluster is the family ID number and the second the number of family member data points (family population).

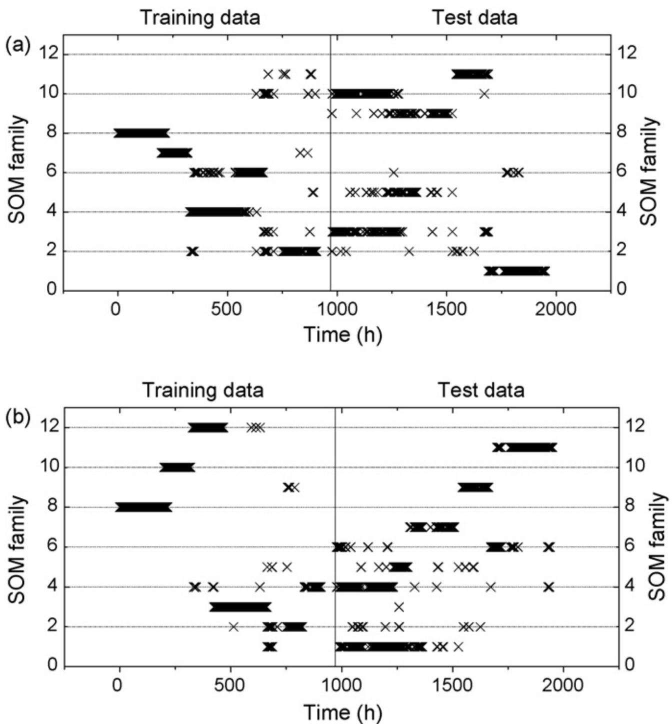


**Fig. 8.** Representation of the families of operational patterns with respect to: (a) permeate flux and (b) percent salt passage.

The fact that ASP models without past-time information ($\Delta t = 0$) are able to capture, in part, features of the time-evolution of the normalized permeate flux but not of the salt passage (Figs. 5 and 6 and Table 2), can be best understood if major RO plant operational data families are identified by clustering similar operation conditions with a Self-Organizing Map (SOM) algorithm [33,34]. In this analysis method, a 27 × 17 cell SOM was used to classify the complete (training plus testing) operational data time-series (Fig. 3) using the five selected input process variables (feed flow rate, feed conductivity, feed pH, overall pressure drop, 2nd stage pressure drop) together with either the normalized permeate flux or salt passage targets. Accordingly, the SOM units or classes were clustered into similar families of operating conditions by means of the Davies–Bouldin index [35]. This clustering process resulted in 11 separate families of operational conditions for the permeate flux and 12 families for the salt passage, as shown respectively in Fig. 7a and b. The corresponding training and test data set distributions into these SOM families is depicted in Fig. 8a and b for the permeate flux and salt passage, respectively. The vertical black line at 976 h in these two figures denotes the separation between the training set data (operational times from 0 to 906 h) and the test set (operational times from 976 to 1950 h).

As illustrated in Fig. 8a, all permeate flux data in the test set (right side of the vertical line) were well represented in the training set families (left side), except for families 1 and 9 which, while being populated in Fig. 7a, were nonetheless observed to share similarities between themselves and with the neighboring SOM classes in Fig. 7a. This explains the reasonable performance of the ASP permeate flux models, as shown in Fig. 5 and Table 2, for the test set of this target variable. On the other hand Fig. 8b shows that the distribution of training and test salt passage data into the SOM families (Fig. 7b) is very unbalanced since test data (right side of Fig. 8b) in families 1, 4, 5, 6, 7, 9 and 11 are poorly or not at all represented in the training data set (left side of Fig. 8b). This training-test data mismatch is also consistent with the poor ASP salt passage models depicted in Fig. 6 and as reported in Table 2. In other words, the characteristics of the test set were not contained in the training set and thus could not be learned during training.

It is clear that, in the absence of past plant information as model input, the lack of a sufficient representation of the test data families in the training set limits the model applicability to the operational families that are well represented in the training time-series. One can resolve the above problem, in part, by dividing the data into training and test sets [36] for the entire set of operating conditions, such that data points in the training set cover the same range of operational conditions as in the test set (i.e., achieved by having the data set points interdispersed throughout the complete data set). However, such an approach would result in an ANN model that is useful for interpolation but not for forecasting RO plant performance (Section 3.2). Even periodic retraining of the ANN model may not assure an acceptable level of predictive accuracy. On the other hand, the inclusion of past target variable information as an additional model input (i.e., target value at $t - \Delta t$) does provide a measure of forecasting horizon. When the above SOM analyses (Figs. 7 and 8) were repeated by adding this target variable information at $t - \Delta t$ all families of operational patterns, for both the permeate flux and salt passage, shared members from both the training and test sets. The above results suggest that forecasting models of plant performance should include the time $\tau$ and/or STM information as discussed Section 2.2 and amplified Section 3.2.

