Heat Rate Predictions in Humid Air-Water Heat Exchangers Using Correlations and Neural Networks

We consider the flow of humid air over fin-tube multi-row multi-column compact heat exchangers with possible condensation. Previously published experimental data are used to show that a regression analysis for the best-fit correlation of a prescribed form does not provide an unique answer, and that there are small but significant differences between the predictions of the different correlations thus obtained. It is also shown that it is more accurate to predict the heat rate directly rather than through intermediate quantities like the j-factors. The artificial neural network technique is offered as an alternative technique. It is trained with experimental values of the humid-air flow rates, dry-bulb and wet-bulb inlet temperatures, fin spacing, and heat transfer rates. The trained network is then used to make predictions of the heat transfer. Comparison of the results demonstrates that the neural network is more accurate than conventional correlations.

[DOI: 10.1115/1.1351167]

Keywords: Artificial Intelligence, Condensation, Convection, Heat Transfer, Heat Exchangers

1 Introduction

For the design of a thermal system it is often necessary, for selection purposes, to predict the heat transfer rates of heat exchangers under specific operating conditions. Heat exchangers are complex devices, the complexity being due both to the geometry and to the physical phenomena involved in the transfer of heat. For a heat exchanger operating with humid air, e.g., in refrigeration and air-conditioning applications, some of the moisture in the air may condense on the fins and tubes. Condensing heat exchangers have been studied by Jacob and Goldschmidt [1], Srivisan and Shah [2], Ramadhyan [3], and Lang et al. [4], among others, using heat and mass transfer coefficients. Physical processes related to the latent heat and modification of the flow field by the water film or droplets increase the computational difficulty of the problem, and numerical calculations entirely based on first principles are not possible. Experiments must be carried out, usually by the manufacturers for each of their models, to determine the heat rates as functions of the system parameters like the flow rates, inlet temperatures and fin spacing. The experimental information must then be transferred in some way to the user of the information who needs it to predict the heat rates under different operating conditions. One way is to provide the heat transfer coefficients. However, these are not constant but vary considerably with operating conditions, and can thus provide only very rough approximations. A better and more common procedure is to compress information about the heat transfer coefficients by means of correlations so that variations with respect to the operating parameters can be taken into account through standard nondimensional groups. Usually the form of the correlation cannot be totally justified from first principles; it is selected on the basis of simplicity and common usage. Estimated errors in heat rates from correlations are normally larger than the experimental error, being mainly due to the data compression that occurs through the correlation process. Another reason for inaccuracy in predictions is that, for most forms of correlating functions that are used, a least-squares analysis of the error gives multiple sets of values for the constants indicating that a local, rather than a global, optimum set of values may have been found.

Contributed by the Heat Transfer Division for publication in the JOURNAL OF HEAT TRANSFER. Manuscript received by the Heat Transfer Division January 12, 2000; revision received October 3, 2000. Associate Editor: R. Mahajan.
percent. This discrepancy may be due to heat losses between the locations at which the temperatures were measured and the heat exchanger. The data for dropwise and film condensation could also have been similarly affected.

McQuiston [15,16] was interested in predicting the air-side heat transfer. He used high Reynolds-number turbulent flow on the water side so that its heat transfer coefficient, which could be estimated from the Dittus-Boelter correlation, was much higher than that on the air side and thus introduced little error. He defined the air-side heat transfer coefficients using the log-mean temperature difference for the dry surface. A similar procedure was followed for determining the air-side heat transfer coefficients under condensing conditions using the enthalpy difference as the driving potential. The correlations that were obtained are

\[ j_s = 0.0014 + 0.2618 \frac{A}{A_{ib}} \delta^{-0.4} \]  
\[ j_t = 0.0014 + 0.2618 \frac{A}{A_{ib}} \delta^{-0.4} \]

for dry surface

\[ j_s = 0.328 \]  
\[ j_t = 0.328 \]

for dropwise condensation

\[ j_s = 0.502 \]  
\[ j_t = 0.502 \]

for filmwise condensation

\[ \delta = \frac{h_s}{\rho_c c_p} \]  
\[ \delta = \frac{h_s}{\rho_c c_p} \]

for dry surface

\[ \delta = \frac{h_s}{\rho_c c_p} \]  
\[ \delta = \frac{h_s}{\rho_c c_p} \]

for dropwise condensation

\[ \delta = \frac{h_s}{\rho_c c_p} \]  
\[ \delta = \frac{h_s}{\rho_c c_p} \]

for filmwise condensation

Gray and Webb [17] added data from other sources to find a different correlation

\[ j_s = 0.14 \Re_D^{0.28} \left( \frac{x_b}{x_a} \right)^{-0.502} \left( \frac{\delta - t}{D} \right)^{0.012} \]

for dry surface

with its corresponding range of applicability.

### 3 Multiplicity of “Best” Correlations

Let us look at a regression analysis to find the best correlation for the dry surface, the procedure for the other cases being similar. The general form of this correlation, as proposed by McQuiston [16], is

\[ j_s = a + b \Re_D^{c} \left( \frac{A}{A_{ib}} \right)^{-d} \]

In a least-squares method the difference between the measured and the predicted values of \( j_s \) is minimized to determine the constants \( a, b, c, \) and \( d \) in the above equation. This is done by calculating the variance of the error defined as

\[ S_{j_s} = \frac{1}{M_1} \sum_{i=1}^{M_1} \left( (j_s)_i - (j_s)_i^\prime \right)^2, \]

where \( (j_s)_i \) and \( (j_s)_i^\prime \) for \( i = 1,...,M_1 \), are the experimental measurements and \( (j_s)_i^\prime \) for \( i = 1,...,M_1 \), are the values predicted by Eq. (8). \( S_{j_s}(C) \) is a smooth manifold in a five-dimensional space, where \( C = (a,b,c,d) \) is the vector of unknown constants, and a search must find \( C \) such that \( S_{j_s}(C) \) is a minimum.

