# Predictive Modeling of Large-Scale Commercial Water Desalination Plants: Data-Based Neural Network and Model-Based Process Simulation

**Khawla A. Al-Shayji and Y. A. Liu***

*Honeywell Center of Excellence in Computer-Aided Design, Department of Chemical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*

This paper presents a methodology and practical guidelines for developing predictive models for large-scale commercial water desalination plants by (1) a data-based approach using neural networks based on the backpropagation algorithm and (2) a model-based approach using process simulation with advanced software tools ASPEN PLUS and SPEEDUP and compares the relative merits of the two approaches. This study utilizes actual operating data from two of the largest multistage flash (MSF) and reverse osmosis (RO) desalination plants in the world. Our resulting neural network and process simulation models are capable of accurately predicting the actual operating data from commercial MSF 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 both MSF and RO desalination plants. Our neural network model compares favorably with recent neural network models developed by others in accurately predicting actual operating data from commercial MSF desalination plants. When compared to a data-based neural network, a properly validated model-based process simulation (as in the case of MSF desalination plants) can more effectively quantify the effects of varying operating variables on the desalination performance variables. When it is difficult to develop a model-based process simulation (as in the case of RO desalination plants), we can use a data-based neural network to accurately predict the desalination performance variables.

## 1. Introduction

This paper describes a methodology and practical guidelines of developing predictive models of large-scale commercial water desalination plants by (1) a data-based approach using a neural network based on the backpropagation algorithm and (2) a model-based approach using process simulation with the advanced software tools ASPEN PLUS and SPEEDUP. Specifically, we deal with a multistage flash (MSF) plant [181 760 $m^3$ per day, or 48 million gallons per day (MGPD)] and a reverse osmosis (RO) plant (56 800 $m^3$ per day, or 16 MGPD) located in Kuwait and Saudi Arabia, respectively.

Both neural networks and process simulation 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.[1]

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.

This work represents by far the most comprehensive development and application of neural network models for large-scale desalination plants in the literature. In contrast to several recent studies[2−4] that use simulated plant data, this work utilizes actual plant data. Prior to our study, Husain et al.[5] presented a SPEEDUP program for the steady-state and dynamic simulations of a MSF desalination plant (22 720 $m^3$ per day, or 6 MGPD) in Abu-Dhabi, United Arab Emirates. For our large-scale MSF desalination plant, we present the methodology for developing a steady-state process simulation using ASPEN PLUS and modify the SPEEDUP model of Husain et al.[5] to develop both steady-state and dynamic process simulations. Although previous reports of model-based process simulations of desalination plants are available, our work is unique in that we compare the performance of both data-based neural network and model-based process simulations and evaluate their relative merits when applied to large-scale desalination plants.

## 2. Process Description

**2.1. Multistage Flash (MSF) Desalination Plant.** The AZ-ZOUR multistage flash (MSF) desalination plant is located approximately 100 km south of Kuwait City, Kuwait. Figure 1 illustrates schematically the MSF desalination plant. It consists of three main sections: the brine heater, the heat-recovery section, and the heat-rejection section. The MSF distillation process is a flash distillation process operating in a vacuum, where the vacuum changes from one stage to the next and the flash temperature decreases from the first to the last stage. Following the brine heater, the heat-recovery section consists of 21 stages, where the recycle brine inside the condensing tubes recovers the

* To whom correspondence should be addressed. Phone: (540) 231-7800. Fax: (540) 231-5022. E-mail: Design@Vt.Edu.

10.1021/ie020077r CCC: $22.00 
Published on Web 10/29/2002

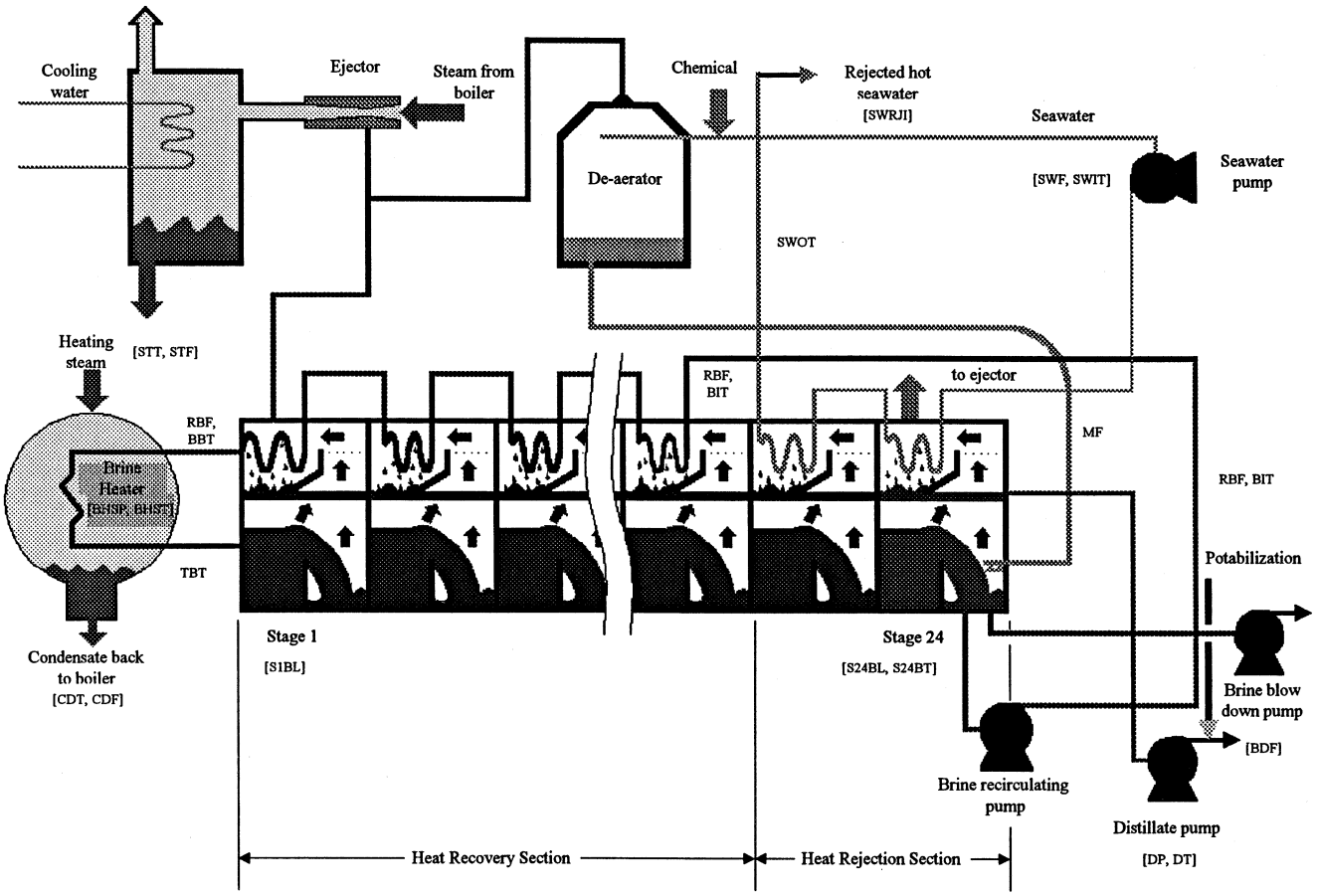


**Figure 1.** Schematic diagram of a multistage flash (MSF) desalination plant.

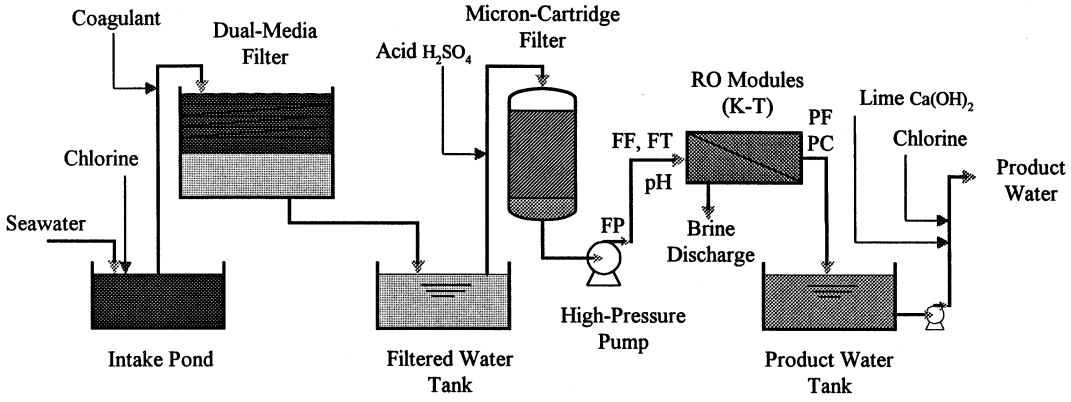


**Figure 2.** Jeddah seawater reverse osmosis plant.

latent heat of vaporization of the flashing brine. The heat-rejection section includes 3 stages to remove the heat added by the brine heater.

**2.2. Jeddah Seawater Reverse Osmosis (RO) Plant.** The Jeddah Seawater RO plant in Saudi Arabia is the largest seawater RO plant in the world. It has a capacity of 56 800 $m^3$ per day or 15 MGPD, and a guaranteed product quality of 625 mg/L in chloride. It has been in operation since March 1994. TOYOBO double-element-type, hollow-fine-fiber RO modules are used for both plants. Figure 2 shows the flowsheet of the RO plant.

## 3. Data-Based Neural Network

**3.1. Methodology of Neural Network Development.** A typical multilayer neural network has an input layer, at least one (normally 1−3) hidden layer(s) and an output layer (see Figure 3). The text by Baughman and Liu[6] describes the fundamentals and applications of neural networks in bioprocessing and chemical engineering, including the backpropagation algorithm for the development of multilayer feedforward networks.

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

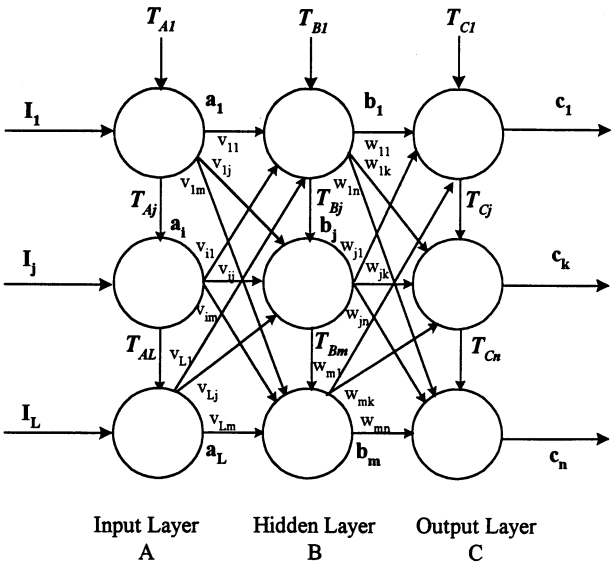


**Figure 3.** Three-layer feedforward neural network.

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. We used NeuralWorks Professional II/Plus (NeuralWare) and SAS (SAS Institute, Inc.) software tools to accomplish this work.