### 3.2. Forecasting models

Evaluations of the applicability of the STSC approach demonstrated that model forecasting was limited to a period of about 30 min and required six and seven consecutive time-instants of previous permeate flux and salt passage data, respectively. These STSC models captured plant performance fluctuations over short

**Table 3**
Summary of the performance of the marching forecast (MF) models for the prediction of the permeate flux and salt passage for the test data sets with the best BP and SVR architectures for three different STM values ($\Delta t = 8$, 12 and 24 h) and five different predictive times ($\delta = 2$, 4, 8, 12, and 24 h).

| Models | Time (h) | | Marching forecasting (MF) | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\Delta t = 8$ h | | | | | $\Delta t = 12$ h | | | | | $\Delta t = 24$ h | | | | |
| | | | $\delta = 2$ | $\delta = 4$ | $\delta = 8$ | $\delta = 12$ | $\delta = 24$ | $\delta = 2$ | $\delta = 4$ | $\delta = 8$ | $\delta = 12$ | $\delta = 24$ | $\delta = 2$ | $\delta = 4$ | $\delta = 8$ | $\delta = 12$ | $\delta = 24$ |
| Permeate flux ($Q_p/Q_{po}$) | 7:3:1 | $\bar{\varepsilon}$ | 0.79 | 0.80 | 1.08 | 1.13 | 0.89 | 0.80 | 0.80 | 1.14 | 1.03 | 1.06 | 0.78 | 0.80 | 1.01 | 1.11 | 1.21 |
| | | $q^2_{te}$ | 0.70 | 0.70 | 0.50 | 0.47 | 0.62 | 0.70 | 0.70 | 0.46 | 0.56 | 0.50 | 0.71 | 0.70 | 0.57 | 0.46 | 0.34 |
| | SVR | $\bar{\varepsilon}$ | 0.82 | 0.89 | 1.14 | 1.13 | 1.13 | 0.82 | 0.91 | 1.15 | 1.14 | 1.04 | 0.81 | 0.86 | 1.04 | 1.04 | 1.02 |
| | | $q^2_{te}$ | 0.69 | 0.65 | 0.45 | 0.43 | 0.40 | 0.68 | 0.63 | 0.45 | 0.47 | 0.49 | 0.69 | 0.67 | 0.54 | 0.54 | 0.52 |
| Salt passage ($SP/SP_o$) | 7:3:1 | $\bar{\varepsilon}$ | 3.10 | 3.65 | 4.31 | 4.46 | 5.87 | 3.23 | 3.94 | 4.49 | 4.78 | 6.60 | 2.89 | 3.27 | 4.82 | 5.16 | 5.86 |
| | | $q^2_{te}$ | 0.90 | 0.87 | 0.81 | 0.81 | 0.70 | 0.90 | 0.85 | 0.82 | 0.77 | 0.57 | 0.92 | 0.90 | 0.77 | 0.74 | 0.66 |
| | SVR | $\bar{\varepsilon}$ | 3.40 | 3.79 | 4.48 | 4.53 | 4.93 | 3.52 | 3.84 | 4.42 | 4.41 | 4.88 | 2.89 | 3.91 | 4.52 | 4.58 | 4.87 |
| | | $q^2_{te}$ | 0.89 | 0.87 | 0.81 | 0.81 | 0.80 | 0.88 | 0.86 | 0.82 | 0.82 | 0.79 | 0.87 | 0.85 | 0.82 | 0.81 | 0.78 |

periods but could not describe longer-term behavior (Fig. 3). Therefore, the STSC modeling approach was deemed unsuitable for the development of practical plant diagnostics tools and early warning systems.

The performance of SF and MF models for the normalized permeate flux and salt passage, for different STM intervals and forecasting time-intervals (i.e., $\Delta t_p$ or $\delta$), is summarized in Table 2 and Table 3. It is apparent that the behavior of the models for each target variable is self-consistent with the following overall trends: (i) the performance of all SF and MF models decreases when the forecasting horizon $\Delta t_p$ or $\delta$ increases, except for some at the longest predictive times close to 24 h where the observed increase in performance is due to the decrease in the number of test data points that was imposed by the continuity of data required during training and testing in both forecasting models; (ii) the STM interval that yielded best performance, with all forecasting models and algorithms, for the permeate flux and salt passage was in the range $\Delta t = 8$–$24$ h; (iii) SF models performed better for permeate flow, particularly for large forecasting intervals ($\Delta t_p \approx \Delta t$), while MF models perform better for salt passage for all forecasting time-intervals ($0 < \delta \leq \Delta t$). It should be noted that in the SF models the same STM information of the target variable was used for all predictions at forecasting times ($\tau$) up to $\Delta t_p$, which facilitated learning when time-variations are small. On the other hand, for the MF models the STM information was updated at each time, capturing better future data trends, particularly for large time-variations of the target variables. In other words, all elements in the input training vectors change for each new time enabling better training or learning; (iv) the MF models yielded reasonable performance for forecasting intervals ($\delta$) greater than the STM interval. The results demonstrated that a forecasting period of 1 day (forward in time) is feasible ($q^2_{te} > 0.5$) with SVR-

