Artificial Neural Networks Based Modeling of the Reverse Osmosis Process Performance

دراسة نمذجة أداء عملية التناضح العكسي باعتماد الشبكات العصبية الاصطناعية

Assistant lecturer / Asseel Majid Rasheed
Electrical & Electronics Eng. Dept. / Collage of Engineering LUniversity of Kerbala

ABSTRACT

This investigation presents a methodology and practical guidelines for developing predictive models for reverse osmosis plants by a data-based approach using neural networks based on the back-propagation algorithm. This study utilizes actual operating data from reverse osmosis (RO) desalination plants. Our resulting neural network model is capable of accurately predicting the actual operating data from RO desalination plants, but the accuracy of a neural network model depends on both the proper selection of input variables and the broad range of data with which the network is trained. A neural network model can handle noisy data more effectively than statistical regression and performs better in predicting the performance variables of RO desalination plants. Permeate flux and salt passage are the key performance parameters. They are mainly influenced by variable parameters such as pressure, temperature, salt concentration, feed flow rate and pH of feed water. When the temperature of feed water is increased, the permeate flux and salt passage increase and permeate flux increases with decreasing pH and concentration of feed water and increasing pressure. The salt passage decreases with increasing pressure, when the concentration of feed and pH decrease too. When increasing the feed flow rate, the permeate flux and salt passage would be increased. A good agreement prediction is obtained using the ANN predictions and the experimental data with a deviation not more than 2% for most of the cases considered. The ANN interpolative levels (which were not represented in the training phase) is shown to be of lesser quality.

الخلاصة :

أن هذا البحث يقدم منهجيه و أدِلَّه عمليه من اجل تطوير نماذج تنبئية لمنظومات التناضح العكسي بواسطة طريقة قاعدة البيانات مستخدماً الشبكة العصبية الاصطناعية المعتمدة على منهاج التكاثر الرجعي. علماً أن هذه الدراسة تستخدم بيانات التشغيل الفعلية من منظومات تحلية المياه بواسطة التناضح العكسي. حيث يعد النموذج الناتج لشبكه العصبية الاصطناعية قادراً على تنبّئ دقيق للبيانات العملية الفعلية من تحلية المياه بواسطة التناضح العكسي، لكن دقة الموديل للشبكة العصبية يعتمد على تحديد مناسب للمتغيرات الداخلة والمدى الواسع للبيانات مع الشبكة المتدربة. يمكن لنموذج الشبكة العصبية أن يعالج بيانات مضطربة بأكثر فعالية من انحسار إحصائي وقد يؤدي الأفضلية في تنبئ فعالية المتغيرات لمنظومات تحلية المياه بواسطة التناضح العكسي. كما أن الجريان الماء النافذ والنفوذ المِلْحِيّ يكونان الأبعاد الفعلية الرئيسية في الدرجة الأولى والتي تتأثر بالمتغيرات مثلا الضغط ودرجة الحرارة وتركيز الملح ومعدل الجريان وpH للماء الداخل. ومع ازدياد درجة حرارة الماء الداخل فان جريان الماء النافذ والنفوذ المِلْحِيّ يزداد, حيث ان جريان الماء النافذ يزداد بانخفاض pH وتركيز ملح الماء الداخل وازدياد الضغط. وكذلك النفوذ المِلْحِيّ يقل مع ازدياد الضغط وانخفض تركيز الملح وpH للماء الداخل أيضا. فإذا أزداد معدل جريان الماء الداخل، فان جريان الماء النافذ والنفوذ المِلْحِيّ سوف يزدادان. أن التنبؤ المنسجم الجيد يكون مُكْتَسَب باستخدام تنبؤات الشبكة العصبية الاصطناعية والبيانات العملية مع انحراف ليس اكثر من 2% لأغلب الحالات. وان مستويات الاستيفاء للشبكة العصبية الاصطناعية (التي ليست مُعَلَّمه في طور التدريب) تكون معروضة لأقل نوعيه.

1. INTRODUCTION

The concepts of "osmosis" and "reverse osmosis" have been known for many years. Studies on osmosis were carried out as early as 1748 by the French scientist Nollet. Many researches reported on reverse osmosis system, most of them using hollow fiber or spiral wound membrane in their research (Shamel and Chung, 2006). The separations performance of reverse osmosis (RO) desalination (e.g., salt passage and permeate flux) and membrane longevity are impacted by numerous factors including, but not limited to different operating parameters (feed water concentration, temperature, pressure and flow rate) effects on membrane performance are examined using RO system (Hawlader et al., 2000). 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 (Jamal et al., 2004).

In recent years, there have been various attempts to use the artificial neural networks (ANN) as a viable approach to develop data-driven models to describe the performance of membrane processes (Niemi et al., 1995; Abbas and Al-Bastaki, 2005; Chen and Kim, 2006; Sahoo and Ray, 2006). Neural networks can be applied to the predictive modeling and optimization of desalination plants. The goals are to achieve better design, improve process efficiency, and enhance operational safety. Desalination plants make good candidates for neural network modeling, because of their computational process complexity, nonlinear behavior, many degrees of freedom, and the presence of uncertainty in the control environment (Rao et al., 1994). Quantitative optimization of operating variables could lead to increased production rates, higher product quality, and better plant performance with less energy consumption and lower operating costs. This optimization can also give the operator an early warning of any decline in unit performance (Al-Shayji and Liu, 2002).

Previously developed ANN models of RO plant performance were based on the use of training data sets whereby the data points for training and testing were inter dispersed throughout the complete data. 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 data that were not covered by the training data set). The ability to forecast membrane plant performance, even for short forward data, 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 (Libotean et al., 2009).

This study is aimed at investigating the reverse osmosis process performance of permeate flux and salt passage and describes a methodology and practical guidelines of developing predictive models of water desalination plants by a data-based approach using a neural network based on the back-propagation algorithm. The performance of RO for different operating parameters has been examined and the best operating condition of RO system with the best permeate quality were utilized to produce fresh water from salt water.

2. PROCESS DESCRIPTION

This research is a pilot plant experimental study and the essential components of the pilot used includes a feed tank of 3 m^3 capacity with a mixer (1500 rpm), a pressure pump (centrifuge pump) of type (KSB/COK-C50-250) of 30000 L/h capacity with speed 2900 rpm to control pressure within range (4 – 6 bar) and the RO membrane module (Figure.1).

The RO unit use in Daura Power Plant is a spiral wound module of TFC-8822HR made of polyamide by Fluid System Company. Also the dimension of membrane (93.5 cm × 105 cm) and the module contains 34 membranes; its surface area is 333795 cm^3, separate the membranes from one another by mesh made of fine plastic material to assist to pass the water between the membranes. For starting the work distills water (conductivity 2 µs/cm) is passed from the RO pilot to determine the initial flux in various pressures of 4, 4.5, 5, 5.5 and 6 bar at an ambient temperature of 25 ºC. Then, the RO system is operated to determine the optimized applying pressure. At last, the optimum values of temperature, pH and NaCl concentration is investigated at the optimum pressure. The amount of permeate flux and salt passage which can pass from the RO membrane is measured. In each step of the experiments, analyses of samples are done after measurement of feed flow; permeate flux and salt passage in a predetermined period of less than 30 min. for about 3 months.

Studying the effect of temperature on RO treatment is accomplished for 18, 19.3, 20.6, 21.9 and 23.2 degrees Centigrade adjusted by using heater with thermocouple and effect of NaCl concentration is examined using of five solutions of NaCl (391, 450, 504, 564 and 618 mg/L). All these synthetic samples are prepared in the laboratory by dissolving certain amounts of a pure salt of food in distilled water and the concentrations of salt is measured by means of electric conductivity. The variation effect of pH on RO performance accomplished for 5, 6, 7, 8 and 9 to control pH by addition little amount of HCl (9.5 N) or NaOH (9.9 N) for all experiments.

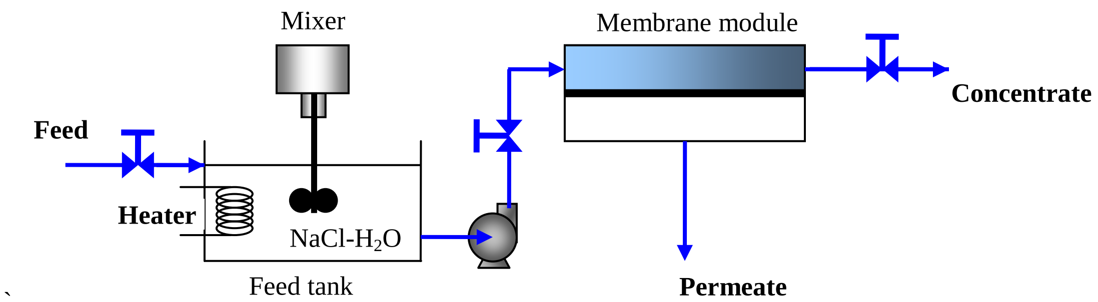


Figure 1. Schematic of the reverse osmosis (RO) pilot plant.

3. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks (ANN) are numeric techniques able to capture and represent complex input-output relationships. They have the ability to learn linear, as well as non-linear correlative patterns between sets of input data and corresponding target values, directly from the data set that is modeled. They can also be successfully used in classification problems, since there are specific algorithms available to group the input patterns in different clusters based on similarities-dissimilarities between them. The ANN are characterized by processing units (neurons) and adjustable parameters (weights) (Bhagat, 1990).