**3.2. Data Collection, Preparation, and Analysis.** To collect relevant data from the operational trial of the plant, it is desirable to carry out the modeling study on a normally operating desalination unit that is relatively free of significant scale deposition and equipped with plant instrumentation.

There are two categories of unit instrumentation. The first consists of all instrumentation located in the distillation control room (DCR). The second consists of instrumentation sited locally throughout the unit. We used data from local instruments to double-check data levels from the DCR. For consistency, we relied only on the data generated by the DCR instrumentation.

In this study, we utilized the operating variables from the MSF desalination plant measured every 2 h for 1 month during both winter and summer operations (300 data sets for each period). We recorded the operating variables from the RO plant every hour for a period of 1 month (568 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. Figure 4 gives an example of the potential influence of outliers on the network performance for predicting the top brine temperature (TBT). We specify the details of

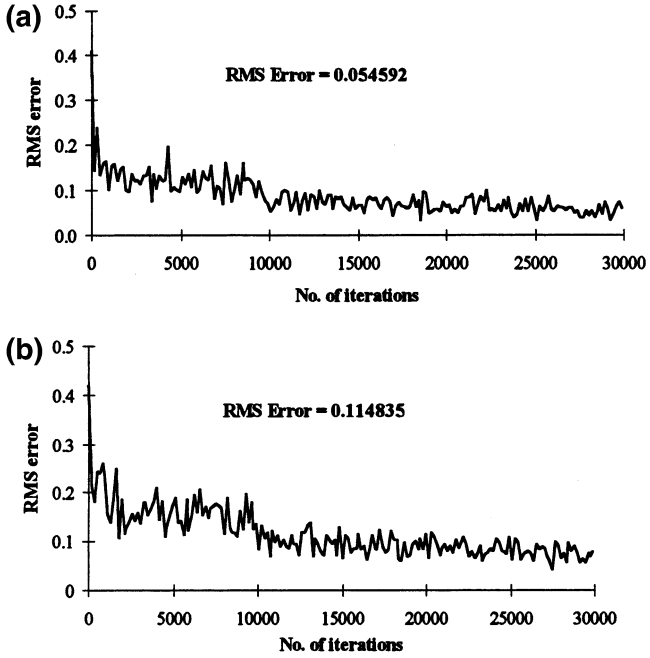


**Figure 4.** RMS error of the top brine temperatue (TBP) network: (a) without outliers and (b) with outliers.

the TBT network later in Tables 3 and 4, after discussing the network development. The root-mean-square (RMS) error decreased from 0.114 835 to 0.054 592 in the TBT network after the outliers had been removed. Removing outliers generally improves network performance.

**3.3. Input-Variable Selection.** The main difficulty in the structural identification of a complex nonlinear system arises from the huge amount of possible relationships among variables. The selection of outputs is straightforward and depends on the modeling goal. However, informed input-variable selection is critical to achieving efficient model performance. We demonstrate how to select input variables of neural networks on the basis of both statistical analysis and engineering knowhow and compare the resulting errors of output predictions.

**3.3.1. Factor Analysis: Principal Component Analysis.** Utojo and Bakashi[7] provide an excellent comparison of multivariate statistical methods and neural networks for data processing. Factor analysis is a technique of multivariate analysis that attempts to account for the covariance among a set of observable random variables (denoted by X) in terms of a minimum number of unobservable or latent random variables called factors. These unobserved factors are assumed to be linear combinations of the variables, which make up the set X. Thus, the objective becomes one of reducing the complexity of the set X into as few linear combinations of the variables within X as possible. There are numerous strategies for performing this reduction of the set X. One such approach is principal component analysis (PCA), which reduces the complexity of the set X via a canonical analysis of the correlation matrix of X. The dominant eigenvectors of the matrix X are then taken to be the principal factors of X. The elements comprising the eigenvectors are then taken to be the weights, which produce a linear combination of the set of variables within X. For instance, if we denote the first factor as $F_1$, then $F_1$ is simply a linear combination of the variables in X, where the weights are determined

**Table 1. Operating Variables, Their Nomenclatures in the MSF Desalination Plant, and Their Formats in the Training Data Files as Input and Output Variables in the Prediction Networks for Top Brine Temperature (TBP), Distillate Produced (DP), and Steam Flow Rate (STF)**

| column no. | variable name | nomenclature | variable type for TBT network | variable type for DP network | variable type for STF network |
|---|---|---|---|---|---|
| 1 | seawater flow rate | SWF | input | input | input |
| 2 | makeup flow rate | MF | input | input | input |
| 3 | seawater recirculating flow rate | SWRF | input | input | input |
| 4 | seawater inlet temperature | SWIT | input | input | input |
| 5 | seawater outlet temperature | SWOT | input | input | input |
| 6 | blow-down flow rate | BDF | input | input | input |
| 7 | brine inlet temperature | BIT | input | input | input |
| 8 | stage 24 brine temperature | S24BT | input | input | input |
| 9 | brine heater inlet temperature | BBT | input | input | input |
| 10 | stage 1 brine level | S1BL | input | input | input |
| 11 | brine heater shell pressure | BHSP | input | input | input |
| 12 | brine heater shell temperature | BHST | input | input | input |
| 13 | steam temperature | STT | input | input | input |
| 14 | condensate temperature | CDT | input | input | input |
| 15 | condensate flow rate | CDF | input | input | input |
| 16 | recirculating brine flow rate | RBF | input | input | input |
| 17 | LP steam flow rate to brine heater | STF | input | input | output |
| 18 | distillate produced | DP | input | output | input |
| 19 | top brine temperature | TBT | output | input | input |

by the elements of the first (most dominant) eigenvector of the correlation matrix of X

$$F_1 = e_{11}x_1 + e_{21}x_2 + ... + e_{p1}x_p \qquad (1)$$

where $e_1 = (e_{11}\ e_{21}\ ...\ e_{p1})$ denotes the most dominant eigenvector of the correlation matrix of X. The elements of $e_1$ are known as the factor loadings for each of the $p$ variables that comprise X. These factor loadings are always between −1.0 and 1.0, and a useful heuristic is that variables whose factor loadings have absolute values greater than 0.5 are related highly to the corresponding factor.

The rotated PCA generally involves the following major steps: (a) selection of the variables, (b) computation of the matrix of correlations among the variables, (c) extraction of the unrotated factors, (d) rotation of the factors, and (e) interpretataion of the rotated factor matrix.

Current estimation and rotation methods require iterative calculations that must be done on a computer. Several computer programs are now available for this purpose. We use the SAS statistical software to carry out the factor analysis.

Ettouney et al.[8] provided a good description of the MSF desalination variables in a recent article. Table 1 summarizes 19 important operating variables from those collected in the distillation control room, of which the primary performance variables of desalination plants are the top brine temperature (TBT), distillate produced (DP), and steam flow rate (STF). We apply the factor analysis to test 19 variables with 4 factors. Table 2 summarizes the resulting rotated factor loadings.

According to the preceding heuristic, the variables whose loading values are greater than 0.5 for a particular factor are taken to represent that factor. Factor 1 ($F_1$) has high loading values for variables S24BT (stage 24 brine temperature), SWOT (seawater outlet temperature), SWF (seawater flow rate), DP (distillate produced), BIT (brine inlet temperature), CDF (condensate flow rate), BDF (blowdown flow rate), and MF (makeup flow rate) and much lower values for all the rest. Factor 2 ($F_2$) has high loadings for variables CDT (condensate temperature), TBT (top brine temperature), BHSP (brine heater shell pressure), BBT (brine heater inlet temperature), BHST (brine heater shell temperature), S1BL (stage 1 brine level), and STF (steam flow rate), with much lower loadings for all other variables. Finally, factor 3 ($F_3$) has high loadings only for variables SWRF (seawater recalculating flow rate), BDF (blowdown flow rate), STF (steam flow rate), and MF (makeup flow rate). Each factor in Table 2, then, is related highly to only a few variables (indicated by values in bold), and the set of related variables differs for each factor.

**Table 2. Estimated Rotated Factor Loading for 19 Operating Variables[a]**

| operating variable | $F_1$ | $F_2$ | $F_3$ | $F_4$ |
|---|---|---|---|---|
| SWF | 0.908 78 | −0.311 66 | 0.047 98 | 0.148 92 |
| SWRF | 0.295 74 | −0.230 82 | 0.855 93 | 0.032 45 |
| MF | 0.612 70 | 0.385 89 | −0.521 36 | 0.070 86 |
| RBF | −0.494 49 | 0.022 60 | −0.051 17 | −0.642 33 |
| BDF | 0.737 23 | 0.029 92 | −0.636 21 | 0.003 86 |
| CDF | 0.737 76 | 0.191 64 | −0.096 35 | 0.406 17 |
| SWIT | −0.378 05 | −0.151 63 | −0.365 83 | 0.497 78 |
| SWOT | −0.947 60 | 0.266 69 | −0.123 96 | −0.037 43 |
| BIT | −0.771 28 | 0.481 12 | −0.270 91 | 0.075 92 |
| BBT | −0.372 88 | 0.867 03 | 0.122 81 | −0.155 22 |
| S24BT | −0.960 54 | 0.093 31 | −0.091 21 | 0.045 16 |
| STT | −0.065 84 | 0.106 95 | 0.053 27 | 0.788 44 |
| CDT | −0.043 27 | 0.959 70 | −0.094 53 | 0.054 09 |
| BHSP | −0.051 89 | 0.946 65 | −0.021 83 | 0.016 74 |
| BHST | 0.164 72 | 0.738 29 | −0.211 71 | 0.087 23 |
| S1BL | −0.102 77 | −0.717 61 | 0.024 79 | −0.525 69 |
| TBT | −0.114 81 | **0.957 00** | 0.050 19 | 0.014 41 |
| DP | **0.859 05** | 0.226 51 | −0.162 89 | −0.257 49 |
| STF | −0.172 05 | 0.548 27 | **0.596 06** | 0.014 70 |

[a] 300 observations.

To determine the relationship between significant input variables and output variables, we must match each output variable to a factor column. We examine the output variable's row of loading values and then identify the largest loading value in that row, and the column in which this value is located indicates the appropriate factor match for that output variable. We repeat this process for each of the output variables, and if there are fewer output variables than columns, we discard the extraneous columns.