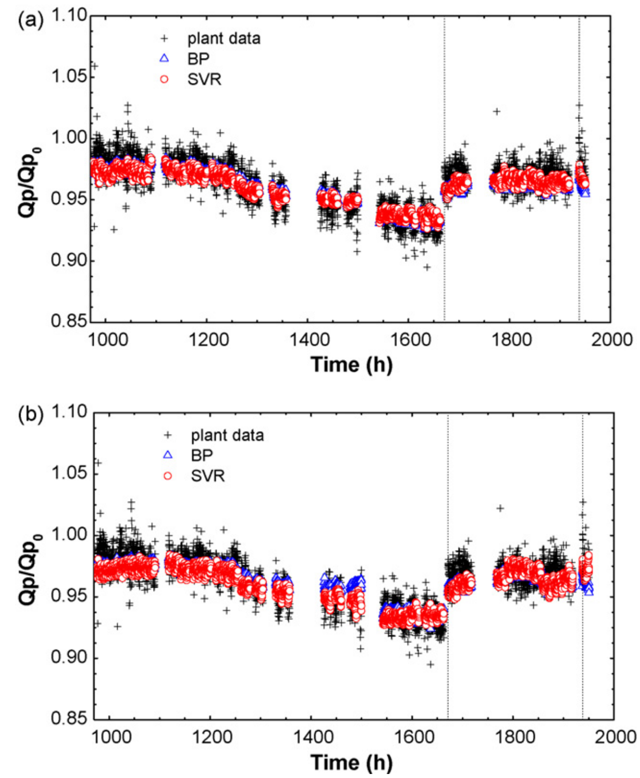


**Fig. 9.** Comparison between experimental $Q_p/Q_{po}$ values and those predicted with BP (8:3:1 architecture) and SVR-based sequential forecasting (SF) models for the test set with $\Delta t = 12$ h. (a) $\Delta t_p = 2$ h and (b) $\Delta t_p = 12$ h.

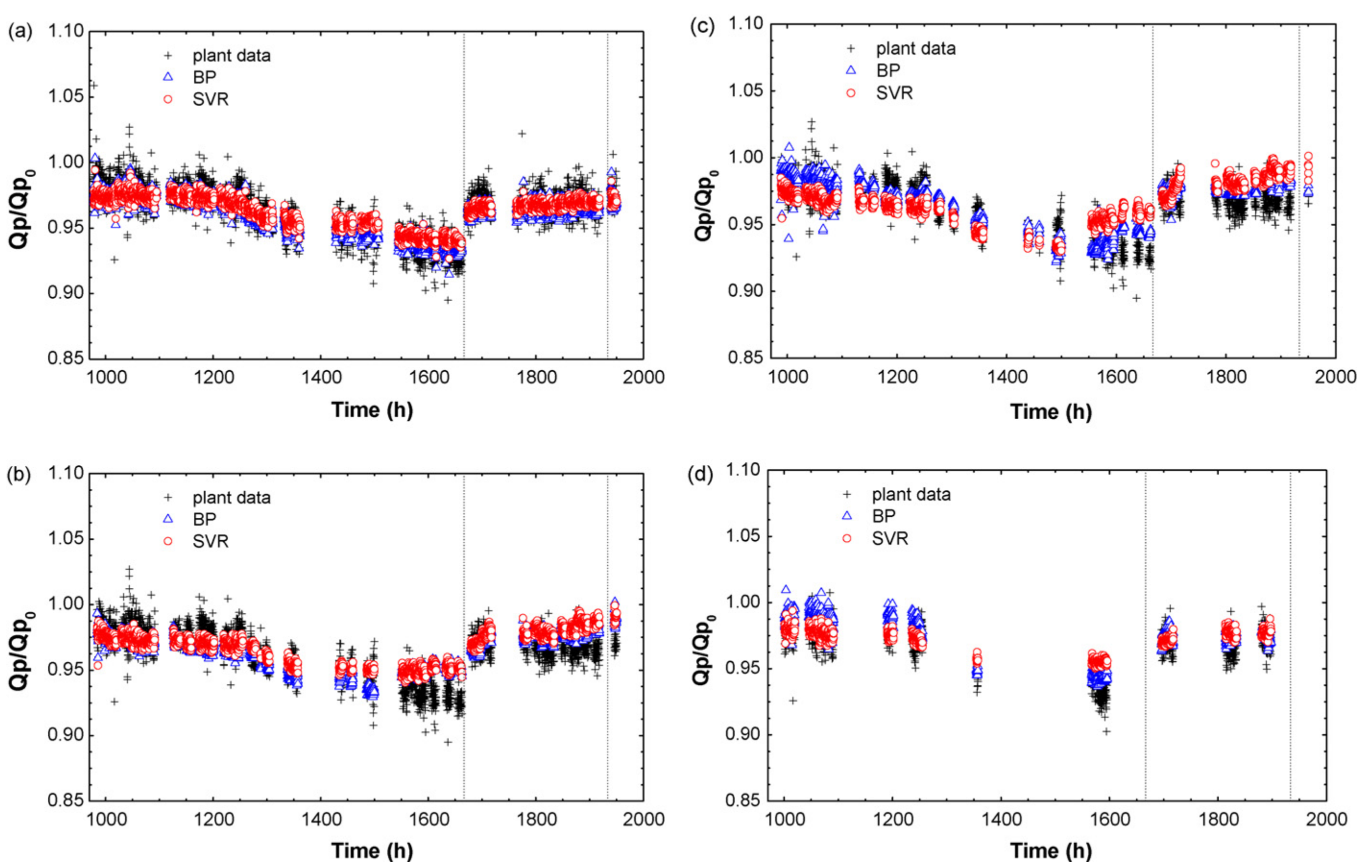


**Fig. 10.** Comparison between experimental $Q_p/Q_{po}$ values and those predicted with BP (7:3:1 architecture) and SVR-based marching forecast (MF) models for the test set with $\Delta t = 12$ h. (a) $\delta = 2$ h; (b) $\delta = 8$ h; (c) $\delta = 12$ h and (d) $\delta = 24$ h.

based permeate flux and salt passage MF models for STM intervals in the range of $\Delta t$ = 8–24 h; (v) both BP and SVR algorithms performed equally well for short forecasting time-intervals. In contrast, the SVR derived models outperformed the BP derived models for the longest forecasting time-intervals (i.e., for $\Delta t_p$ or $\delta \approx 24$ h) and time-variability of the data as was the case, for example, for the salt passage.