ANN models were built, separately, for the permeate flux and salt passage (i.e., target variables) using a back-propagation algorithm (Bhagat, 1990) to establish the relationships between the selected model inputs and the target variables (permeate flux and salt passage). Baughman and Liu, 1995 described the fundamentals and applications of neural networks in bioprocessing and chemical engineering, including the back-propagation algorithm for the development of multilayer feed forward networks. Back-propagation is a neural network training method based on a forward flowing of information, and back-propagated error corrections (Bhagat, 1990; Vapnik, 1995). The back-propagation networks are usually organized in layers of neurons, as the architecture presented in the figure.2.

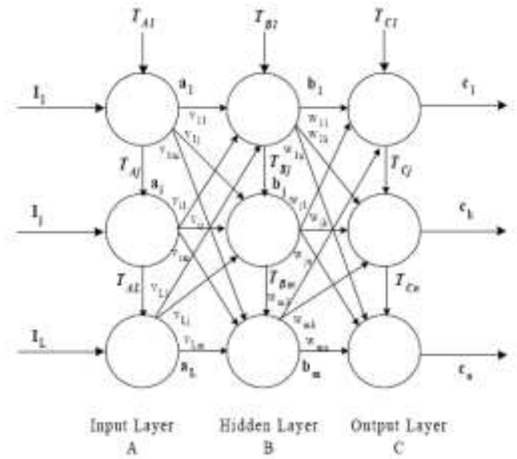


Figure 2. Multilayer feed forward neural network (Al-Shayji and Liu, 2002).

Connections are made between the neurons of adjacent layers: a neuron is connected so that it receives signals from each neuron in the preceding layer and transmits signals to each neuron in the immediately succeeding layer. Each processing element (neuron) receives a number of inputs, using the neuron's assigned weights, which is transformed by an activation function to produce a single output signal that is sent to the neurons in the succeeding layer. The activation function defines the output of the neuron in terms of the activity level at its input. Briefly, the current back-propagation architectures consist of one input layer (with the number of inputs required by each model tested) which receives the input data, one or more hidden layers in which a different number of neurons are used for different models to evaluate the performance of various model architectures and one output layer (one output target variable). Additionally, a bias neuron that supplies an invariant output is connected to each neuron in the hidden and output layer (Bishop, 2002). Different expressions can be used for the neuron's activation function like a sigmoid, hyperbolic tangent or linear transfer functions. Linear and hyperbolic tangent transfer functions are utilized for the input and output layers, respectively and a hyperbolic tangent transfer function is used for the hidden layer (Bishop, 2002; Libotean et al., 2009).

In the ANNs approaches, data normalization is necessary before starting the training process, to ensure that the influence of the input variable in the course of model building is not biased by the magnitude of their native values, or their range of variation. The zero-mean normalization technique used for the input/output variables to the range [-1, +1], having an average value set at zero. This technique utilizes the entire range of the hyperbolic tangent transfer function, and every input variable in the data set has a similar distribution range (Al-Shayji and Liu, 2002).

The back-propagation training consists of two passes of computation: a forward pass and a backward pass. In the forward pass an input pattern vector is applied to the neurons in the input layer. The signals from the input layer propagate to the units in the first hidden layer, each one producing an output as described above. The outputs of these neurons are propagated using the same algorithm to units in subsequent layers until the signals reach the output layer where the actual response of the network to the input vector is obtained. Extending the formula for calculating the output of a single neuron for the general case of any unit from any layer, leads to the networks' weights, that are fixed during the forward pass, are all adjusted during the backward pass in accordance with a back-propagated error signal for minimizing an error function (Chitra, 1993). The error gradient is calculated based on the difference between the target value and the neuron's output, while for the hidden layer neurons the error gradient is determined by calculating the weighted sum of errors at the previous layer. The principle used for weights adaptation is also known as generalized delta rule. Once the error gradients are evaluated for every layer, the biases and the weights are updated (Hinton, 1992).

A more efficient method used for weights adaptation as the Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963), which is a combination between the gradient descent rule and the Gauss-Newton method. The algorithm uses a parameter to decide the step size, which takes large values in the first iterations (equivalent with the gradient descent algorithm), and small values in the later stages (equivalent with the Gauss-Newton method). It combines the ability of both methods (i.e., convergence from any initial state in the case of gradient descent, and rapid convergence when reach the vicinity of the minimum error in the case of Gauss-Newton method) while avoiding their drawbacks (Bishop, 2002; Hagan and Menhaj, 1994). For the learning phase, the data must be divided in two sets: the training data set, which is used to calculate the error gradients and to update the weights, and the validation data set, which allows to select the optimum number of iterations in which the networks learns general information from the training set. As the number of iterations increases, the training error drops whereas the validation data set error begins to drop, then reaches a minimum and finally increases. Continuing the learning process after the point when the validation error arrives to a minimum leads to a process called over-fitting, when the network became specific to the pattern vectors that form the training data set. After finishing the learning process, another data set (test set) is used to validate and confirm the prediction accuracy (Delgrange et al., 1998).

The goal of a neural network is to map a set of input patterns onto a corresponding set of output patterns. The network accomplishes this mapping by first learning from a series of past examples defining sets of input correspondences for the given system. The network then applies what it has learned to a new input pattern to predict the appropriate output. Developing a neural network requires three phases: training, recall, and generalization. The training phase repeatedly presents a set of input-output patterns to the network, adjusting the weight of interconnections between nodes until the specified inputs yield the desired outputs. The recall phase subjects the network to a wide array of input patterns seen in training to test its memory. The generalization phase tests the network with new input patterns, for which the system will hopefully perform properly (Al-Shayji and Liu, 2002). We used Neural Network Toolbox / MatLab (the Language of Technical Computing, Version 6.1.0.450 Release 12.1) manufactured by The MathWorks, Inc. (2001) software tools to accomplish this work.

4. DATA PREPROCESSING AND ANALYSIS

To permit accurate monitoring of the operation of the RO unit, we collect data of varied natures and log them properly. In any experiment, as the number of observations increases, the resulting statistical correlations become increasingly reliable. Therefore, the investigator should use a large sample size wherever possible. In this study, we recorded the operational variables from the RO pilot plant measured every 30 min. for a period of 3 months (4375 data sets for each period).

Real-life data often contain outliers, which are observations that do not reasonably fit within the pattern of the bulk of the data points and are not typical of the rest of the data. Some outliers are the result of incorrect measurements and can be immediately rejected and removed from the data set. Other outliers are observations caused by unusual process phenomena that are of vital interest. Data require careful inspection and examination to observe this distinction. The inclusion of outliers in training data forces the network to consider a larger solution space and can therefore reduce the overall precision of the resulting network. This is observed as occasional large differences between actual and predicted values of output variables. The root-mean-square (RMS) error decreased after the outliers had been removed. Removing outliers generally improves network performance.

5. MODEL DEVELOPMENT AND OPTIMIZATION

5-1. Network Parameters

To control the prediction ability of neural network for the modeling and optimization of RO process that is to say, the practical guidelines for developing and optimizing neural network models for RO desalination processes. Importantly, these guidelines represent effective starting points for neural network modeling of RO process.

The first step in neural computing, prior to training a neural network, is to initialize the weight factors between any two nodes of the hidden layers. Since no prior information about the system being modeled is available, it is preferable to set all the free weight factors of the network to random numbers that are uniformly distributed (Gaussian weight-factor distribution) inside a small zero-mean range of values, say, between -0.5 and +0.5. A multilayer prediction network trained with the back propagation algorithm will, in general, learn faster when the transfer function built into the network is symmetric (hyperbolic tangent function, with values between -1 and +1) rather than nonsymmetrical (sigmoid function, with values between 0 and 1). Therefore, we use a hyperbolic tangent transfer function throughout this study.

We normalize the input and output data sets between limits of -1 and +1, having an average value set at zero. This zero-mean normalization method utilizes the entire range of the hyperbolic tangent transfer function, and every input variable in the data set has a similar distribution range in conformity with (Al-Shayji and Liu, 2002).

5-2. Effect of Number of Training Data

Neural network is designed to predict the behavior of output variable based on number of training data for input variable and it is also unable to consider the interactions between input variables. Development of neural networks for predictive modeling of water desalination process, as our work utilizes actual operating data (not simulated data). Two subsets of data are used to build a model: a training set and a testing set. The training phase needs to produce a neural network that is both stable and convergent. Therefore, selecting what data to use for training a network is one of the most important steps in building a neural-network model.

Training is the process by which the neural network systematically adjusts the weights of interconnections between nodes so that the network can predict the correct outputs for a given set of inputs. For the best “learning” possible, we need a large and robust set of historical input/output data. It adjusts the weight factors and the process continues, iteratively, until the error (i.e., sum-of-square errors) between the predicted and actual outputs has been minimized for effective initial weight-factor distributions used in training back propagation networks based on number of training data of input variables and total number of nodes in the hidden layer(s). Figure.3 shows affect the number training data which increase the number of training data led to increase the degree of simulation between the predicted and actual data by means of Neural Network Toolbox / MatLab.