As seen in Table 2, marked in bold, we find $F_1$, $F_2$, and $F_3$ to be the DP (distillate produced), TBT (top brine temperature), and STF (steam flow rate) variables, respectively.

**Table 3. Input Variables Recommended for the MSF Network by Engineering Knowhow and by Factor Analysis**

| | factor analysis | | |
|---|---|---|---|
| engineering knowhow | top brine temperature | distillate produced | steam flow rate |
| 1. seawater flow rate | 1. condensate temperature | 1. stage 24 brine temperature | 1. seawater recirculating flow rate |
| 2. seawater recirculating flow rate | 2. brine heater shell pressure | 2. seawater outlet temperature | 2. steam temperature |
| 3. makeup flow rate | 3. brine heater inlet temperature | 3. seawater flow rate | 3. recirculating brine flow rate |
| 4. recirculating brine flow rate | 4. steam temperature | 4. steam temperature | 4. blowdown flow rate |
| 5. blowdown flow rate | 5. brine heater shell temperature | 5. brine inlet temperature | 5. stage 1 brine level |
| 6. condensate flow rate | 6. stage 1 brine level | 6. condensate flow rate | 6. makeup flow rate |
| 7.seawater inlet temperature | 7. recirculating brine flow rate | 7. blowdown flow rate | 7. seawater inlet temperature |
| 8.seawater outlet temperature | 8. steam flow rate | 8. recirculating brine flow rate | |
| 9. brine inlet temperature | 9. seawater inlet temperature | 9. makeup flow rate | |
| 10. brine heater inlet temperature | | 10. stage 1 brine level | |
| 11. stage 24 brine temperature | | 11. seawater inlet temperature | |
| 12. steam temperature | | | |
| 13. condensate temperature | | | |
| 14. brine heater shell pressure | | | |
| 15. brine heater shell temperature | | | |
| 16. stage 1 brine level | | | |
| 17. top brine temperature | | | |
| 18. distillate produced | | | |
| 19. steam flow rate | | | |

**3.3.2. Engineering Knowhow.** In using statistical analysis, there is always the risk that significant input variables might be excluded if one does not utilize the particular functional relationship in that testing method. Our emphasis is on selecting an appropriate subset of these variables for use in the final prediction model. Therefore, we investigate various specifications of the input variables, based on the plant design and engineering knowhow, and retain those that are deemed worthy of further study. On the basis of a detailed analysis of operating variables and their constraints presented in Al-Shayji,[9] we decided to include 19 operating variables as input variables in training the MSF network, as listed in Table 1.

Table 3 specifies the details of the top brine temperature (TBP) prediction network, and Table 4 gives the format of data files used to training the TBP network. Figure 5a and 5b compares the performance of the TBP network with input variables selected by engineering knowhow and by factor analysis. It is clear that the network with input variables selected by engineering knowhow performs better than that with variables selected by factor analysis. Table 5 shows that the RMS errors for the engineering knowhow network are much lower in value and fluctuation than those for the factor analysis network

We conclude that the degree of success in input-variable selection greatly influences the resulting network's predictive ability. Statistical methods can aid in the process of variable selection, but the wise engineer will not hesitate to use his engineering knowhow when it comes down to the final decisions.

**3.4. Training and Testing Sets.** We use two subsets of data to build a model: a training set and a testing 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.

To effectively visualize how well a network performs recall and generalization, we often generate a learning curve, which represents the average error for both the recall of training data sets and the generalization of the testing data sets as a function of the number of examples in the training data set.

**Table 4. Specifications of the Prediction Network for Top Brine Temperature, Distillate Produced, and Steam Flow Rate**

| | | | |
|---|---|---|---|
| network type | backpropagation | | |
| training file names | WMIMO.nna, WTBT.nna, WDP.nna, and WSTF.nna | | |
| transfer function (input layer) | linear | | |
| transfer function (hidden layer) | tanh | | |
| transfer function (output layer) | tanh | | |
| learning rule | delta rule | | |
| summation | sum | | |
| error | standard | | |
| network weight distribution | normal distribution, $3\sigma$ limits of [−1,1] | | |
| **Input Layer** | | | |
| training iteration | 5000 | | |
| noise | 0.0 | | |
| learning rate | 0.9 | | |
| momentum coefficient | 0.6 | | |
| error tolerance | 0.0 | | |
| **Hidden Layer 1** | | | |
| training iterations | 10 000 | 30 000 | 70 000 |
| noise | 0.0 | 0.0 | 0.0 |
| learning rate | 0.3 | 0.15 | 0.375 |
| momentum coefficient | 0.4 | 0.2 | 0.05 |
| error tolerance | 0.1 | 0.1 | 0.1 |
| **Hidden Layer 2** | | | |
| training iterations | 10 000 | 30 000 | 70 000 |
| noise | 0.0 | 0.0 | 0.0 |
| learning rate | 0.25 | 0.125 | 0.031 25 |
| momentum coefficient | 0.4 | 0.2 | 0.05 |
| error tolerance | 0.1 | 0.1 | 0.1 |
| **Output Layer** | | | |
| training iterations | 10 000 | 30 000 | 70 000 |
| noise | 0.0 | 0.0 | 0.0 |
| learning rate | 0.15 | 0.075 | 0.018 75 |
| momentum coefficient | 0.4 | 0.2 | 0.05 |
| error tolerance | 0.1 | 0.1 | 0.1 |

Figure 6 shows a learning curve for training the top brine temperature (TBP) network. We can use a learning curve (1) to find the number of training examples required to achieve a fixed average error and (2) to estimate the minimum average error that can be obtained by adding data sets. This figure displays the lowest average recall and generalization errors below 0.1 with 150 or more training examples. Therefore, we

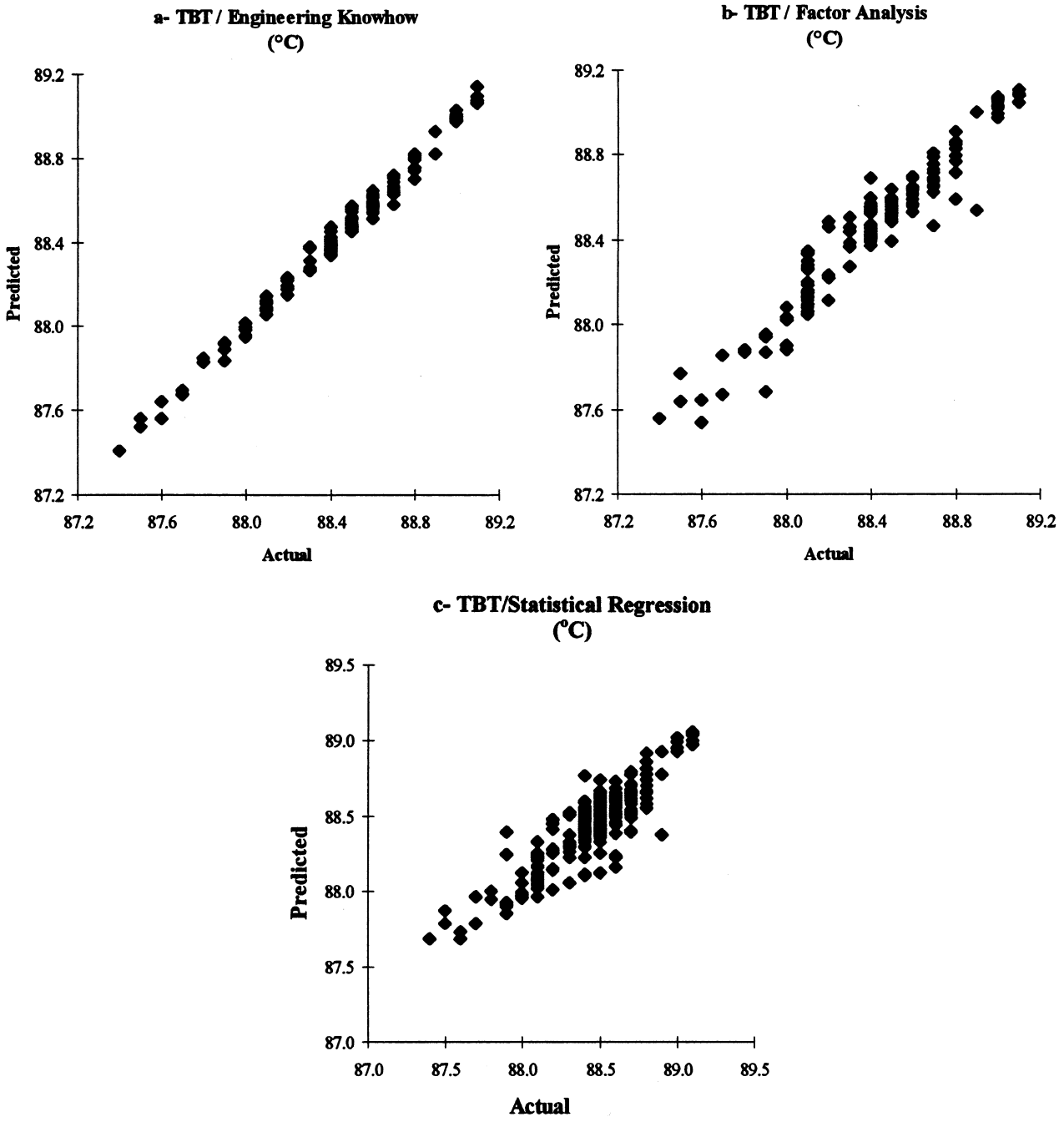


**Figure 5.** Actual and predicted top brine temperature for the MSF network by (a) neural network with input variables selected by engineering knowhow, (b) neural network with input variables selected by factor analysis, and (c) statistical regression.

**Table 5. RMS Errors of the Top Brine Temperature (TBT), Distillate Produced (DP), and Steam Flow Rate (STF) from Neural Networks Using Different Approaches to Input-Variable Selection and from Statistical Regression**

| | TBT (°C) | DP (ton/h) | STF (ton/h) |
|---|---|---|---|
| neural network | | | |
| 1. engineering knowhow | 0.0546 | 0.0621 | 0.0679 |
| 2. factor analysis | 0.0986 | 0.1425 | 0.1705 |
| statistical regression | 0.1286 | 11.6521 | 2.9963 |

use a training data set and a testing data set of 150 samples each throughout this study.