SF and MF permeate flux models (Eqs. (4a) and (5a), respectively) were derived using BP neural networks (NN) with optimal NN architectures of 8:3:1 and 7:3:1, respectively. Eight input nodes were utilized for the SF model to predict permeate flux in accordance with Eq. (4a) (i.e., the selected five plant parameters, values of the permeate flux at times $t$ and $t - \Delta t$, and time $\tau$ within the predictive interval $\Delta t_p$). The dimension of the input vector for the MF models was reduced to seven since time $\tau$ was excluded. The same input vectors were used in the corresponding SVR-based SF and MF models. For both the BP and SVR derived permeate flux models, the best performances for SF models (Table 2) were obtained with STM intervals of $\Delta t$ = 8 h and $\Delta t$ = 12 h and forecasting intervals of $\Delta t_p$ = 2 h and $\Delta t_p$ = 4 h. In contrast, performance of the MF models (Table 3), derived using both BP and SVR algorithms, all STM intervals considered resulted in similar performance for short forecasting intervals of $\delta$ = 2 h and $\delta$ = 4 h. Best model performances were obtained with the shortest STM intervals and forecasting intervals with a tendency for decreasing performance for a given STM interval ($\Delta t$) with increasing $\Delta t_p$ or $\delta$ and for increasing $\Delta t$ for a given $\Delta t_p$ or $\delta$. A survey of performance of all the forecasting models developed suggests that a STM interval $\Delta t$ = 12 h is a reasonable practical choice in terms of performance model and robustness.

The SF models developed with an 8:3:1 BP architecture (Table 2) predicted the normalized standardized permeate flux with mean absolute errors $\bar{\varepsilon}$ in the range [0.82–0.89%] and quality indices $q_{te}^2$ in the interval [0.64–0.70] over the test set, for $\Delta t$ = 8 h and $\Delta t$ = 12 h for the range of forecasting intervals considered ($\Delta t_p$ = 2–24 h). The quality indices for the models decreased to $q_{te}^2 = 0.61$ for $\Delta t$ = 24 h and $\Delta t_p$ = 4 h. Similar performance was obtained for the SVR derived SF models, with mean absolute errors increasing to at most 0.009 and 0.012 permeate flux units for $\Delta t_p$ = 2 h and 24 h, respectively, with $q_{te}^2$ decreasing to 0.59 and 0.31 for the same forecasting time-intervals (Table 2). The results show that in all cases, model performance degrades as the forecasting interval $\Delta t_p$ approaches the size of the STM interval. This decrease in model accuracy can be attributed to the smoothing effect of the STM term in Eq. (4b) which masks the temporal permeate flux fluctuations. This effect is also observed for the largest STM interval of 24 h for which the BP and SVR derived forecasting models also show similar performance. As an illustration of the model performance, the experimental permeate flux and the values forecasted with the BP and SVR-based SF models, for the STM interval of $\Delta t$ = 12 h and two forecasting intervals of $\Delta t_p$ = 2 and 12 h are shown in Fig. 9 for the complete test set. It is clear that the forecasted $Q_p(t)/Q_{po}$ values follow the trend of experimental data.

The performances of both the BP and SVR-based MF models (Table 3) were similar to the corresponding SF models for short and moderate forecasting times ($\delta$ = 2–8 h). However, the MF models did not require the additional input of time-sequence information ($\tau$) as in the case of the SF models (Section 2.2). The performance of the BP-based MF models was slightly inferior relative to the corresponding SVR-based MF models. The accuracy of the BP/MF models, in terms of the mean absolute error varied from approximately $\bar{\varepsilon} \approx 0.79\%$ to $\bar{\varepsilon} \approx 1.2\%$ as the forecasting time, $\delta$, was increased from 2 h to 24 h for $\Delta t \leq 24$ h. The error range of $\bar{\varepsilon} \approx 0.82 - 1.15\%$, were obtained with the SVR/MF models over the same forecasting time range $\delta$ (Table 3). For the BP and SVR-derived MF models, the quality of the fit was in the range of $q_{te}^2 = 0.34$–$0.71$ and $q_{te}^2 = 0.40$–$0.69$, respectively, for $\Delta t \leq 24$ h and $\delta \leq 24$ h. The BP and SVR-derived MF models are compared in Fig. 10, for $\Delta t$ = 12 h and $\delta$ = 2, 8, 12 and 24 h. The performance of these models slightly degrades as the forecasting interval approaches or exceeds the value of the STM interval. The quality index for longer term forecasting ($\delta$ = 24 h) using STM interval of $\Delta t$ = 12 h are close to $q_{te}^2 = 0.50$ for both the BP and the SVR models. Table 3 shows that the performance of the SVR-based MF model for the permeate flux is maintained above 0.5 when the STM interval is increased to $\Delta t$ = 24 h while that of the BP-based models drops to $q_{te}^2 = 0.34$. It is important to recognize that both the SF and MF forecasting models for permeate flux (Tables 2 and 3) outperform the ASP model (Table 2). These forecasting models demonstrated that the mean absolute errors decrease and the quality of the fit indices increase when past target information is included as input information, regardless of the algorithm used, even for values of $\delta$ close to the STM interval.