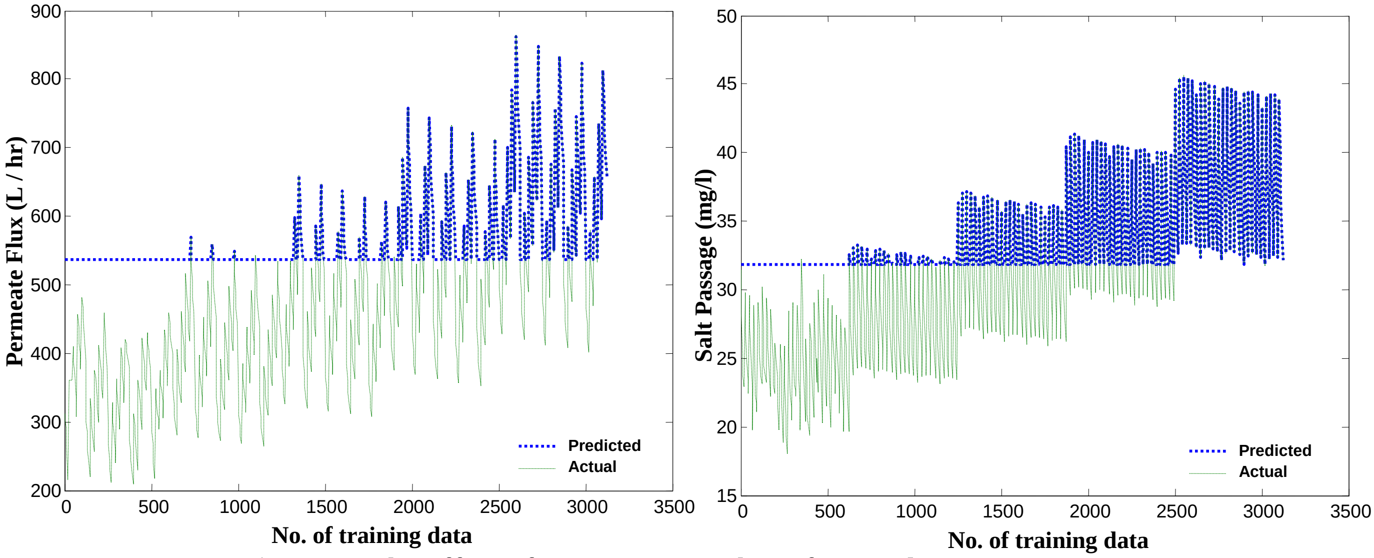


Figure 3. The effect of increasing number of examples in training sets.

5-3. Effect of Number of Nodes in the Hidden Layer(s)

The number of input and output nodes corresponds to the number of network inputs and desired outputs, respectively. The choice of the number of hidden layers and the nodes in the hidden layer(s) depends on the network application (Libotean, 2009; Baughman and Liu, 1995). Determining the number of hidden layers is a critical part of designing a network and it is not straightforward as it is for input and output layers. To determine the optimal number of hidden layers, and the optimal number of nodes in each layer, we train the network using various configurations, and then select the configuration with the fewest number of layers and nodes that still yield the minimum (RMS) error quickly and efficiently. The network consists of 5 input variables, and 2 output variable. It uses a back-propagation network, the delta learning rule, the hyperbolic tangent transfer function, 0.3 for the learning rate, and 0.4 for the momentum coefficient. We use 325 data sets to train these configurations with 10,000 iterations. The network is tested with one- and two-hidden-layer configurations with an increasing number of nodes in each hidden layer(s). Figure.4 illustrates the network response as the number of nodes in one- and two-hidden-layer networks increases. The results show that the 2-hidden-layer network performs significantly better than the 1-hidden-layer network in accordance with (Al-Shayji and Liu, 2002).

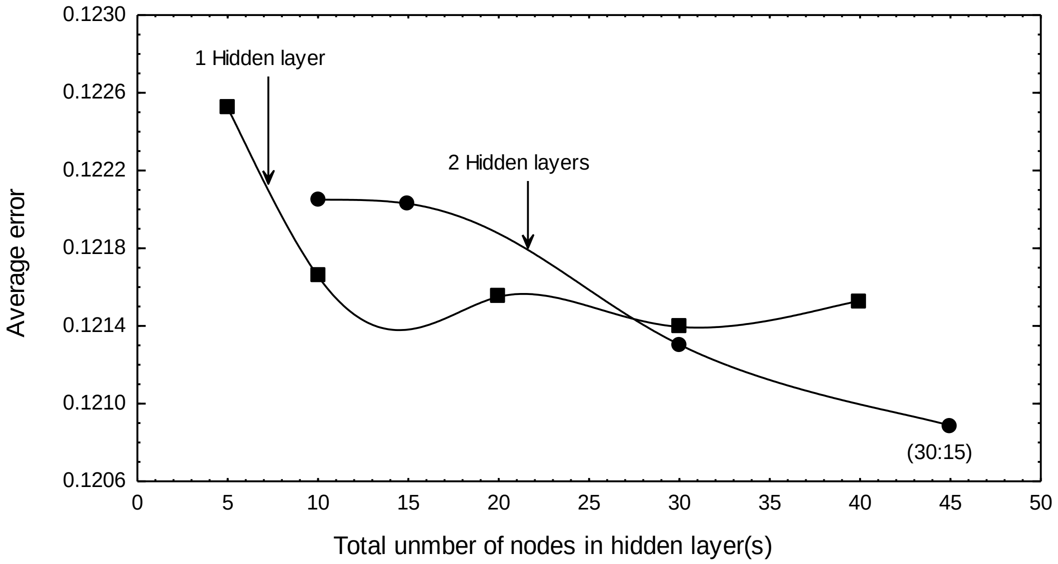


Figure 4. Comparison of the average errors for the prediction network for the permeate flux and salt passage trained with various one- and two-hidden-layer configurations.

The average errors for the best performing architecture for each length of training time considered are in agreement with the former analysis based on explained variance in prediction and number nodes in hidden layers, as illustrated in figure.5. The optimal configuration in 2-hidden-layer networks with minimum average error is 30:15 (i.e., with 30 nodes in hidden layer 1 and 15 nodes in hidden layer 2). This configuration agrees with what (Baugham and Liu, 1995) recommend as effective hidden nodes in their text and it will be used throughout our work.

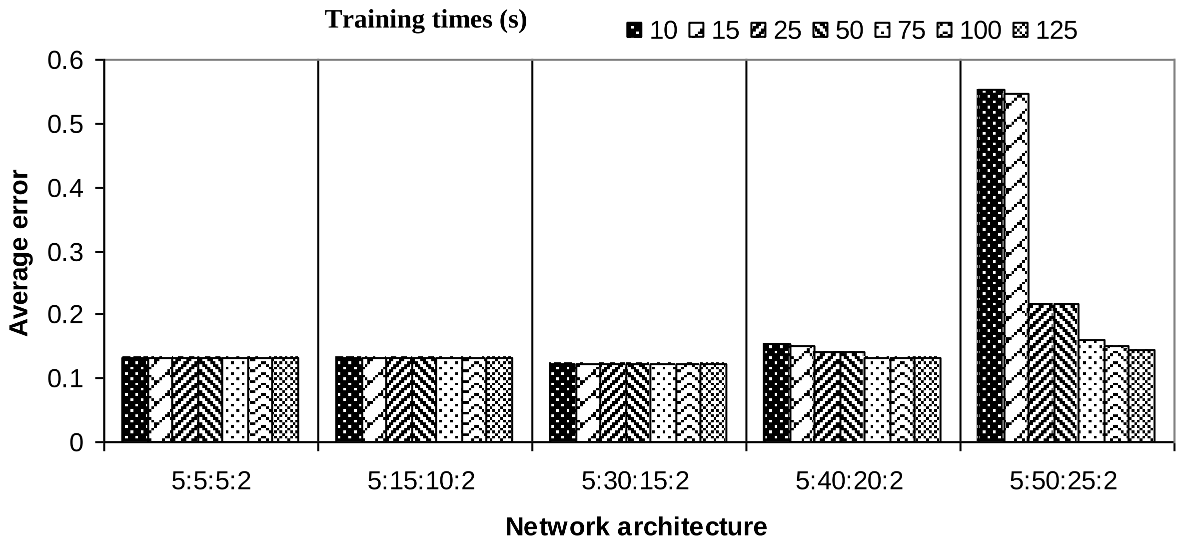
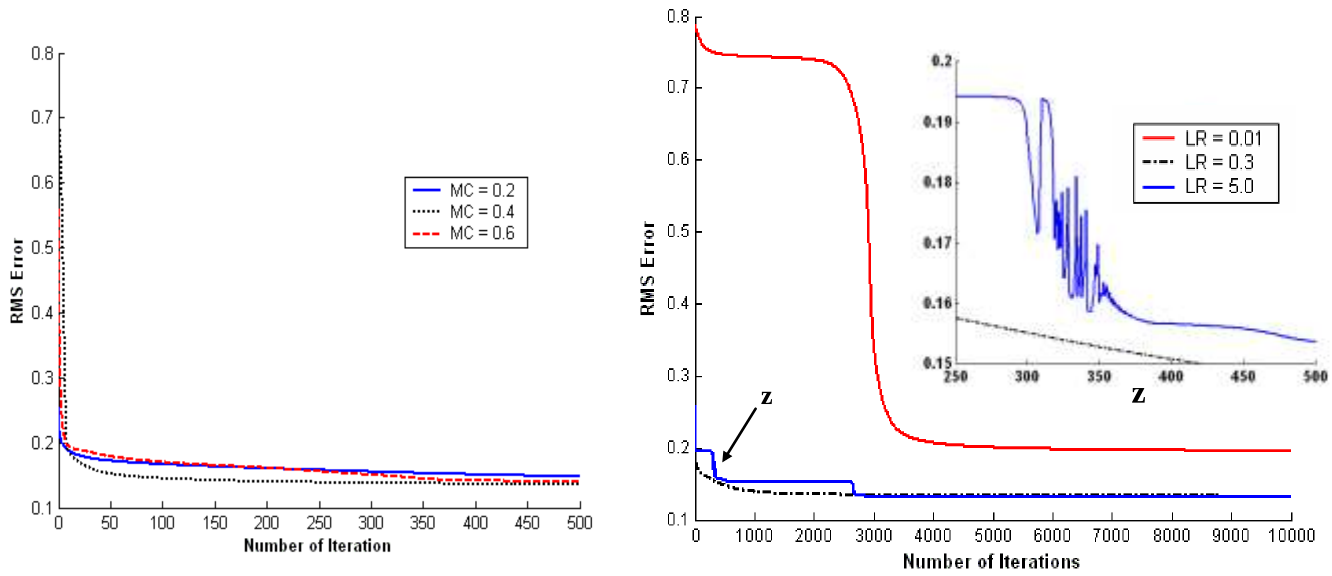


Figure 5. Average error for the best architecture for each length of training time for the prediction network for the permeate flux and salt passage trained with two-hidden-layer configurations.