**3.5. Model Development and Optimization. 3.5.1. Selection of Key Network Parameters.** Figure 7 shows the steps in determining key parameters that control the prediction ability of a multilayer feed-forward network for the modeling and optimization of commercial desalination plants. Our focus here is to

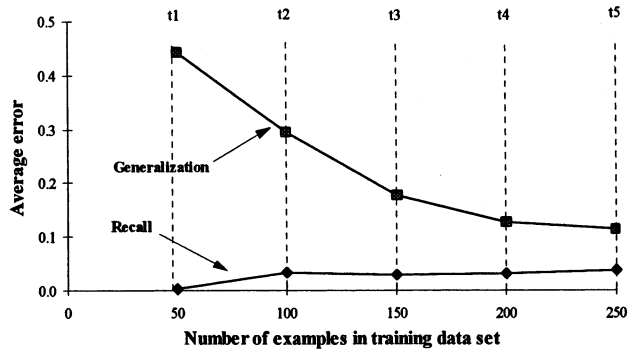


**Figure 6.** Learning curve for training the top brine temperature network.

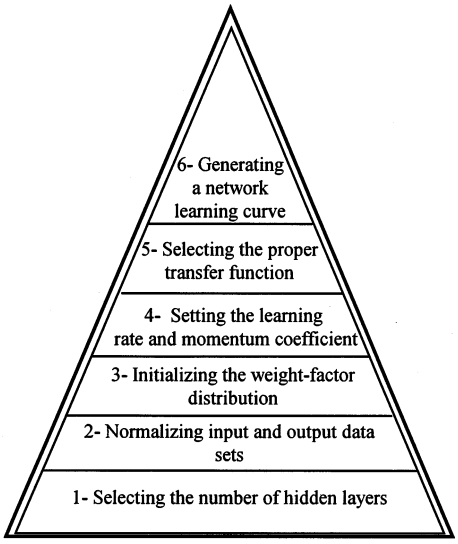


**Figure 7.** Steps to determine the key parameters controlling the prediction ability of a multilayer feedforward network.

present the practical guidelines for developing and optimizing neural network models for commercial desalination processes. Importantly, these guidelines represent effective starting points for neural network modeling of other large-scale commercial processes.

We first initialize the weight factors between any two nodes within the network. Because no prior information about the system being modeled is available, we prefer to set all of the free weight factors of the network to random numbers that are normally distributed (Gaussian weight-factor distribution) inside a zero-mean range of values, say, between −0.5 and +0.5.

A multilayer prediction network trained with the backpropagation 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.

**3.5.2. Number of Nodes in the Hidden Layer(s).** The number of input and output nodes corresponds to the number of inputs into the network and the number of desired outputs of the network, respectively. The choice of the number of nodes in the hidden layer(s) depends on the network application. Although using a single hidden layer is sufficient in solving many functional approximation problems, some problems might be easier to solve with a two-hidden-layer configuration.

Consider the prediction network for the top brine temperature (TBT) specified in Tables 3 and 4. The network consists of 18 input variables and 1 output variable. It uses a backpropagation 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 150 data sets to train these configurations with 60 000 iterations. We tested the network with one- and two-hidden-layer configurations with an increasing number of nodes in each hidden layer(s).

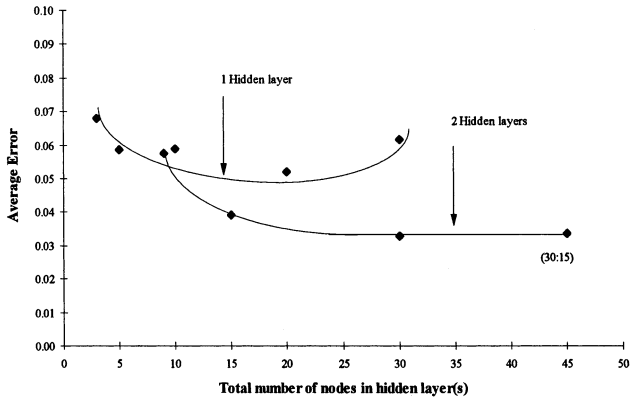


**Figure 8.** Comparison of the average errors for the prediction network for the top brine temperature (TBT) trained with various one- and two-hidden-layer configurations.

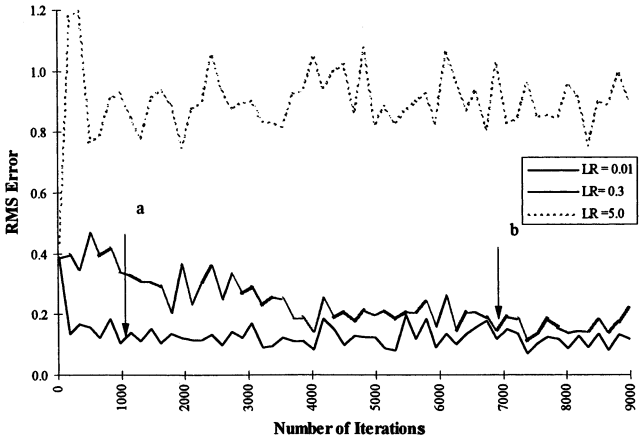


**Figure 9.** Comparison of the RMS error in training the TBT network with different learning rates (LRs).

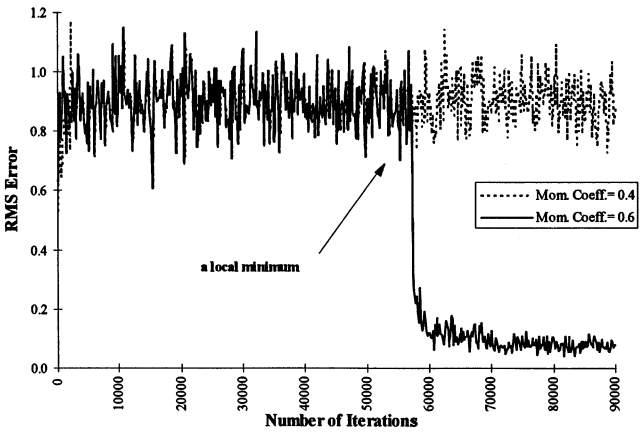


**Figure 10.** Effect of introducing a momentum coefficient on the local minimum.

Figure 8 shows that the two-hidden-layer network performs significantly better than the one-hidden-layer network. The optimal two-hidden-layer network configuration with a minimum average error is 30:15 (i.e., with 30 nodes in hidden layer 1 and 15 nodes in hidden layer 2).

**3.5.3. Effects of Learning Rate and Momentum Coefficients.** The learning rate is a positive parameter that regulates the relative magnitude of weight changes during learning. 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.

**Table 6. Comparison of Input and Output Variables for the Present Work and Recent Studies**

| | 1. This work | | 2. Selverajs and Despande[4] | | 3. Abdulbary et al.[2] | | 4. Parentis et al.[3] | |
|---|---|---|---|---|---|---|---|---|
| no. | input variables (16) | outpu variables (3) | input variables (2) | output variables (3) | input variables (4) | output variables (3) | input variables (4) | output variables (8) |
| 1 | seawater flow rate | top brine temperature | recirculating flow rate | top brine temperature | top brine temperature | recirculating brine salinity | recirculating brine flow rate | steam flow rate |
| 2 | makeup flow rate | distillate produced | steam flow rate | distillate produced | recirculating brine flow rate | distillate produced | recirculating brine temperature | makeup flow rate |
| 3 | seawater recycle flow | steam flow rate | | stage 1 brine level | makeup flow rate | performance ratio | recirculating brine salinity | distillate produced |
| 4 | seawater inlet temperature | | | | steam temperature | | seawater inlet temperature | blowdown flow rate |
| 5 | seawater outlet temperature | | | | | | | gain output ratio |
| 6 | blowdown flow rate | | | | | | | evaporated brine outflow |
| 7 | brine inlet temperature | | | | | | | evaporated brine output salinity |
| 8 | stage 24 brine temperature | | | | | | | evaporated brine output temperature |
| 9 | brine heater inlet temperature | | | | | | | |
| 10 | stage 1 brine level | | | | | | | |
| 11 | brine heater shell temperature | | | | | | | |
| 12 | brine heater shell pressure | | | | | | | |
| 13 | steam temperature | | | | | | | |
| 14 | condensate temperature | | | | | | | |
| 15 | condensate flow rate | | | | | | | |
| 16 | recirculating brine flow rate | | | | | | | |

Consider again the prediction network for the top brine temperature (TBT) specified in Tables 3 and 4, except for the values of learning rate and momentum coefficient used, as discussed below.

Figure 9 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 lower learning rate results in slower convergence. When the learning rate is low (0.01), the network takes a longer (4000 iterations) to reach an RMS error of 0.16 (point b in Figure 9). This follows because the lower the learning rate, the smaller the changes to the weights in the network 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 0.16 (point a in Figure 9) in a shorter time (200 iterations). If we increase the learning rate to 5.0, so that larger steps are taken, the algorithm becomes unstable; the oscillating error fluctuations increase instead of decaying and, thus, do not reach a minimum.

Therefore, to avoid the danger of instability and improve convergence as the learning rate is increased, we introduced a momentum coefficient to smooth out the oscillation. In backpropagation with momentum, the weight changes in a direction that is a combination of the current gradient and the previous gradient. This helps in moving the minimization routine, if during training, it is trapped in a local minimum, as Figure 10 illustrates. To prevent excessive weight changes and possible oscillation, the algorithm slows the weight changes by a term that is proportional to the previous weight change and the momentum coefficient.

In general, we should assign a lower learning rate in the last layers than in the front-end layers, because the last layers tend to have larger local gradients than the layers at the front end of the network. In contrast, we should keep the momentum coefficient constant between layers and decrease it with increasing in the number of iterations. The training schedule for the TBP network summarized in Table 4 represents a recommended schedule for training a two-hidden-layer backpropagation network used for modeling large-scale commercial processes.

**Table 7. Average Recall, Generalization, and RMS Errors of the Top Brine Temperature (TBP), Distillate Produced (DP), and Stage 1 Brine Level (S1BL) of the Model by Selvarajs and Deshpande[4] and Our Model**

| error | TBT (°C) | DP (ton/h) | S1BL (cm) |
|---|---|---|---|
| | Model 1 | | |
| recall | 0.2397 | 9.5117 | 12.3059 |
| generalization | 0.1886 | 20.5378 | 14.8734 |
| RMS | 0.3054 | 0.2222 | 0.2055 |
| | Model 2 | | |
| recall | 0.2400 | 9.5157 | 12.2910 |
| generalization | 0.1892 | 20.5340 | 14.7740 |
| RMS | 0.2985 | 0.2164 | 0.2076 |
| | Model 3 | | |
| recall | 0.0769 | 6.8592 | 12.7257 |
| generalization | 0.1202 | 18.1674 | 12.8100 |
| RMS | 0.1311 | 0.1942 | 0.2563 |
| | Our Model | | |
| recall | 0.0287 | 0.9329 | [a] |
| generalization | 0.1764 | 13.0878 | |
| RMS | 0.0546 | 0.0621 | |

[a] Not considered as an output.