Compared to the RO plant performance with respect to permeate flux, forecasting of salt passage should provide a better indication of the adequacy of the forecasting approach in describing plant performance since the data revealed greater temporal salt passage variations (Figs. 3 and 6). Indeed, the first noticeable trend in the performances of salt passage models, summarized in Tables 2 and 3, is that salt passage forecasting models developed using both the BP and SVR algorithms, which yielded the similar accuracy and goodness of the fit (with the SVR-based models better capturing long-term effects) performed better ($q_{te}^2 \approx 0.53$–$0.92$) than the permeate flux models ($q_{te}^2 \approx 0.31$–$0.71$). Mean absolute errors for salt passage were larger ($\bar{\varepsilon} \approx 4\%$) because the span of data was also greater. The second characteristic is that the MF models yielded better salt passage predictions (Table 3) than the SF models (Table 2). Salt passage values were well predicted with the MF mod-

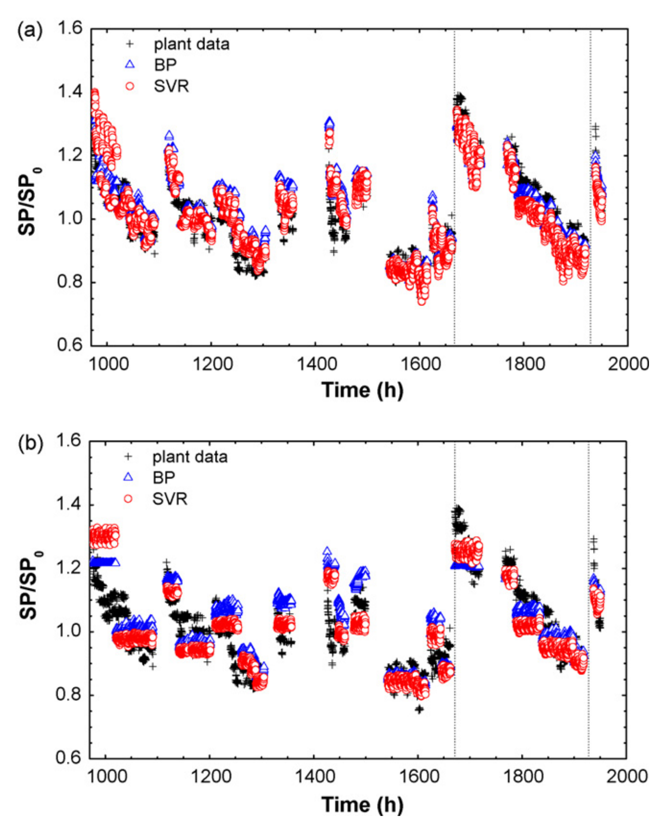


**Fig. 11.** Comparison between experimental SP/SP$_o$ values and those predicted with BP (8:3:1 architecture) and SVR-based SF models for test set with $\Delta t$ = 8 h. (a) $\Delta t_p$ = 2 h and (b) $\Delta t_p$ = 8 h.