5-4. Effects of Learning Rate and Momentum Coefficient.



(a) (b)

Figure 6. (A-B) Comparison of the RMS error in training the permeate flux and salt passage network with different learning rates (LRs) and momentum coefficients (MCs).

The learning rate and momentum coefficient are two important parameters that control the effectiveness of the training algorithm. The learning rate is a positive parameter that regulates the relative magnitude of weight changes during learning. However, how would a change in the learning rate change the performance of the algorithm? To understand the effect of the learning rate on the network training, let us consider the prediction network for the permeate flux and salt passage with 325 training examples. We use a back-propagation network with the 30:15 hidden-layer configuration, the delta learning rule, the hyperbolic tangent transfer function, and zero momentum coefficient. Figure.6 (b) compares the RMS error using a low learning rate of 0.01, a moderate learning rate of 0.3, and a high learning rate of 5.0. In general, a smaller learning rate results in slower convergence. When the learning rate is low (0.01), the network takes a longer time (4,000 iterations) to reach an RMS error of 0.2. This is due to the fact that the smaller the learning rate, the smaller will the changes to the weights in the network be from one iteration to the next, and the larger the number of update steps needed to reach a minimum.

However, when the learning rate is 0.3, the network reaches an RMS error of 0.2 in a shorter time (200 iterations). If increase the learning rate to 5.0, such that we will be taking larger steps, the algorithm will become unstable; the oscillating error fluctuations will increase instead of decaying and thus not reaching a minimum (see point (z) in Figure.6(b)).

Figure.6 (a) illustrates the effect of increasing the momentum coefficient on the speed of convergence. The network takes a longer time (500 iterations) to reach an RMS error of 0.15 when no momentum coefficient is used. By contrast, it reaches the same error at a shorter time (50 iterations) when the momentum coefficient is increased to 0.4. Therefore, to avoid the danger of instability and improve convergence as we increase the learning rate, a momentum coefficient is introduced, which will smooth out the oscillation. The momentum coefficient is a constant, between 0 and 1, used to promote stability of weight adaptation in a learning rule, and it tends to accelerate descent in a steady downhill direction. In back-propagation with momentum, the weight changes in a direction that is a combination of the current gradient and the previous gradient. This will help in moving the minimization routine out, if during training, it is prevented excessive weight changes and possible oscillation, the algorithm slows down the weight changes by a term that is proportional to the previous weight change and the momentum coefficient. Accordance with (Al-Shayji and Liu, 2002), a smaller learning rate results in a slower convergence and that as the learning rate decreases, a larger momentum coefficient increases the speed of convergence. In other words, it decreases the training time. The use of high learning rate and momentum coefficient causes oscillation in the RMS error during learning.

5-5. Neural Networks versus Statistical Regression

We use three subsets of data to build a model: a training set, a testing set and a generalization set. Neural networks interpolate data very well, but they do not extrapolate. Therefore, we should choose the training set to include data from all regions of desirable operation and we pick the sets as equally spaced points throughout the original data. To effectively visualize how well a network performs recall and generalization, we often generate a training curve, which represents the RMR error for both the recall of training data sets and the generalization of the testing data sets.

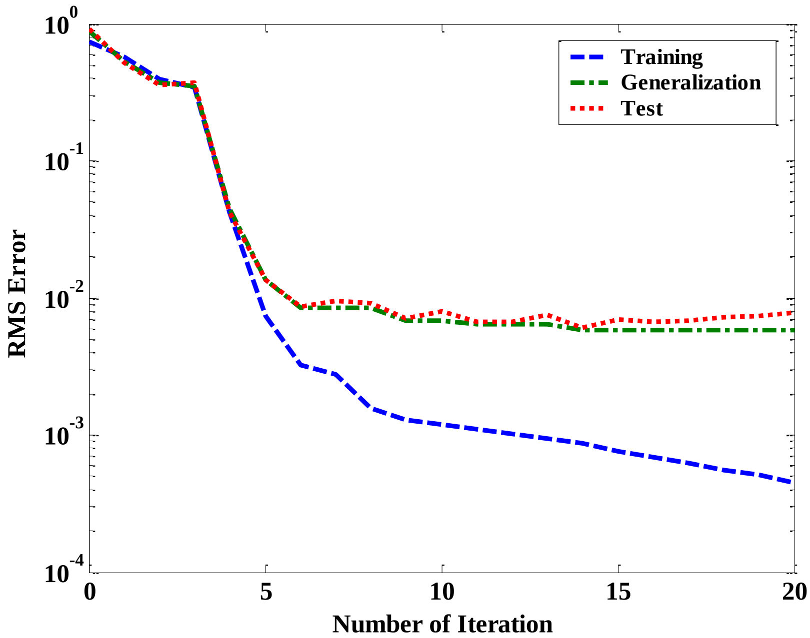
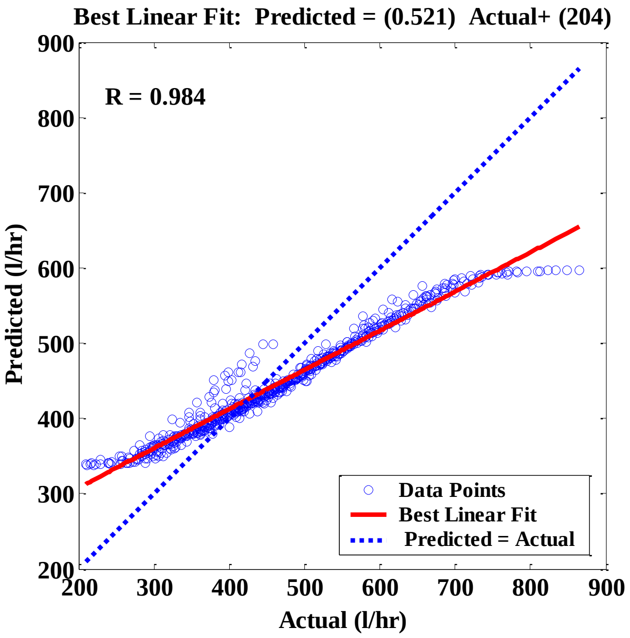


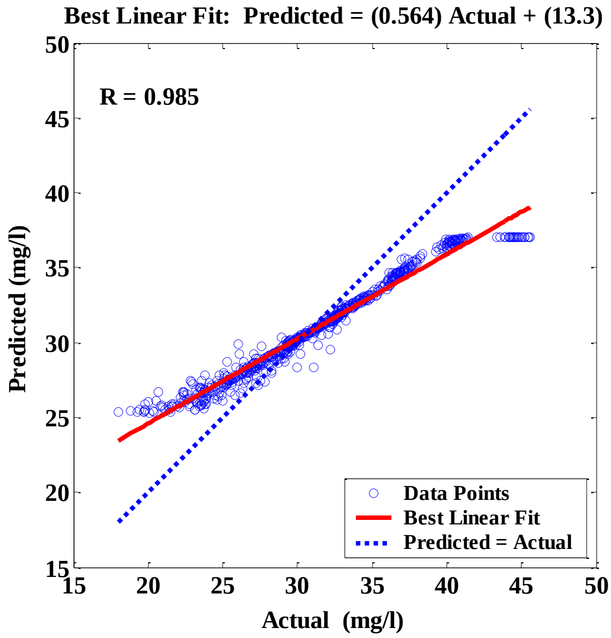
Figure 7. Training curve to comparison the RMS error for both the recall of training data sets and the generalization of the testing data sets.

Figure.7 shows the training stopped after 20 iterations because the generalization error increased. It is a useful diagnostic tool to plot the training, generalization and test errors to check the progress of training. The result here is reasonable, since the test set error and the generalization set error have similar characteristics, and it doesn't appear that any significant overfitting has occurred.

The next step is to perform some analysis of the network response. We will put the entire data set through the network (training, generalization and test) and will perform a linear regression between the network outputs and the corresponding targets. In this case, we have two outputs, so we perform two regressions. Figure.8 present the predicted and actual outputs from the neural networks and from statistical regression. The correlation coefficient R values for the permeate flux and salt passage are: 0.984 and 0.985, respectively. Despite the very good R values, the neural network still outperforms the regression analysis. Neural networks have been very effective in predicting and optimizing the performance variables of the RO desalination plants. It also outperforms the regression models in prediction problems.