**3.6. Neural Network versus Statistical Regression.** Figure 5a−c compares the predicted and actual top brine temperature from the neural network (specified in Tables 3 and 4) and from the statistical regression. Table 5 summarizes the corresponding RMS errors

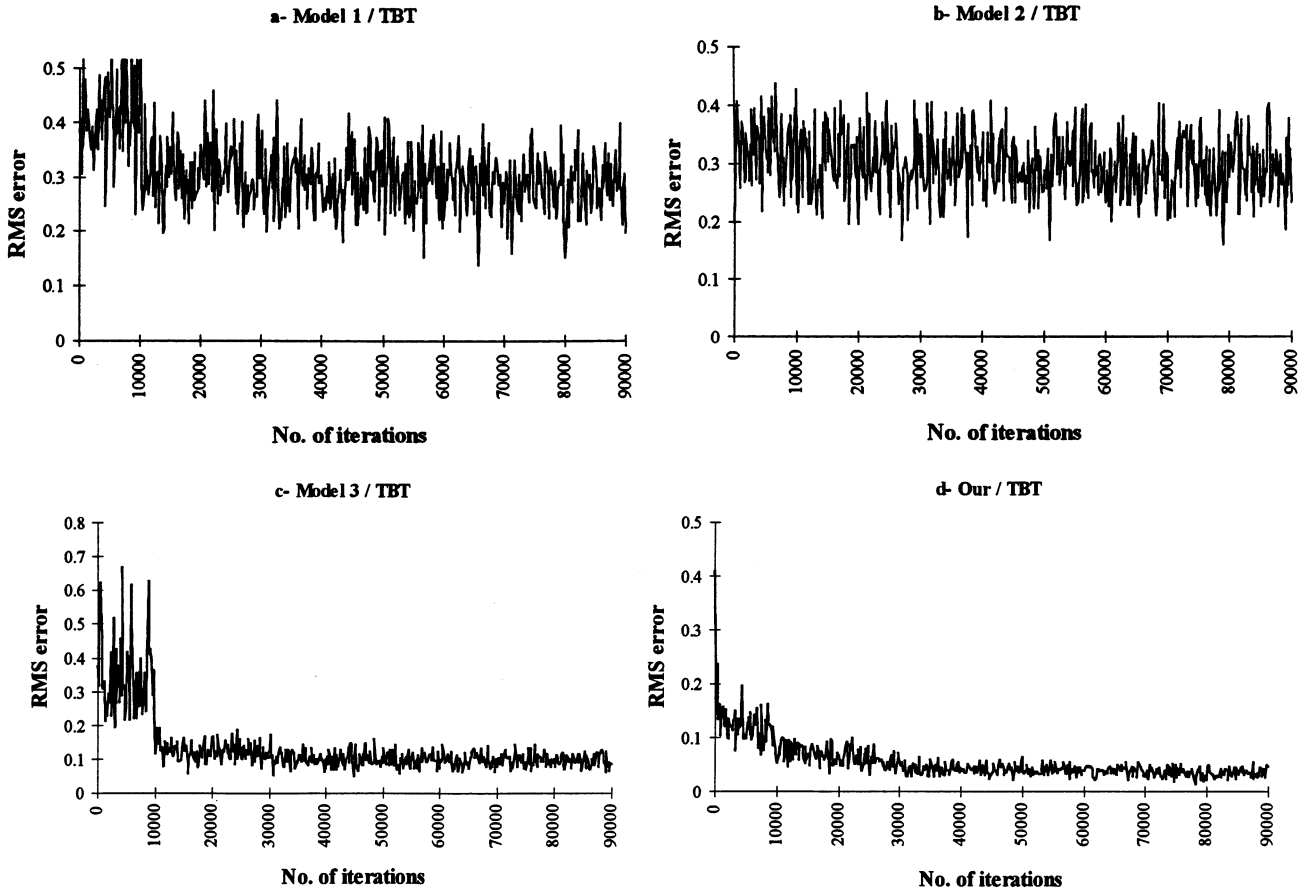


**Figure 11.** (a−d) Comparison actual and predicted distallate produced of our model and three versions of the model by Selvarajs and Despande[4] using actual AZ-ZOUR MSF desalination of plant data.

of the top brine temperature, distillate produced, and steam flow rate from both approaches. The results show that the neural network outperforms statistical regression, especially for noisy data such as distillate produced (DP) and steam flow rate (STF). The RMS errors for TBT (top brine temperature), DP, and STF in the statistical regression model are 57.5, 99.5, and 97.7%, respectively, larger than those of neural networks. After fitting the regression model to the given data, we evaluate its adequacy of fit to the data. The most widely used measure is the square of the multiple correlation coefficient $R^2$, which has a range between 0 and 1. When the model fits the data well, the value of the $R^2$ is close to unity. With a good fit, the observed and predicted values are close to each other and the residual is small. The $R^2$ values for the TBT, DP, and STF models are 0.767 846, 0.433 603, and 0.184 604, respectively.

We conclude that neural networks have been very effective in predicting the performance variables of the MSF desalination plants and are capable of handling complex and nonlinear problems. They also outperform the regression models in prediction problems.

**3.7. Comparison with Recent Neural Network Models.** Table 6 compares the network architectures between recent studies (with all using simulated plant data) and the present work (using actual plant data). We compare our model with that of Selvarajs and Deshpande (SD)[4] in three different ways using actual operating data for winter and summer from the AZ-ZOUR MSF desalination plant, as follows: (1) Model 1 includes SD's variables + SD's network architecture + actual plant data. (2) Model 2 includes SD's variables + our network architecture + actual plant data. (3) Model 3 includes Our variables + SD's network architecture + actual plant data. The model responses with summer and winter data are essentially identical. We will limit our discussion of the three models to winter data.

Table 7 compares the recall, generalization, and RMS errors of TBT (top brine temperature), DP (distillate produced), and S1BL (stage 1 brine level) networks of our model with those of the other three models. Only the generalization error for model 2 for TBP (0.1202) is slightly smaller that that for our model (0.1764). For all 35 other data points, our model performs better in terms of the recall error, generalization error, and RMS error.

Models 1 and 2 do not show a significant difference in their error values. Their RMS errors are trapped in local minima, unstable and fluctuating. The inefficient network performance in models 1 and 2 could be attributed to the type of data used in their model (simulated data) and the variable selected. Model 3 shows a better improvement in the network performance than models 1 and 2. Figure 11 illustrates the actual and predicted distillate produced and the RMS errors from the above models.

This comparison further supports our key observation that the accuracy of a neural network model depends on the proper selection of input variables. The input variables recommended by recent papers are simply insufficient for accurate predictive modeling of actual operating data from large-scale MSF desalination plants.

**3.8. Application of the Neural Network Modeling Methodology to RO Plant.** We have successfully applied the preceding modeling methodology to the

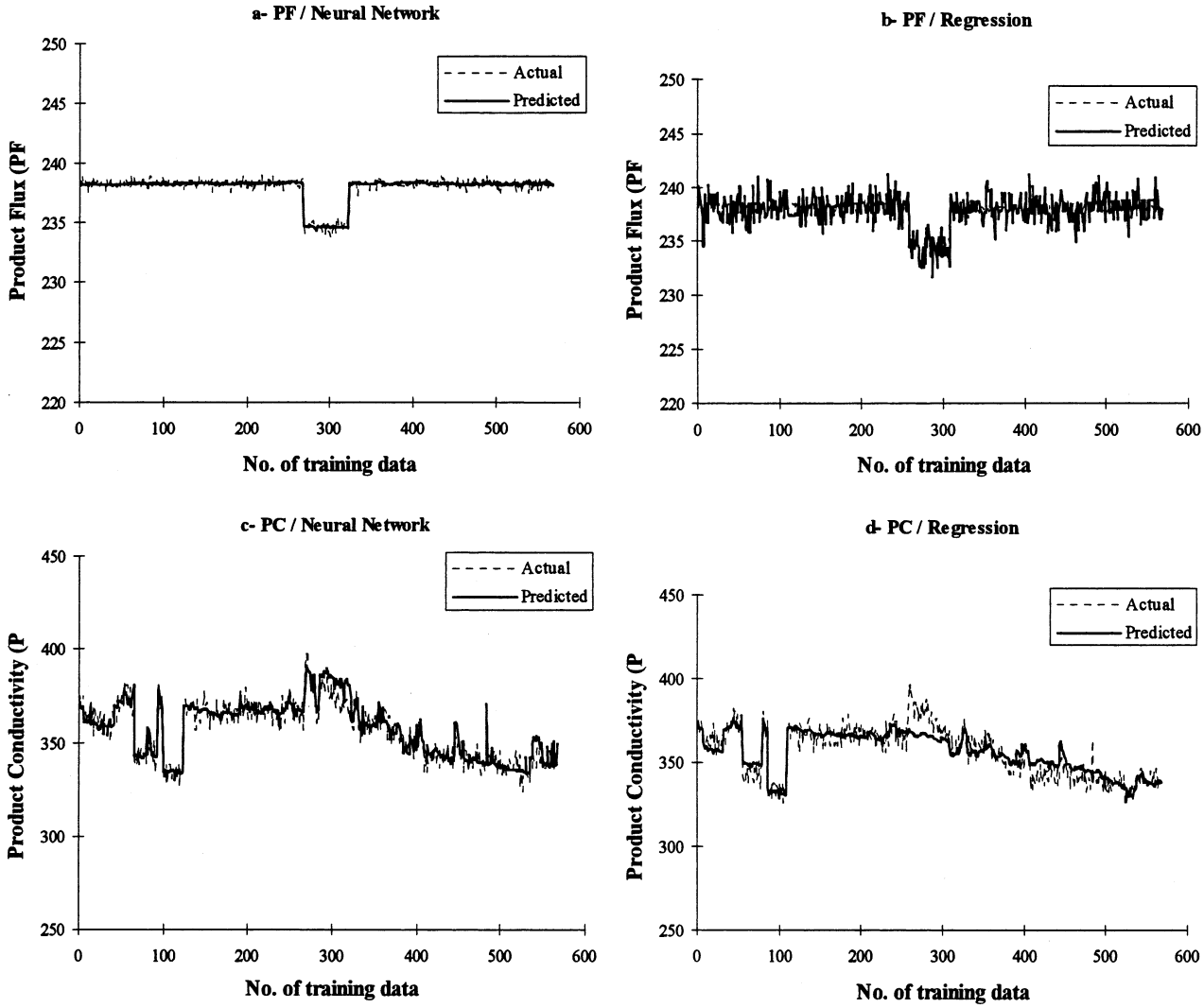


**Figure 12.** Actual and predicted product flux (PF) and product conductivity (PC) from the RO plant by neural network and by statistical regression.

development of an accurate predictive neural network model for the Jeddah Seawater RO plant in Saudi Arabia, and the details are available in Al-Shayji[9] and in Al-Shayji and Liu.[10]

The most important operating variables for the RO plant are feed temperature (FT), feed pressure (FP), feed pH (pH), product flow rate (PF), and product conductivity (PC), of which the most significant ones are PF and PC. Figure 12 compares the predicted and actual outputs from neural networks and statistical regression. The correlation coefficients $R^2$ values for the PF and PC are 0.9167 and 0.9512, respectively. Despite the very good $R^2$ values, the neural network still outperforms the statistical regression. Because it is difficult to develop a model-based process simulation of a large-scale RO desalination plant, our work demonstrates that we can use a data-based neural network to accurately predict the desalination performance variables.