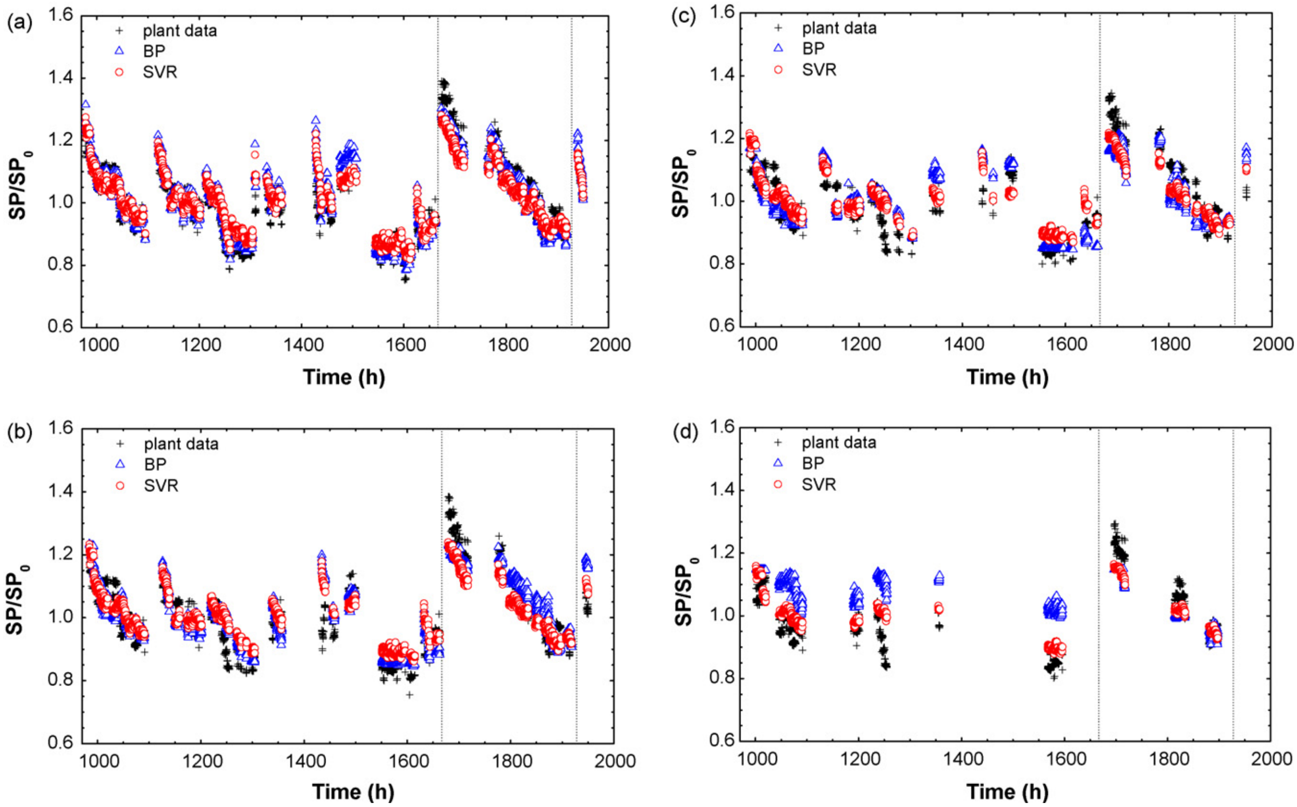


**Fig. 12.** Comparison between experimental SP/SP$_0$ values and those predicted with BP (7:3:1 architecture) and SVR-based MF models for the test set with $\Delta t = 8$ h. (a) $\delta = 2$ h; (b) $\delta = 8$ h; (c) $\delta = 12$ h and (d) $\delta = 24$ h.

els (Table 3) with $\bar{\varepsilon} \le 4.9\%$ and $q^2_{te} \ge 0.80$ for STM intervals, within the range of $\Delta t \le 24$ h and forecasting times of $\delta \le 24$ h, using both the BP and SVR algorithms. In contrast, reasonable performance for the SF-derived salt passage model (Table 2), over the same STM range $\Delta t \le 24$ h, was only feasible for much shorter forecasting intervals, i.e., $\Delta t_p \le 8$ h.

The performance of the BP and SVR-based SF and MF models for salt passage, for the complete test data set with $\Delta t = 8$ h, is illustrated in Figs. 11 and 12, respectively. It is seen that the models follow the trend of data reasonably, with the MF outperforming the SF-based model for all forecasting times (see bottom parts of Tables 2 and 3). It should be noted that marching forecast models, developed with either BP or SVR algorithms, are robust in forecasting salt passage even up to 1 day forecasting (Fig. 12). For example, the SVR-based MF models with $\Delta t = 8$, 12 and 24 h yielded salt passage predictions at $\delta = 24$ h with mean absolute errors of 0.049 salt passage units and $q^2_{te} \approx 0.8$ for these three cases (Table 3). The SF models are less accurate yielding predictions with higher mean absolute errors. For example, for an STM interval of $\Delta t = 24$ h and a forecasting time-interval of $\Delta t_p = 24$ h, the SVR-based SF model performed with a mean absolute error of $\bar{\varepsilon} \approx 5.8\%$ and lower quality of fit ($q^2_{te} \approx 0.71$) (Table 2). It should be recognized that the short-term memory information of the target variables in the SF models is maintained for all forecasting time-instants $\tau$ within $\Delta t_p$, thereby leading to smearing in the training process, especially when forecasted values change significantly with time and with increasing forecasting time-instants $\tau$ within the interval $\Delta t_p$. Notwithstanding, the forecasting models developed in the present study (e.g., Figs. 11 and 12) demonstrated superior sensitivity to following the temporal salt passage variability. In contrast, the ASP models (Fig. 6), which have no forecasting capability, had little or no ability to describe the temporal salt passage variability demonstrating poor model performance with $\bar{\varepsilon} \approx 10\%$ and $q^2_{te} \approx 0$ (Table 2).

## 4. Conclusions

ANN modeling approach with BP and SVR algorithms, introducing a STM time-interval as an input parameter, was evaluated for describing and forecasting the time-variability of plant performance. An ASP model and two types of forecasting models (sequential forecasting and matching forecast) for permeate flux and salt passage were investigated using real-time RO plant performance data. Analysis of various BP and SVR architectures and time-intervals demonstrated that it is feasible to model plant performance to a reasonable level of accuracy with respect to both permeate flux and salt passage, with a short-term memory interval of up to about 24 h. The present study focused on performance variability that was of short duration. Notwithstanding, the results of the present study suggest that there is merit in applying the present approach to plant operations that may involve longer time scales of performance degradation as is the case when membrane fouling and scaling occur. Current work is ongoing to extend and incorporate data-driven neural network-based models in a control strategy and early warning system of the deterioration of RO plant performance.