(A) Permeate Flux / Statistical Regression



(B) Salt Passage / Statistical Regression

Figure 8. Actual and predicted RO plant outputs from neural network and statistical regression: (A) Permeate Flux; and (B) Salt Passage.

6. RESULTS AND DISCUSSION

6-1. Effect of pressure on permeate flux and salt passage (ANN predictions)

The permeate flux and salt passage with different pressures, pHs and concentrations are studied by varying operating pressure (Feed water pressure) from 4 - 6 bar. Consistent with (Ujang and Anderson, 1998) and present in figure.9 (a-b) shows the increase of pressure will improve quantity of permeate flux. This is due to the driving force of reverse osmosis membrane is transmembrane pressure. All of the water fed to membrane resulted in a similar trend. The results presented in figure.9 (a) indicate that concentration of 450 mg/l results highest permeate flux, followed by 564 mg/l and 618 mg/l concentrations. On the other hand, permeate flux increases with decreasing pH and increasing pressure, as shows in figure.9 (b) that membrane can produce permeate flux up to 540 l/hr at pH of 6 and 516 l/hr at pH of 9 at 6 bar.

The next study, the influence of pressure to salt passage content is depicted in Figure.7(c-d) indicates that operating pressure will tend to increase salt passage and then the lower it. During membrane operation, membrane will retain solid molecule at its surface, and the solid will accumulate at membrane surface that is in agreement with (Mulder, 1996). With existence of pressure increase operates for, hence bait which can penetrate membrane layer more and more so that sum up solid which retained at membrane surfaces also more and more. This matter will result membrane ability to retain solid component will be on the wane. Descend of this membrane ability expressed in salt passage tend to downhill along with pressure increase. It is clear from the figure.9(c-d) that the salt passage decreases with high pressure and lower concentration of feed and pH.

The ANN's prediction of permeate flux and salt passage is shown as lines in figure.9. Solid symbols are used to identify experimental data points which are used in the training phase of the ANN while open symbols are used for data points not included in the training phase, i.e., points used in the generalization phase. According to points not included in the training phase, i.e., of the ANN while open symbols are used for data figure.9, it can be seen that the ANN successfully predicts the nonlinear behaviour of permeate flux and salt passage vs pressure. A good prediction is obtained for all concentration and pH levels with very small deviation (< 2%) between the experimental data and the ANN predictions to compare with predictions by (Al-Zoubia et al., 2007) has deviation (< 10%) between the experimental data and the ANN predictions .

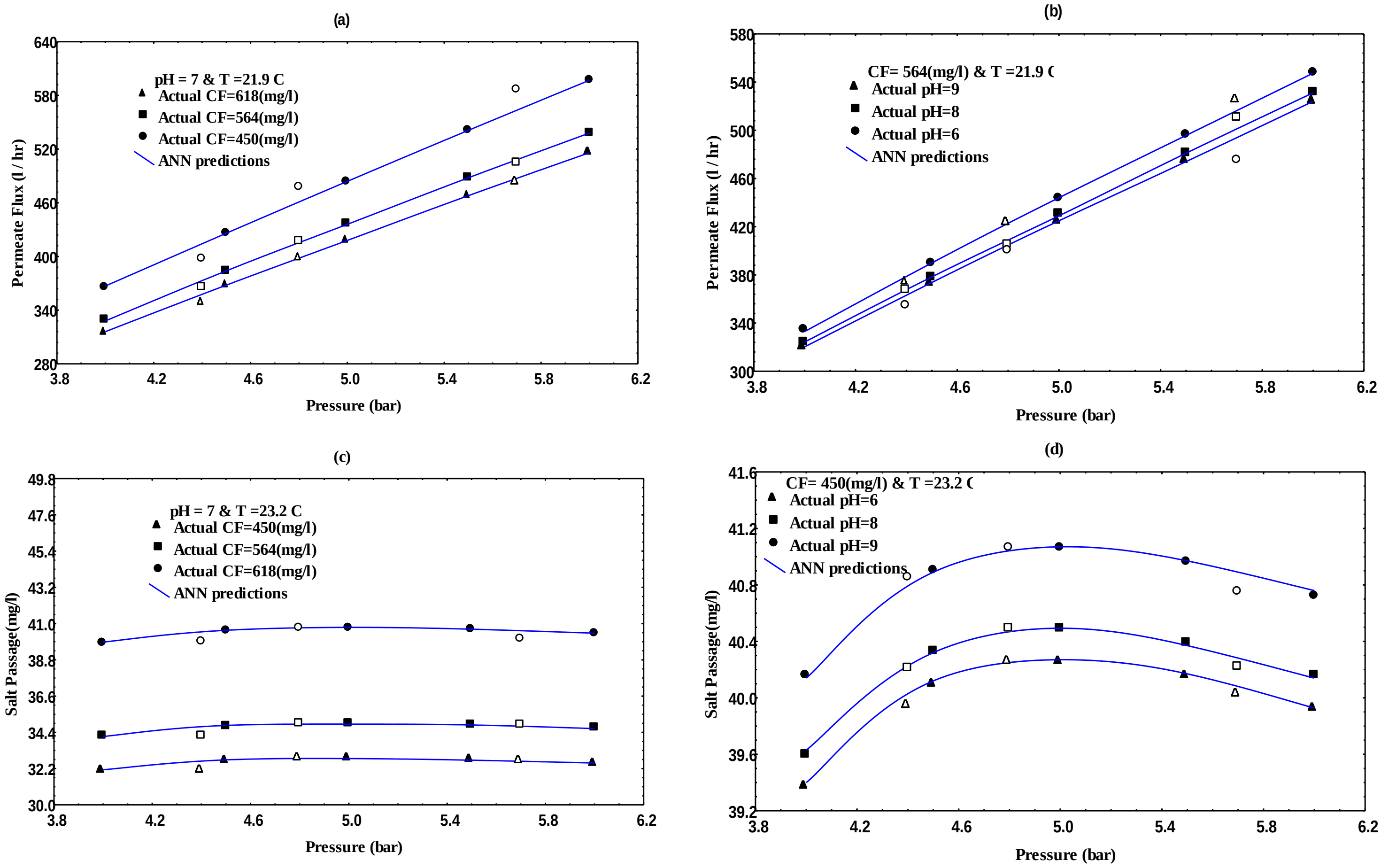


Figure 9. Neural network prediction of the permeate flux and salt passage vs. pressure using two effective factors: (a & c) concentrations of feed (CF), (b & d) pH. Experimental data used in the training process are marked as symbols, whereas lines represent the best fit of the net prediction.

6-2. Effect of pH on permeate flux and salt passage (ANN predictions)

According to (Bellona and Drewes, 2005) and shown in figure.10 (a-d), the pH is decreasing with the permeate flux and salt passage. We attach importance to the solution pH on the membrane performance, especially related to possible changes on the surface charge of membrane. The phenomenon is likely due to the charged membrane and the charged solute which leads to a Donnan potential. The charged membrane attracts ions of opposite charged ion to achieve equilibrium. At the same time, the membrane will repel the same charged ions by an electrostatic force. In addition, the opposite charged ions will also be rejected due to an electroneutrality in the solution. Because of these phenomena, the water can pass through the membrane. Our study find that the charge of the membrane surface is shifted from negative to positive for acidic solutions with pH less than 5.0 that is in accordance with (Ku et al., 2005) and, on the other hand, those differences in electrical charge between the membrane and solute caused concentration polarization phenomenon, blocking the diffusion of solute to coincide with (Akbari et al., 2002).

The ANN's prediction of permeate flux and salt passage is shown as lines in figure.10. Solid symbols are used to identify experimental data points which are used in the training phase of the ANN while open symbols are used for data points not included in the training phase, i.e., points used in the generalization phase. The optimum pH should be explained that the surface charge of a RO membrane is dependent to the ionic strength of the surrounding solution and the best performance of a membrane is expected when its surface charge becomes similar to the electrical charges of the molecules in solution witch obviously may be happened only in a definite pH. According to points not included in the training phase, i.e., of the ANN while open symbols are used for data figure.10, it can be seen that the ANN successfully predicts the nonlinear behavior of permeate flux and salt passage vs pH. It is clear from the figure.10 (a-d) that the permeate flux and salt passage increases with increasing pH and feed flow rate and decreasing concentration of feed. A good prediction is obtained for all concentration and feed flow rate levels with very small deviation (< 2%) between the experimental data and the ANN predictions.