## 4. Model-Based Process Simulation

**4.1. Modeling Assumptions.** We make the following assumptions in the development of steady-state and dynamic models: (1) The distillate leaving any stage is salt-free. (2) The system is well insulated so that there is no heat loss to the surroundings. (3) The tube temperature is equal to the vapor temperature. (4) All liquid and vapor phases are well-mixed, so their temperatures, concentrations, and specific enthalpies are the same as those of the outlet brine from the flash chamber. (5) There is no entrainment of mist by the vapor flashed from the brine. (6) No subcooling of condensate leaving the brine heater occurs. (7) The low-pressure saturated steam is supplied to the brine heater. (8) Noncondensable gases have negligible effects on the efficiency of heat transfer because they are removed by the ejectors at particular stages in the plant. (9) The contractor's total temperature loss is considered.

**4.2. Steady-State Process Simulation.** We describe the methodology of modeling the brine heater, heat-recovery section, and heat-rejection section below. Interested readers can refer to Al-Shayji[9] for detailed model parameters.

**4.2.1. Brine Heater.** Figure 13 compares the actual plant flowsheet and ASPEN PLUS model for the brine heater. The brine heater consists of two inlet streams and two outlet streams. One inlet stream is the recirculating brine ready to be heated to the desired temperature. The second inlet stream is the steam used to heat the recirculating brine. Both streams exit accordingly, without mixing.

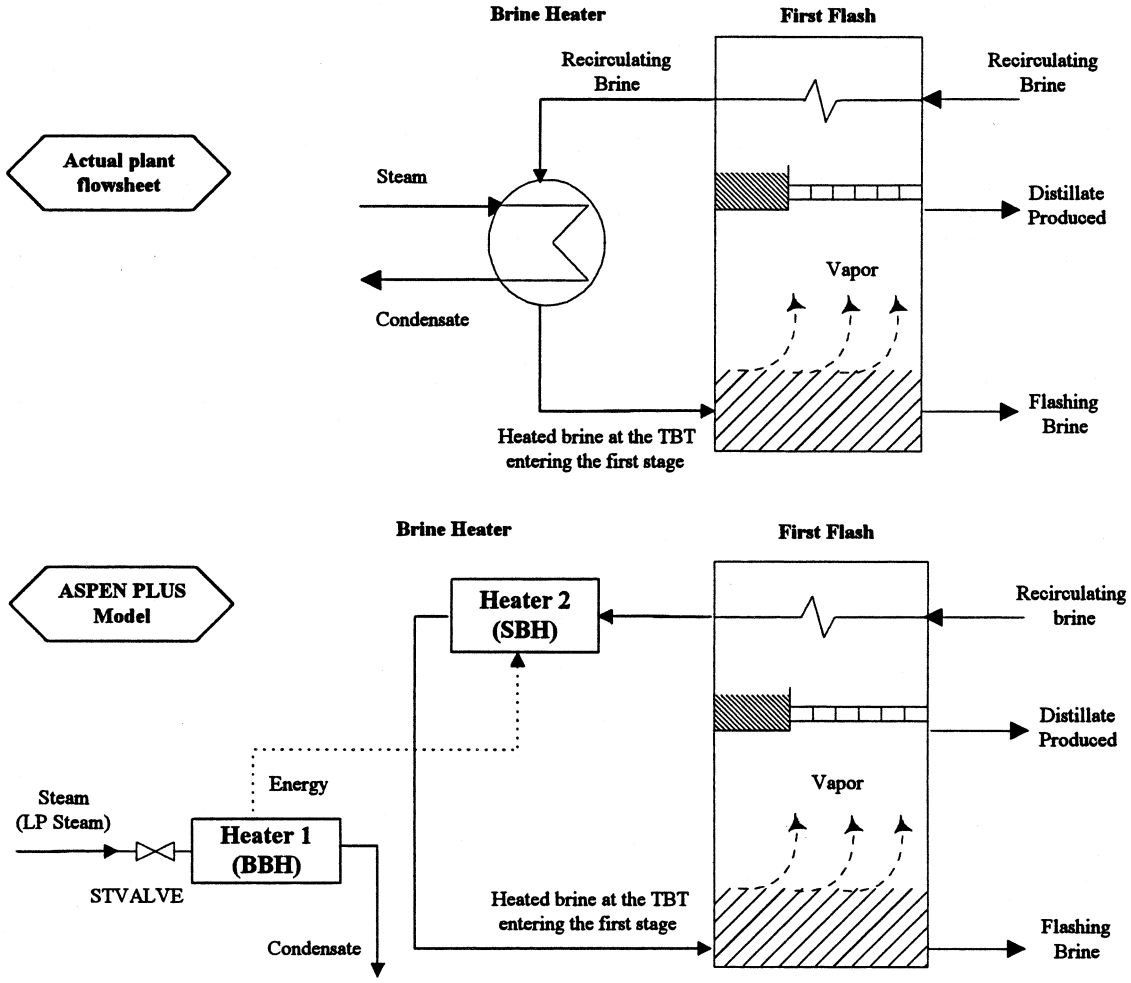


**Figure 13.** Comparison of the actual plant flowsheet and the ASPEN PLUS model for the brine heater.

We use two heaters (heater 1 and heater 2) together with a heat stream in ASPEN PLUS to simulate the brine heater.

**4.2.2. Single Flash Stage.** Figure 14 shows the actual flowsheet of a single flash stage and its representation using ASPEN PLUS. As the brine flows into the stage, it becomes superheated and ultimately flashed off to give pure vapor as a result of pressure reduction. In the actual process, the vapor passes through demisters, where the salt carried with the vapor is removed. Because of the limitations of ASPEN PLUS, we cannot include demisters in the flash model, and we assume the vapor to be salt-free. A temperature drop exists across the demister because of the pressure drop, as well as across the stage because of the boiling-point elevation. To account for this temperature change, we include a block, heater 1, in the ASPEN PLUS model.

As the vapor passes through demister, it condenses on the cooling tubes into a water box. The cooling tubes contain recirculating brine, which uses the heat released by condensation to preheat the feed. We include a block, heater 2, to simulate the phase change from vapor to liquid. Here, the vapor enters the heater and condenses into a liquid while generating heat. At the same time, the generated energy serves as a heat stream and enters heater 3 to simulate the preheating of the recirculating brine flowing through tubes that run across the top of the stage.

The condensed vapor is collected as distillate in the distillate tray inside the flashing stage. The distillate produced from one stage is combined with that from previous stages into a distillate duct. We use a block, MIXER, to combine the distillate produced from one stage with the distillate produced from all of the successive stages.

**4.2.3. Heat-Recovery Section.** The heat-recovery section consists of 21 flash stages connected to one another. The block diagram for each stage of the ASPEN PLUS model is identical to that of Figure 14.

**4.2.4. Heat-Rejection Section.** The heat-rejection section consists of three stages. We simulate all of the stages in the heat-rejection section in the same manner as in the heat-recovery section, except for the last stage. Figure 15 compares the actual plant flowsheet and ASPEN PLUS model for the last stage of the heat-rejection section.

The last flash includes one additional input (makeup seawater) and one additional output stream (recirculating brine). The cooling water in the heat-rejection tubes is seawater and not brine as in the heat-recovery section. The feed seawater (makeup) and a large mass of the last-stage brine mix together to give the recirculating brine that circulates through the heat-recovery tubes. In the actual plant, this mixing takes place inside the last stage. However, because this is not possible with the two-phase flash model (FLASH2) in ASPEN PLUS, we assume outside mixing after rejecting part of the last-stage outlet brine (blowdown) back to the sea.