## Acknowledgments

This work was supported, in part, by the United States Environmental Protection Agency, the National Water Research Institute, the California Department of Water Resources, the International

Desalination Association, and the UCLA Water Technology Research Center. Financial support was also received from the Catalan Government (2005SGR-00735) and the Spanish Ministry of Education (CICYT, CTQ2006-08844). Francesc Giralt holds the "Distinció a la Recerca de la Generalitat de Catalunya."

**Nomenclature**

*Symbols*

| | |
|---|---|
| $C$ | capacity term in the support vector regression (SVR) model |
| $C_{b,a}$ | measured concentrate (brine) concentration at a given time, mg/L |
| $C_{b,s}$ | concentrate (brine) concentration at standard (reference) conditions, mg/L |
| $C_{f,a}$ | measured feed concentration at a given time, mg/L |
| $C_{f,s}$ | feed concentration at standard (reference) conditions, mg/L |
| Cn | conductivity, μS/cm |
| $EPF_a$ | RO permeate flow rate, GPM |
| $EPF_s$ | RO permeate flow at standard (reference) conditions, GPM |
| GPM | Gallons per min |
| $N_{\Delta t_p}$ | number of $\Delta t_p$ time-intervals |
| $N_{\Delta \tau}$ | number of $\Delta \tau$ time-intervals |
| $P_{c,a}$ | measured concentrate pressure, kPa |
| $P_{c,s}$ | concentrate pressure at standard (reference) conditions, kPa |
| $P_{f,a}$ | measured feed pressure, kPa |
| $P_{f,s}$ | feed flow pressure at standard (reference) conditions, kPa |
| $P_{p,a}$ | permeate pressure, kPa |
| $P_{p,s}$ | permeate pressure at standard (reference) conditions, kPa |
| $q^2$ | model quality index |
| $Q_{po}$ | standardized permeate flux at operational time $t = 0$ |
| $Q_{p,a}$ | measured permeate flow rate, GPM |
| $Q_{p,s}$ | standardized permeate flow rate, GPM |
| $SP_o$ | standardized salt passage at operational time $t = 0$ |
| $\%SP_a$ | measured percent salt passage |
| $\%SP_s$ | standardized percent salt passage |
| $t$ | overall operational time, min, h or days |
| $t_N^o$ | time-instants at the beginning of the $N_{\Delta t}$ intervals, min or h |
| $T$ | temperature, K |
| $TCF_a$ | temperature correction factor for actual process conditions |
| $TCF_s$ | temperature correction factor for standardization |
| $TDS_{NaCl}$ | total dissolved salts, expressed as equivalent NaCl concentration, mg/L |
| $Y$ | recovery |
| $\bar{y}$ | average y value |
| $y_i$ | experimental $y$ value |
| $\hat{y}_i$ | predicted $y$ value |
| $\delta$ | forecasting time in MF models, min or h |
| $\Delta t$ | short-term memory interval (STM) |
| $\Delta t_p$ | predictive interval in SF models, min or h |
| $\varepsilon$ | maximum allowed error in the $\varepsilon$-tube of the SVR models |
| $\bar{\varepsilon}$ | mean absolute error (%) |
| $\Delta \tau$ | plant data sampling interval (10 min) |
| $\pi_{b,a}$ | concentrate (brine) osmotic pressure at a given time, kPa |
| $\pi_{b,s}$ | concentrate (brine) osmotic pressure at standard (reference) conditions, kPa |
| $\pi_{p,a}$ | permeate osmotic pressure at a given time, kPa |
| $\pi_{p,s}$ | permeate osmotic pressure at the standard (reference) condition, kPa |
| $\sigma$ | dispersion of the radial basis function kernels used in SVR models |
| $\tau$ | time within the interval $\Delta t_p$, min or h |