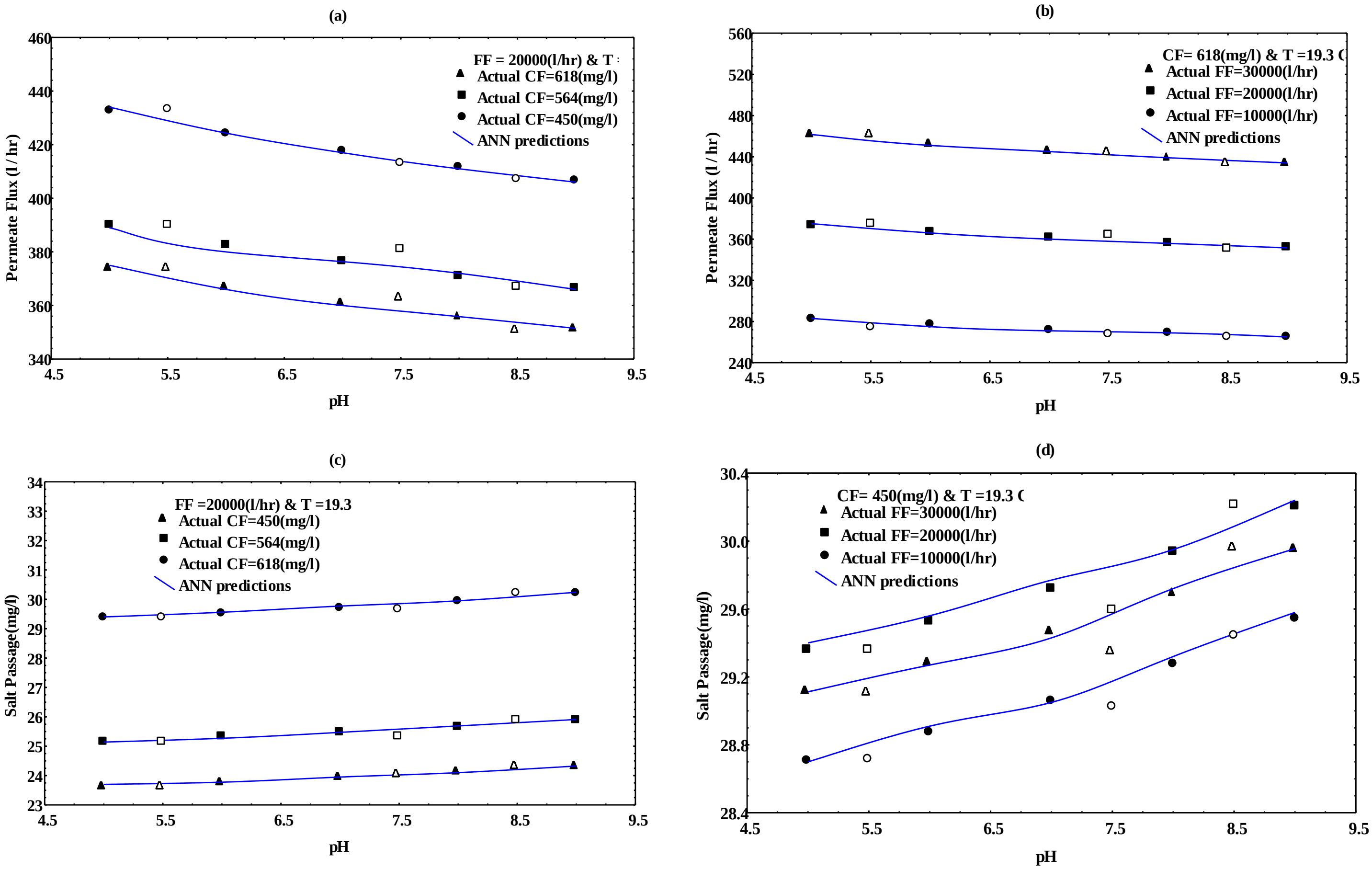


Figure 10. Neural network prediction of the permeate flux and salt passage vs. pH using two effective factors: (a & c) concentrations of feed (CF), (b & d) feed flow rate (FF). Experimental data used in the training process are marked as symbols, whereas lines represent the best fit of the net prediction.

6-3. Effect of temperature on permeate flux and salt passage (ANN predictions)

The effect of temperature on membrane performance is the most important parameter. Figure.11 (a-b) shows that increasing the feed temperature increases the permeate flux. This is attributed to the effect of the temperature of the feed water. As this temperature increases, on one hand, this will decrease the net driving pressure due to an increase in osmotic pressure and, on the other hand, will lead to increasing in water permeability coefficient due to the decrease in both viscosity and density. The later one, will overcomes the effect of net driving pressure thus the permeate flux is increased this is in accordance with (Shamel and Chung, 2006).

The increase pressure and decrease concentration of feed water with increasing feed temperature lead to increase permeate flux as shown in figure.11 (a-b). As the decrease pressure and increase concentration of feed water with increasing feed temperature lead to increase salt passage as shown in figure.11(c-d). The rate of water permeation through the membrane increases as the feed water temperature increases since the viscosity of the solution is reduced and higher diffusion rate of water through the membrane is obtained to coincide with (Sourirajan, 1979). Increasing feed water temperature will yield higher salt passage due to higher diffusion rate for salt through the membrane that is in accord with (Cadotte et al., 1980).

The ANN's prediction of permeate flux and salt passage is shown as lines in figure.11. Solid symbols are used to identify experimental data points which are used in the training phase of the ANN while open symbols are used for data points not included in the training phase, i.e., points used in the generalization phase. According to points not included in the training phase, i.e., of the ANN while open symbols are used for data figure.11, it can be seen that the ANN successfully predicts the nonlinear behaviour of permeate flux and salt passage vs temperature. When temperature of feed water is increased for constant product flow the required applied feed pressure decreases and the product water salinity increases. Energy consumption is decreased as the applied pressure decreases. If the permeate flux is let to increase as the temperature increase fewer membrane elements will be required. Permeate flux and salt passage increase with increasing the feed water temperature. A good prediction is obtained for all concentration and pressure levels with very small deviation (< 2%) between the experimental data and the ANN predictions.

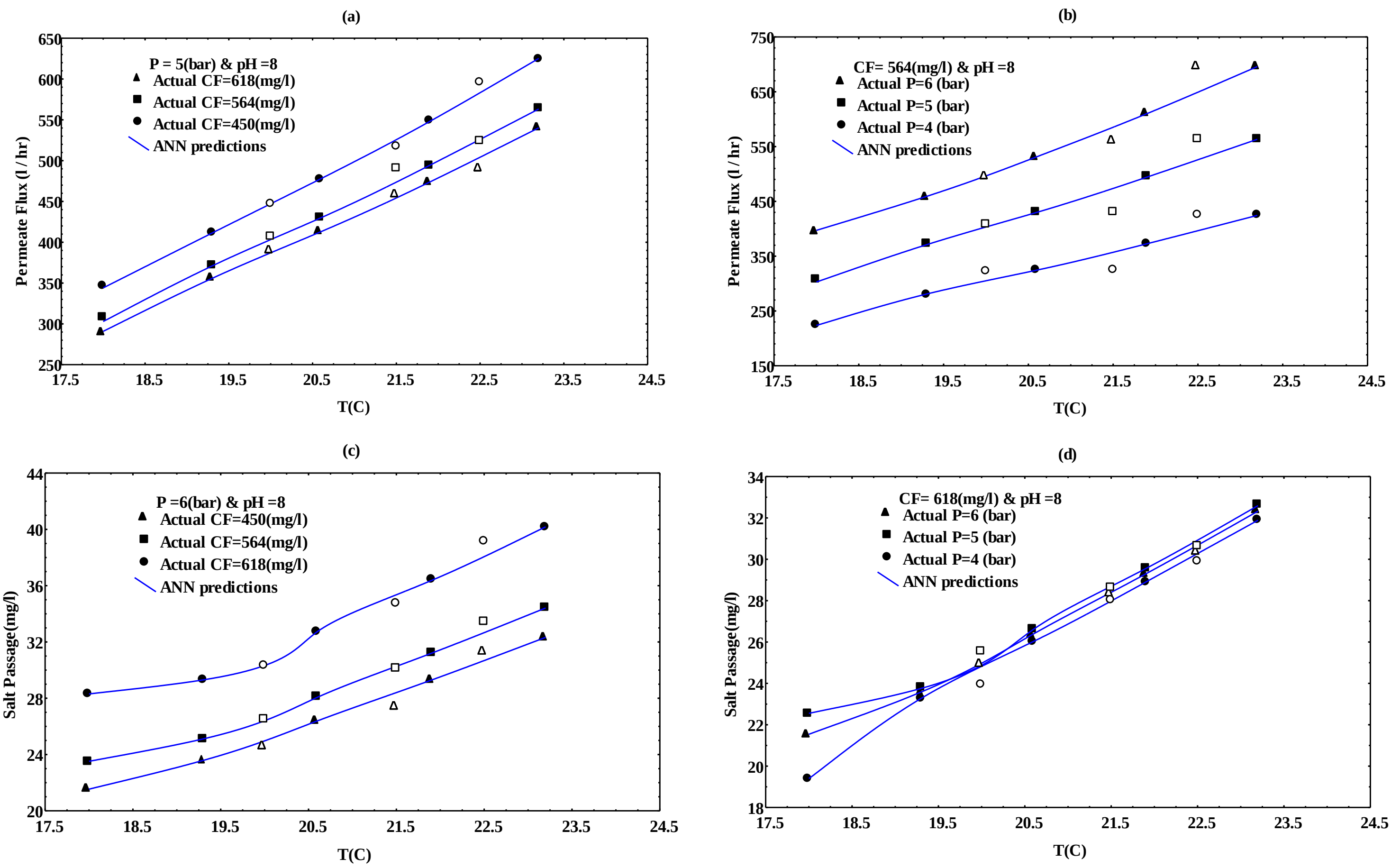


Figure 11. Neural network prediction of the permeate flux and salt passage vs. temperature using two effective factors: (a & c) concentrations of feed (CF), (b & d) feed pressure (P). Experimental data used in the training process are marked as symbols, whereas lines represent the best fit of the net prediction.