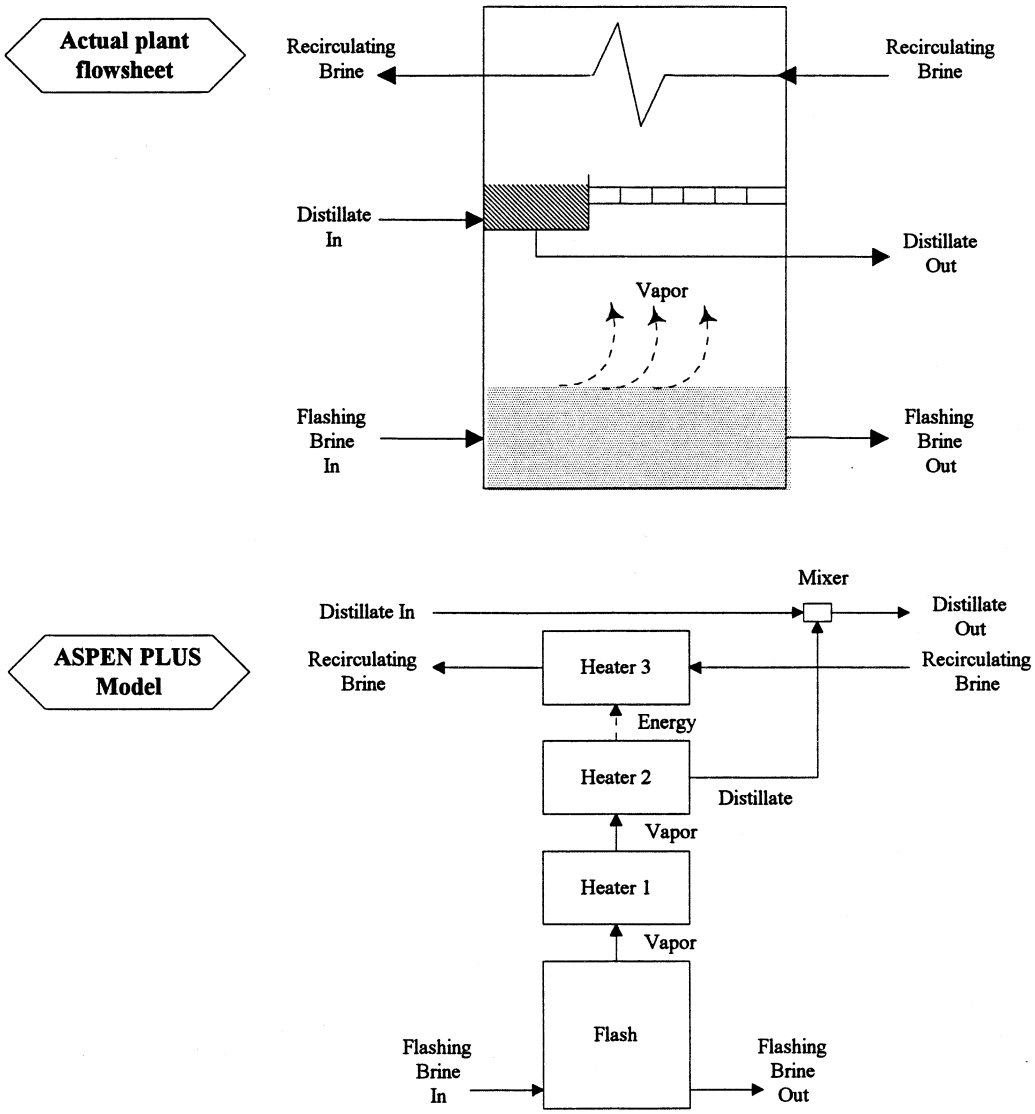


**Figure 14.** Comparison of the actual plant flowsheet and ASPEN PLUS for a single flash stage.

**4.3. Steady-State Model Verification.** Figure 16 compares the simulated results of the outlet pressure, P_OUT (bar) for summer operation from each stage with design and operating data.

Table 8 summarizes a similar comparison for three performance variables for summer operation. The MSF desalination plant can operate without any problems within 5 °C of design values for the top brine temperature. Therefore, a temperature difference of 1.46 °C between the simulated and design values of the top brine temperature is not significant.

We conclude from this comparison that a model-based process simulation using ASPEN PLUS is able to accurately simulate the design and operational data of the MSF desalination plant.

**4.4. Steady-State Model Applications.** We illustrate the applications of the ASPEN PLUS model the prediction of desalination performance under new operating conditions.

In designing a single flash stage, we must specify the temperature and pressure of the flash operation. We choose the temperature on the basis of actual plant data and apply the ASPEN PLUS model to determine the outlet pressure of the flashing brine using a sensitivity analysis. For example, we wish to find the required outlet pressure to produce 41.1 ton/h of vapor from the first flash stage, when the inlet and outlet temperature of the flashing brine are 90.56 and 88.9 °C, respectively. Figure 17 shows the effect of the flash outlet pressure on the amount of vapor produced for the first flash stage. We see that the outlet pressure required to produce 41.1 ton/h of vapor is 0.6574 bar.

As another application, we investigate the effect of changing the seawater temperature on the resulting top brine temperature. The sensitivity analysis tool available in ASPEN PLUS allows us to simulate this effect easily. Figure 18 shows a resulting plot of the top brine temperature versus the seawater inlet temperature.

This application shows that, when compared to a data-based neural network, a properly validated simulation model can more effectively quantify the effects of varying operating variables on the desalination performance variables.

**4.5. Steady-State and Dynamic Process Simulations.** Husain et al.[5] developed a SPEEDUP model for simulating the steady-state and dynamic performance of an 18-stage MSF desalination plant (22 720 $m^3$ per day, or 6 MGPD) located at Abu Dhabi, United Arab

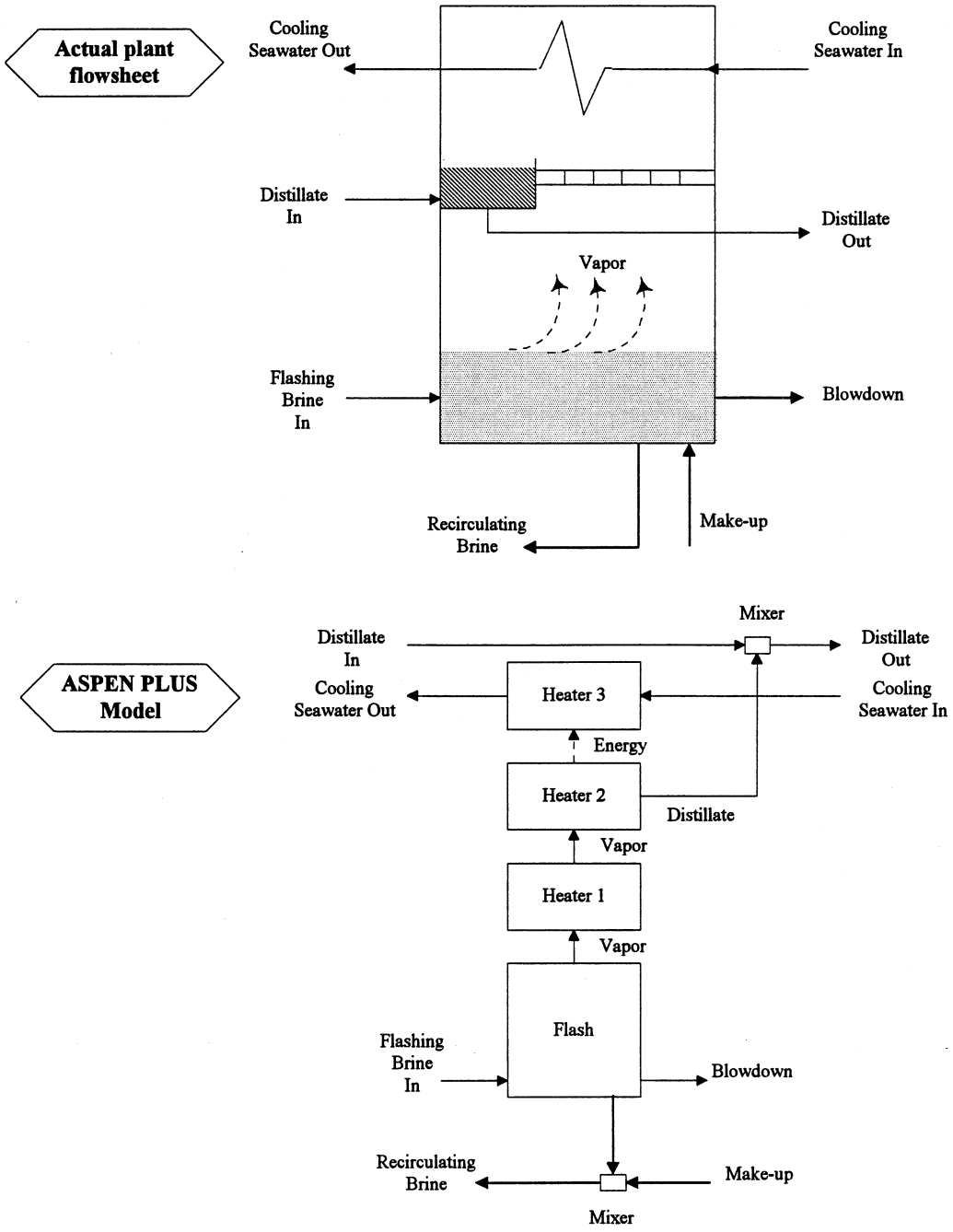


**Figure 15.** Comparison of the actual plant flowsheet and ASPEN PLUS model for the last stage of the heat-rejection section.

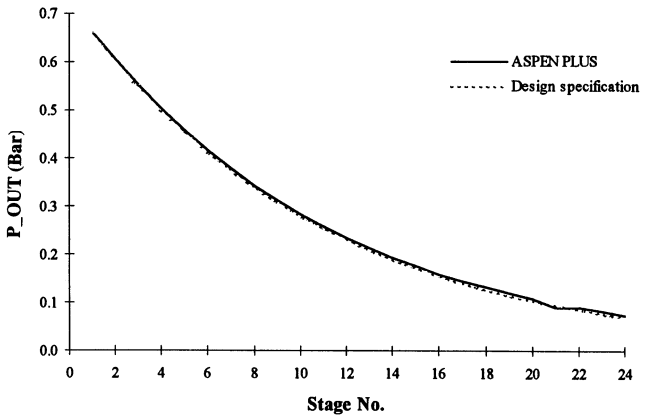


**Figure 16.** Comparison of simulated values and design specifications for the outlet pressure from each flash stage for summer operation.

Emirates. The model consists of mass and energy balances for all sections in the MSF plant, together with the Fortran correlations for heat transfer and physical properties of liquid and vapor streams in the plant. We have modified this model for specific application to the AZ-ZOUR MSF desalination plant in Kuwait. This plant has a total of 24 stages, including 21 heat-recovery stages and 3 heat-rejection stages. Our focus is to use the resulting SPEEDUP model to evaluate the relative merits of the model-based process simulation and the data-based neural network when applied to large-scale MSF desalination plants.

Figure 19 shows the components of the SPEEDUP model for our plant. Interested readers can refer to Husain et al.[5] and Al-Shayji[9] for detailed model equations and parameters.

**Table 8. Comparison of Model-Based Steady-State Process Simulation Using ASPN PLUS with Design Specifications**

| performance variable | units | Aspen PLUS | design specification |
|---|---|---|---|
| top brine temperature | °C | 89.1 | 90.56 |
| recycle brine flow rate | ton/min | 238.09 | 238.1 |
| distillate produced | ton/min | 18.816 | 18.795 |

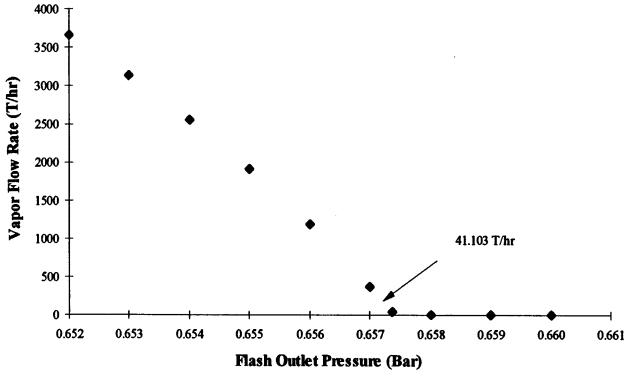


**Figure 17.** Effect of flash outlet pressure on the amount of vapor produced for the first flash stage.

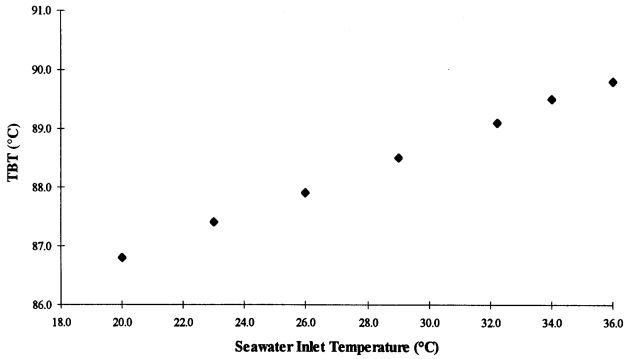


**Figure 18.** Effect of changing seawater inlet temperature on the top brine temperature.

Table 9 compares the actual and simulated values of the recirculating brine temperature entering the brine heater (TF), temperature of the final distillate produced (TD), and the top brine temperature (TB), as well as the flow rate of distillate produced (D), blowdown (B), and steam to brine heater (S). In addition, this table shows the simulated and actual performance ratio, which is the ratio of distillate produced to steam consumed. It is clear that the simulated results match the design and operating data well.