## References

[1] K. Jamal, M.A. Khan, M. Kamil, Mathematical modeling of reverse osmosis systems, Desalination 160 (1) (2004) 29–42.
[2] K.L. Chen, L.F. Song, S.L. Ong, W.J. Ng, The development of membrane fouling in full-scale RO processes, J. Membr. Sci. 232 (1–2) (2004) 63–72.
[3] M. Dornier, M. Decloux, G. Trystram, A. Lebert, Dynamic modeling of crossflow microfiltration using neural networks, J. Membr. Sci. 98 (3) (1995) 263–273.
[4] H. Niemi, A. Bulsari, S. Palosaari, Simulation of membrane separation by neural networks, J. Membr. Sci. 120 (1995) 185–191.
[5] W.R. Bowen, M.G. Jones, H.N.S. Yousef, Prediction of the rate of crossflow membrane ultrafiltration of colloids: a neural network approach, Chem. Eng. Sci. 53 (22) (1998) 3793–3802.
[6] N. Delgrange, C. Cabassud, M. Cabassud, L. Durand-Bourlier, J.M. Laine, Modelling of ultrafiltration fouling by neural network, Desalination 118 (1–3) (1998) 213–227.
[7] N. Delgrange, C. Cabassud, M. Cabassud, L. Durand-Bourlier, J.M. Laine, Neural networks for prediction of ultrafiltration transmembrane pressure—application to drinking water production, J. Membr. Sci. 150 (1998) 111–123.
[8] N. Delgrange-Vincent, C. Cabassud, M. Cabassud, L. Durand-Bourlier, J.M. Laine, Neural networks for long term prediction of fouling and backwash efficiency in ultrafiltration for drinking water production, Desalination 131 (2000) 353–362.
[9] G.R. Shetty, S. Chellam, Predicting membrane fouling during municipal drinking water nanofiltration using artificial neural networks, J. Membr. Sci. 217 (1–2) (2003) 69–86.
[10] M.A. Razavi, A. Mortazavi, M. Mousavi, Application of neural networks for crossflow milk ultrafiltration simulation, Int. Dairy J. 14 (1) (2004) 69–80.
[11] A. Abbas, N. Al-Bastaki, Modeling of an reverse osmosis water desalination unit using neural networks, Chem. Eng. J. 114 (2005) 139–143.
[12] S. Chellam, Artificial neural network model for transient crossflow microfiltration of polydispersed suspensions, J. Membr. Sci. 258 (1–2) (2005) 35–42.
[13] H.Q. Chen, A.S. Kim, Prediction of permeate flux decline in crossflow membrane filtration of colloidal suspension: a radial basis function neural network approach, Desalination 192 (1–3) (2006) 415–428.
[14] G.B. Sahoo, C. Ray, Predicting flux decline in crossflow membranes using artificial neural networks and genetic algorithms, J. Membr. Sci. 283 (1–2) (2006) 147–157.
[15] R. Sitte, J. Sitte, Neural network system technology in the analysis of financial time series, in: T. Leondes Cornelius (Ed.), Intelligent Knowledge-based Systems, vol.5, Neural Networks, Fuzzy Theory and Genetic Algorithms, Kluwer Academic Publishers, 2005, pp. 59–110.
[16] R. Rallo, J. Ferre-Gine, A. Arenas, F. Giralt, Neural virtual sensor for the inferential prediction of product quality from process variables, Comp. Chem. Eng. 26 (2002) 1735–1754.
[17] F. Giralt, A. Arenas, J. Ferre-Gine, R. Rallo, G.A. Kopp, The simulation and interpretation of turbulence with a cognitive neural system, Phys. Fluids 12 (7) (2000) 1826–1835.
[18] ASTM D 4516-00, Standard Practice for Standardizing Reverse Osmosis Performance Data, in American Society of Testing Materials, 2000.
[19] T.D. Wolfe, Membrane Process Optimization Technology, Bureau of Reclamation, Desalination and Water Purification Research and Development Report No. 100, 2003.
[20] OLI, OLI Analyzer 2.0, OLI Systems, Morris Plains, NJ, 2005.
[21] G. Box, G.M. Jenkins, G. Reinsel, Time Series Analysis: Forecasting & Control, 3rd edition, Prentice Hall, Upper Saddle River, NJ, 1994.
[22] P. Bhagat, An introduction to neural nets, Chem. Eng. Prog. 86 (8) (1990) 55–60.
[23] V. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 1995.
[24] B. Schölkoft, C.J.C. Burges, A.J. Smola (Eds.), Advances in Kernel Methods: Support Vector Learning, MIT Press, Cambridge, 1999.
[25] O. Ivanciuc, Applications of support vector machines in chemistry, in: K.B. Lipkowitz, T.R. Cundari (Eds.), Reviews in Computational Chemistry, vol. 23, Wiley-VCH, Weinheim, 2007, pp. 291–400.
[26] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 2002.
[27] G.E. Hinton, How Neural Networks Learn from Experience, Sci. Am. 267 (3) (1992) 145–151.

[28] S.P. Chitra, Use neural networks for problem-solving, Chem. Eng. Prog. 89 (4) (1993) 44–52.
[29] A. Tropsha, P. Gramatica, V.K. Gombar, The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models, QSAR Comb. Sci. 22 (1) (2003) 69–77.
[30] R. Carbo-Dorca, D. Robert, Ll. Amat, X. Girones, E. Besalu, Molecular Quantum Similarity in QSAR and Drug Design, Lectures Notes in Chemistry, vol. 73, Springer Verlag, Berlin, 2000.
[31] A. Tropsha, Variable selection QSAR modeling, model validation, and virtual screening, in: Y. Martin (Ed.), Ann. Rev. Comp. Chem., Chapters 4 and 7, Elsevier, 2006, pp. 113–126.
[32] Guidance Document on the Validation of (Quantitative) Structure–Activity Relationships [(Q)SAR] Models, Organization for Economic Cooperation and Development, Paris, 2007.
[33] T. Kohonen, The self-organizing map, Neurocomputing 21 (1–3) (1998) 1–6.
[34] T. Kohonen, The Self-Organizing Map, Proc. IEEE 78 (9) (1990) 1464–1480.
[35] D.L. Davies, D.W. Bouldin, Cluster separation measure, IEEE Trans. Pattern. Anal. Mach. Intell. 1 (2) (1979) 224–227.
[36] G.J. Bowden, H.R. Maier, G.C. Dandy, Optimal division of data for neural network models in water resources applications, Water Resour. Res. 38 (2) (2002) 2.1–2.11.

Please cite this article in press as: D. Libotean, et al., Neural network approach for modeling the performance of reverse osmosis membrane desalting, J. Membr. Sci. (2008), doi:10.1016/j.memsci.2008.10.028