6-4. Effect of feed NaCl concentration on permeate flux and salt passage (ANN predictions)

The profiles of permeate flux and salt passage change when the feed concentration change that is in accordance with (Shamel and Chung, 2006) and shown in figure.12 (a-d). When increasing salt concentrations will decrease permeate flux and increase salt passage. This is because the osmotic pressure difference across the membrane increases. Much higher driving force, for the same applied pressure to the feed, is due to the osmotic pressure which is directly related to the salt concentration. The higher feed concentration also leads to surface coating or fouling by salt.

Osmotic pressure is a function of the type and concentration of salts or organics contained in feed water. As salt concentration increases, so does osmotic pressure. The amount of feed water driving pressure necessary to reverse the natural direction of osmotic flow is, therefore, largely determined by the level of salts in the feed water. Temperature also affects permeate flux because increases in temperature result in increases in osmotic pressure and solute and solvent permeability; the increase in solvent permeability results in an increase in permeate flux that is in conformity with (Bhattacharya and Williams, 1992) and shown in figure.12(a). This permeate flux can be often described by Arrhenius temperature dependence on pure water permeability constant. Pure permeate flux change with temperature can also be predicted by water viscosity changes. The permeate flux through the membrane increase as the feed water temperature increases since the viscosity of the solution is reduced and higher diffusion rate of water through the membrane is obtained.

Figure.12(c) demonstrates that, the concentration gradient across the membrane acts as a driving force for the flow of salt through the membrane. As feed concentration increases membrane salt passage increases. Also increasing feed water temperature will yield higher salt passage due to higher diffusion rate for salt through the membrane that is in conformity with (Cadotte et al., 1980).

Increasing the feed flow rate increases the permeate flux and salt passage. This is attributed to the effect of the feed flow rate of the water. As this feed flow rate increases, the salt concentration on the feed-brine side of the membrane increases, on one hand, this will cause an increase in salt passage across the membrane, on the other hand, will lead to increase the osmotic pressure because a higher salt concentration in the feed-brine solution, on the other hand, this will reduce the net driving pressure and consequently reducing the permeate flux and increasing salt passage according to figure.12 (b, d).

The ANN’s prediction of permeate flux and salt passage is shown as lines in figure.12. Solid symbols are used to identify experimental data points which are used in the training phase of the ANN while open symbols are used for data points not included in the training phase, i.e., points used in the generalization phase. According to points not included in the training phase, i.e., of the ANN while open symbols are used for data figure.12, it can be seen that the ANN successfully predicts the nonlinear behaviour of permeate flux and salt passage vs feed NaCl concentration. The maximum feed temperature and feed flow rate of water in any RO system usually depends not on the limiting osmotic pressure, but on the concentration of the salt present in the feed water and their tendency to precipitate on the membrane surface. Chemical treatment of feed water can help in preventing salt precipitation and cause significant increase in permeate flux and salt passage. Good prediction is obtained for all temperature and feed flow rate levels with very small deviation (< 2%) between the experimental data and the ANN predictions.

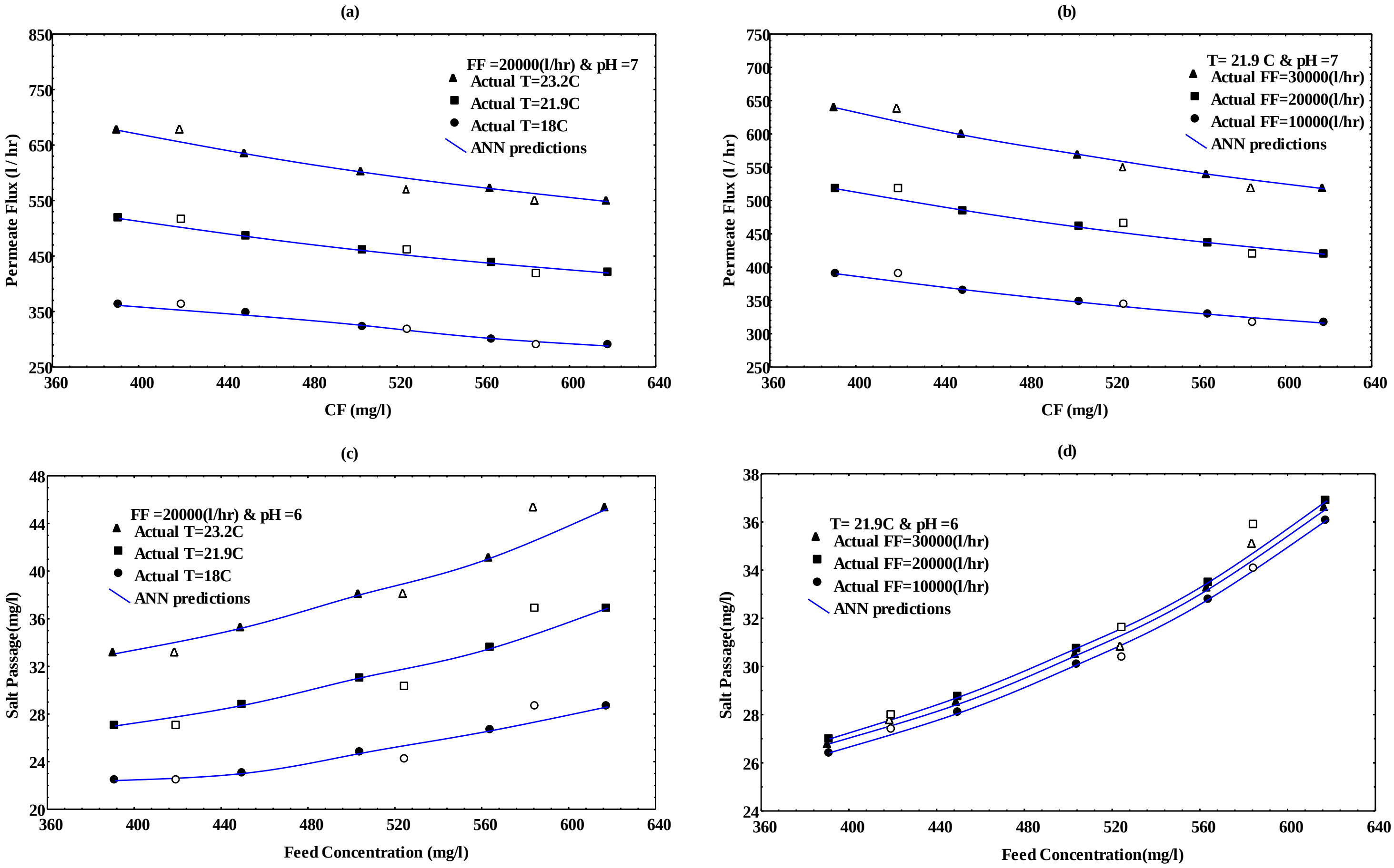


Figure 12. Neural network prediction of the permeate flux and salt passage vs. feed NaCl concentration using two effective factors: (a & c) temperature (T), (b & d) feed flow rate (FF). Experimental data used in the training process are marked as symbols, whereas lines represent the best fit of the net prediction.

6-5. Effect of feed flow rate on permeate flux and salt passage (ANN predictions)

The permeate flux increases and salt passage increases at first and decreases at last when the feed flow rate change are shown in figure.13 (a-d). Feed flow rate of water to the RO is increased and increasing the pH of the feed impacts both the water chemistry and characteristics of the RO membrane. Increasing pH can change the water chemistry by affecting charge, size or solubility of specific constituents in the feed. Increasing pH can also influence the charge of an RO membrane and open the highly crosslinked molecules that form the polyamide structure. These changes in water chemistry and membrane characteristics can influence decreasing permeate flux and increasing salt passage that is in agreement with (Bellona and Drewes, 2005) and shown in figure.13 (b, d).

Identify with (Shamel and Chung, 2006), the feed flow rate increases, when the salt concentration in the feed-brine side of the membrane increases, therefore this will cause an increase in salt passage across the membrane, on one hand, will lead to increase the osmotic pressure because a higher salt concentration in the feed-brine solution, on the other hand, this will reduce the net driving pressure and consequently reducing the permeate flux and increasing salt passage according to figure.13(a, b).

The ANN's prediction of permeate flux and salt passage is shown as lines in figure.13. Solid symbols are used to identify experimental data points which are used in the training phase of the ANN while open symbols are used for data points not included in the training phase, i.e., points used in the generalization phase. According to points not included in the training phase, i.e., of the ANN while open symbols are used for data figure.13, it can be seen that the ANN successfully predicts the nonlinear behaviour of permeate flux and salt passage vs feed flow rate. It is clear from figure.13(a-d) that increasing feed flow rate, the permeate flux increases when low pH and concentration of feed and salt passage increases with high concentration of feed and pH and good prediction is obtained for all concentration and pH levels with very small deviation (< 2%) between the experimental data and the ANN predictions.