## 5. Comparison of Data-Based Neural Network and Model-Based Process Simulation

In the preceding sections, we have concluded that the accuracy of a neural network model depends on the proper selection of input variables. Additionally, this accuracy depends on the broad range of data with which the network is trained. To fully demonstrate the latter significant feature of a data-based neural network, especially in relation to model-based process simulation, we purposely compare the results of process simulations using ASPEN PLUS and SPEEDUP with a neural network that is trained with summer operating data at high ranges of top brine temperature (104−105.6 °C) and distillate produced (21.85−23.83 ton/h). The comparison in Table 10 shows that the neural network might not seem at first to have produced favorable results, but this is due to the range of data with which the network is trained. Neural network can interpolate data accurately, but it often does not generalize or extrapolate well. As a result, testing the network with the designed distillate product flow rate of 18.79 ton/h at a top brine temperature of 90.56 °C is expected to predict a higher distillate flow rate. To produce favorable results, the network needs to be trained with a wider range of operating data.

To model the desalination plant completely in ASPEN PLUS, we need to use the design-specification tool to manipulate some independent variables to obtain the specified value of the desired distillate flow rate. Therefore, the distillate product flow rate obtained by ASPEN PLUS is very close to the design value, with an error of only 0.14%. The sensitivity-analysis tool available in ASPEN PLUS enables us to effectively quantify the effects of varying operating variables on the desalination performance variables, as Figures 16 and 17 illustrate.

SPEEDUP produces the desired results, because we accurately model the plant with mass and energy balances, as well as with physical property correlations. However, the input preparation for SPEEDUP takes a considerable amount of time. To model our MSF desalination plant, SPEEDUP has to solve a total of 975 equations (including 1362 variables and parameters), which are further grouped into 238 blocks. Consequently, convergence problems can often occur, and it proves costly to redo simulations for any small change in the input conditions.

## 6. Conclusions

(1) Neural networks have been very effective in predictive modeling of the performance variables of large-scale commercial MSF and 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.

(2) 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.

(3) We recommend the following network parameters and functions for modeling a neural network for a large-scale commercial 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; and (5) initial values for both the learning rate and the momentum coefficient of {0.3, 0.5, and 0.7} and {0.4}, respectively.

(4) Our comparative studies show that our neural network model performs favorably compared to recent models developed by others in accurately predicting commercial desalination plant data. The input variables recommended by recent studies are simply insufficient for the accurate predictive modeling of actual operating data from large-scale MSF desalination plants.

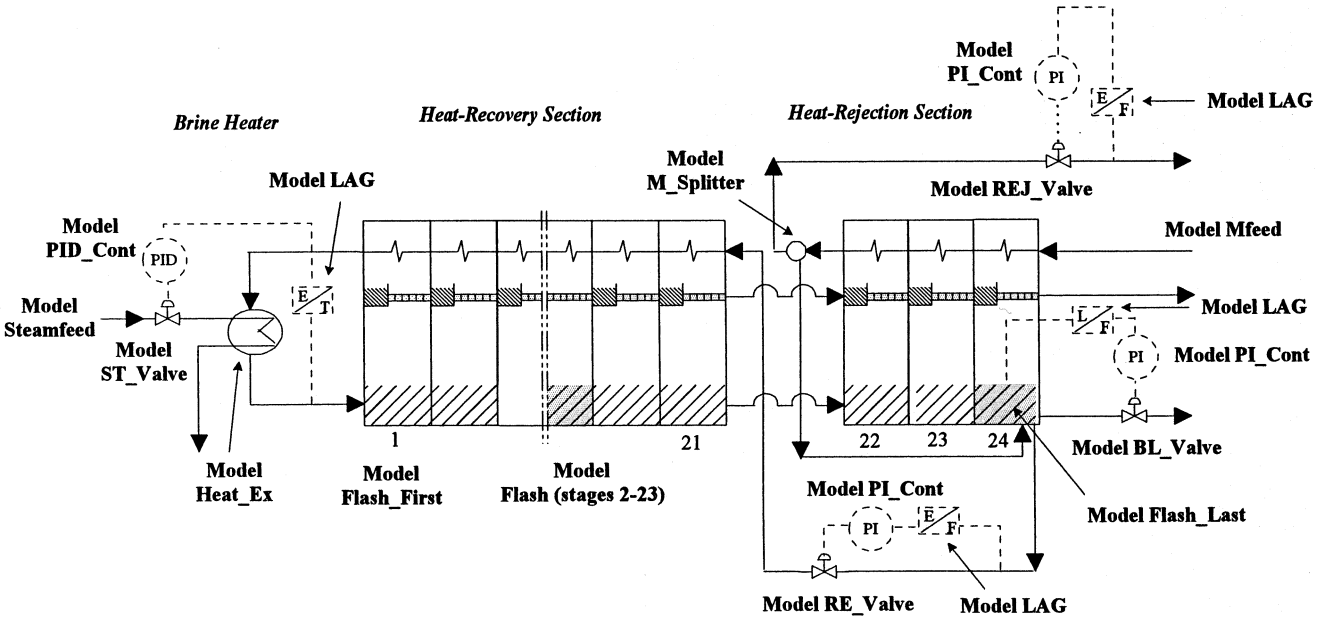


**Figure 19.** Components of the SPEEDUP model for the MSF desalination plant.

**Table 9. Comparison of Model-Based Steady-State Process Simulation Using SPEEDUP with Design Specifications**

| variable | units | SPEEDUP | design |
|---|---|---|---|
| TF_IN | °C | 84.43 | 84.89 |
| TD_OUT | °C | 36.54 | 38.60 |
| TB_OUT | °C | 38.44 | 40.50 |
| D_OUT | ton/min | 19.33 | 18.80 |
| B_OUT | ton/min | 29.43 | 29.96 |
| S_OUT | ton/min | 2.50 | 2.35 |
| performance ratio | kg/540 kcal | 7.76 | 8.00 |

**Table 10. Comparison of Data-based Neural Network with Model-based Process Simulation.**

| modeling approach | distillate produced (ton/min) |
|---|---|
| summer operation at a top brine temperature of 90.56 °C | 18.79 |
| neural network | 23.92 |
| Aspen PLUS | 18.82 |
| SPEEDUP | 19.33 |

(5) A properly validated model-based process simulation (as in the case of the MSF desalination plants) can more effectively quantify the effects of varying operating variables on the desalination performance variables. When it is difficult to develop a model-based process simulation (as in the case of RO desalination plants), we can use a data-based neural network to accurately predict the desalination performance variables.

## Acknowledgment

We thank the Honeywell International Foundation and Aspen Technology (particularly Dr. Jila Mahalac, Director, Worldwide University Programs) for supporting the computer-aided design educational program at Virginia Tech. We gratefully acknowledge the support of Mr. Abudiullah Al-Azaz at Saline Water Conservation Corporation, Saudi Arabia, and the Ministry of Electricity and Water, Kuwait, for providing us with actual plant data. We also thank NeuealWare, Inc., Pittsburgh, PA, for providing a free copy of their PC-based Professional II/Plus for use in our work. Finally, we thank Christine Petrella for her help in this study.

## Literature Cited

(1) Rao, G. P.; Darwish, D. M.; Hassan, A.; Kurdali, A. Toward Improved Automation for Desalination Processes. Part II. Intelligent Control. *Desalination* **1994**, *97*, 507.

(2) Abdulbary, A. F.; Lai, L. L.; Reddy, K. V.; Al-Gobaisi, D. M. Artificial Neural Networks as Efficient Tools of Simulation. In *Proceedings of the IDA World Congress on Desalination and Water Sciences*; International Desalination Association: Topsfield, MA, 1995.

(3) Parenti, R.; Bogi, S.; Massarani, A. Industrial Application of Real-Time Neural Networks in Multistage Desalination Plant. In *Proceedings of the IDA World Congress on Desalination and Water Sciences*; International Desalination Association: Topsfield, MA, 1995.

(4) Selvarajs, R.; Deshpande, P. B., Neural Networks for the Identification of MSF Desalination Plants. *Desalination* **1995**, *101*, 185.

(5) Husain, A.; Woldai, A.; Al-Radif, A.; Kesou, A.; Borsani, R.; Sultan, H.; Deshpandey, P. B. Modeling and Simulation of a Multistage Flash (MSF) Desalination Plant. *Desalination* **1994**, *97*, 555.

(6) Baughman, D. R.; Liu, Y. A. *Neural Network in Bioprocessing and Chemical Engineering*; Academic Press: San Diego, CA, 1995.

(7) Utojo, U.; Bakshi, B. R. A Unified View of Artificial Neural Networks and Multivariate Statistical Methods. In *Neural Networks in Bioprocessing and Chemical Engineering*; Baugham, D. R., Liu, Y. A.; Academic Press: San Diego, CA, 1995; pp 435−459.

(8) Ettouney, H. M.; El-Dessouky, H. T.; Alatiqi, I. Understand Thermal Desalination. *Chem. Eng. Prog.* **1999**, *95* (9), 43.

(9) Al-Shayji, K. A. Modeling, Simulation and Optimization of Large-Scale Commercial Desalination Plants. Ph.D. Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA, 1998.

(10) Al-Shayji, K, A.; Liu, Y. A. Neural Networks for Predictive Modeling of Large-Scale Commercial Water Desalination Plants. In *Proceedings of the IDA World Congress on Desalination and Water Reuse*; International Desalination Association: Topsfield, MA, 1997.

*Received for review* January 25, 2002
*Revised manuscript received* September 19, 2002
*Accepted* September 25, 2002

IE020077R