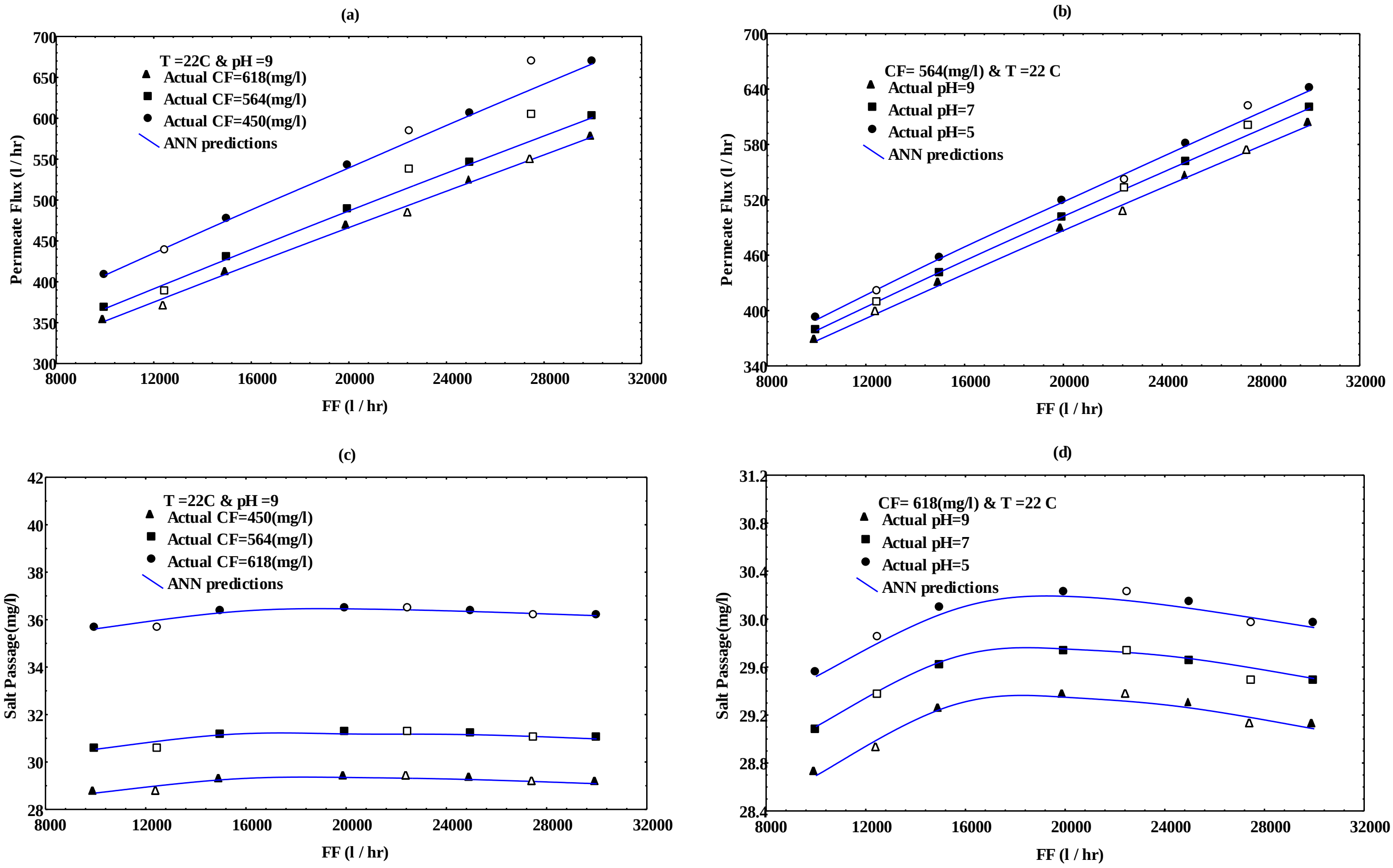


Figure 13. Neural network prediction of the permeate flux and salt passage vs. feed flow rate using two effective factors: (a & c) concentrations of feed (CF), (b & d) pH. Experimental data used in the training process are marked as symbols, whereas lines represent the best fit of the net prediction.

7. CONCLUSION

The performance of reverse osmosis system is studied and analyzed theoretically. Neural networks have very effective in predictive modeling of the performance variables of RO desalination plants and are capable of handling complex and nonlinear problems. A neural network model can handle noisy data more effectively than statistical regression and performs better in predicting desalination plant performance. The accuracy of a neural network model depends on the proper selection of input variables and the broad range of data with which the network is trained. Statistical analysis can aid in the selection of input variables, but the wise engineer will not hesitate to use engineering judgment when it comes to the final decision.

We recommend the following network parameters and functions for modeling a neural network for a (RO) desalination plant:

(1) zero-mean normalization method for input variables.
(2) Gaussian weight-factor distribution for initial values.
(3) Hyperbolic tangent transfer function.
(4) Initial architecture including 30 nodes in hidden layer 1 and 15 nodes in hidden layer 2.
(5) Initial values for both the learning rate and the momentum coefficient of {0.3} and {0.4}, respectively.

Our predictive model by means of ANN can interpolate and predict permeate flux and salt passage with increasing and decreasing of effective variable parameters such as pressure, temperature, salt concentration, feed flow rate and pH of feed water.

REFERENCE

1. Abbas, A. and Al-Bastaki, N., Modeling of a reverse osmosis water desalination unit using neural networks. Chem. Eng. J., 114: 139–143, (2005).
2. Akbari, A., Remigy, J.C., and Aptel, P. Treatment of textile dye effluent using a polyamidebased nanofiltration membrane. Chem. Eng. Proc., 41: 601–609, (2002).
3. Al-Shayji, K. A. and Liu, Y. A., Predictive Modeling of Large-Scale Commercial Water Desalination Plants: Data-Based Neural Network and Model-Based Process Simulation. Ind. Eng. Chem. Res., 41: 6460-6474, (2002).
4. Al-Zoubia, H., Hilala, N., Darwishb, N.A., Mohammadc, A.W., Rejection and modelling of sulphate and potassium salts by nanofiltration membranes: neural network and Spiegler–Kedem model. Desalination 206: 42–60, (2007).
5. Baughman, D. R. and Liu, Y. A., Neural Network in Bioprocessing and Chemical Engineering. Academic Press: San Diego, CA, (1995).
6. Bellona, C. and Drewes, J. The role of membrane surface charge and solute physico-chemical properties in the rejection of organics acids by NF membranes. J. Membr. Sci., 249: 227–234, (2005).
7. Bhagat, P., An Introduction to Neural Nets. Chemical Engineering Progress, 86 (8): 55-60, (1990).
8. Bishop, C.M., Neural Networks for Pattern Recognition. Oxford University Press, (2002).
9. Cadotte, J.E., Petersen, R.J., Larson, R.E. and Erickson, E.E. A new thin-film composite seawater reverse osmosis membrane. Desalination, 32: 25, (1980).
10. Chen, H.Q. and Kim, A.S., 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, (2006).
11. Chitra, S.P., Use Neural Networks for Problem-Solving. Chemical Engineering Progress, 89 (4): 44-52, (1993).
12. Delgrange, N., Cabassud, C., Cabassud, M., Durand-Bourlier, L. and Laine, J.M., Modelling of ultrafiltration fouling by neural network. Desalination 118(1-3): 213-227, (1998).
13. Hagan, M.T. and Menhaj, M.B., Training Feed forward Networks with the Marquardt Algorithm. Ieee Transactions on Neural Networks, 5(6): 989-993, (1994).
14. Hawlader, M., Ho, J. and Chua K., Desalination of Seawater: An Experiment with RO Membrane. Desalination, 132: 275-280, (2000).
15. Hinton, G.E., How Neural Networks Learn from Experience. Scientific American, 267(3): 145-151, (1992).
16. Jamal, K., Khan, M.A. and Kamil, M., Mathematical modeling of reverse osmosis systems. Desalination, 160 (1): 29–42 (2004).
17. Ku, Y., Lee, P.L. and Wang, W.Y. Removal of acidic dyestuffs in aqueous solution by nanofiltration. J. Membr. Sci., 250: 159–165, (2005).
18. Levenberg, K., A Method for the Solution of Certain Non-linear Problems in Least Squares. Quarterly of Applied Mathematics, 2(2): 164-168, (1944).
19. Libotean, D., Giralt, J., Giralt, F., Rallo, R., Wolfec, T. and Cohen, Y., Neural network approach for modeling the performance of reverse osmosis membrane desalting. J. Membr. Sci., 326 (2): 408–419, (2009).
20. Marquardt, D.W., An Algorithm for Least-Squares Estimation of Nonlinear Parameters. Journal of the Society for Industrial and Applied Mathematics, 11(2): 431-441, (1963).
21. Mulder, M., Basic Principle of Technology Membrane. second edition, Netherlands, (1996).

22. Niemi, H., Bulsari, A. and Palosaari, S., Simulation of membrane separation by neural networks. J. Membr. Sci., 120: 185–191, (1995).
23. Rao, G. P., Darwish, D. M., Hassan, A. and Kurdali, A., Toward Improved Automation for Desalination Processes. Part II. Intelligent Control. Desalination, 97: 507, (1994).
24. Sahoo, G.B. and Ray, C., Predicting flux decline in crossflow membranes using artificial neural networks and genetic algorithms. J. Membr. Sci., 283 (1–2): 147–157, (2006).
25. Shamel, M. and Chung, O. T., Drinking Water from Desalination of Seawater: Optimization of Reverse Osmosis System Operating Parameters. J. Eng. Sci. Technol., 1 (2): 203-211, (2006).
26. Sourirajan, S. Pure and Applied Chemistry. 50: 593, (1979).
27. Ujang, Z. and Anderson, G.K. Performance of low pressure reverse osmosis membrane (LPROM) for separating mono- and divalent ions. Wat Sci Tech., 38(4–5): 521–528, (1998).
28. Vapnik, V., The Nature of Statistical Learning Theory. Springer, New York, (1995).