This procedure was carried out for the \( M_1 \) data sets. It was found that \( S_{j_s} \) had multiple local minima, the following two being examples.
Table 1 Comparison of percentage errors in $j_s$, $j_t$, and $\dot{Q}$ predictions between various correlations and the ANN

<table>
<thead>
<tr>
<th>Surface</th>
<th>Prediction method</th>
<th>$j_s$</th>
<th>$j_t$</th>
<th>$\dot{Q}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry</td>
<td>McQuiston (1978b)</td>
<td>14.57</td>
<td>14.57</td>
<td>6.07</td>
</tr>
<tr>
<td></td>
<td>Eq. (10)</td>
<td>8.1</td>
<td>8.1</td>
<td>3.84</td>
</tr>
<tr>
<td></td>
<td>Eq. (15)</td>
<td>9.3</td>
<td>9.3</td>
<td>3.84</td>
</tr>
<tr>
<td></td>
<td>ANN</td>
<td>1.002</td>
<td>1.002</td>
<td>0.928</td>
</tr>
<tr>
<td>Dropwise</td>
<td>McQuiston (1978b)</td>
<td>8.50</td>
<td>7.55</td>
<td>8.10</td>
</tr>
<tr>
<td></td>
<td>Minimizing $S_t$</td>
<td>9.12</td>
<td>9.43</td>
<td>9.91</td>
</tr>
<tr>
<td></td>
<td>Minimizing $S_Q$</td>
<td>-17.76</td>
<td>4.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ANN</td>
<td>3.32</td>
<td>3.87</td>
<td>1.446</td>
</tr>
<tr>
<td>Filmwise</td>
<td>McQuiston (1978b)</td>
<td>9.01</td>
<td>14.98</td>
<td>10.25</td>
</tr>
<tr>
<td></td>
<td>Minimizing $S_t$</td>
<td>8.32</td>
<td>10.89</td>
<td>10.44</td>
</tr>
<tr>
<td></td>
<td>Minimizing $S_Q$</td>
<td>-18.97</td>
<td>4.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ANN</td>
<td>2.58</td>
<td>3.15</td>
<td>1.960</td>
</tr>
<tr>
<td>Combined</td>
<td>ANN</td>
<td>4.58</td>
<td>5.05</td>
<td>2.69</td>
</tr>
</tbody>
</table>

The same global search procedure has been followed to find the best correlations for dropwise and film condensation by first fixing the constants found for the dry surface and then searching for the additional constants that Eqs. (3) and (4) have. The results of the correlations for the dry surface are shown in Table 1 as “Eq. (10),” and for the wet surfaces as “Minimizing $S_j$.” The table shows the ability of the correlations to predict the same data from which they were derived. The errors indicated are the root-mean-square (rms) values of the percentage differences between the predicted and experimental data. The results of the correlations of McQuiston [16] for all surfaces and Gray and Webb [17] for dry surfaces only are also shown; as claimed, the latter is seen to perform better than the former. The correlations found here using the global search technique are much better than the previous correlations for the dry surface, slightly worse for dropwise and slightly better for filmwise condensations. Some of the differences could be due to possible elimination of outliers by previous investigators; none of the published data has been excluded here.

4 Direct Correlation of Heat Rate

Using the method of the previous section to determine the heat rate, which is what the user of the information is usually interested in, one must find the $j$-factor first, then the heat transfer coefficient, and finally the heat rate. The procedure gives some generality to the results since different fluids and temperatures can be used. If accuracy is the goal, however, it may be better to correlate the heat rate directly, as will be investigated in this section. The dry-surface case is discussed in some detail, and for the wet surfaces the procedure is similar.

Using the water-side heat transfer coefficient defined on the basis of the log-mean temperature difference [15], the heat rate is

\[
\dot{Q} = (T_w - T_a) \left( \frac{1}{(m_c p_w)^{-1} - (m_a c_p a_d)^{-1}} \right) \left( \frac{1}{\eta G_c c_p a_d j_s} \right)
\]

where

\[
U = \left( \frac{A}{A_w h_w} + \frac{P_{2/3}}{\eta G c_p a_d j_s} \right)^{-1}
\]

and $j_s$ is given in Eq. (8). To find the constants in the correlation, the mean square error of the heat transfer rates must be minimized. A genetic-algorithm based global search results in a correlation, which when converted to a $j$-factor is

\[
S_{\dot{Q}} = \frac{1}{M_1} \sum_{k=1}^{M_1} [\dot{Q}_k - \dot{Q}_k']^2
\]
which is different from Eqs. (1), (10), and (11), all of which have been obtained from the same data.

The results of using direct correlation of the heat rate are also shown in Table 1 as “Eq. (15).” The percentage error is 3.84 percent, which is the same as that given by Eq. (10). However, if we compare the absolute, dimensional rms error, direct correlation of heat rate shows an error of only 158.6 W, compared to an error of 272 W from the j-factor correlation; this is an improvement of 41.7 percent. The global minimization of $S_Q$ does not necessarily imply the global minimization of $S_{j}$. The difference of the predictions in $j$, is 14.8 percent between the two methods.

A similar procedure was followed to find the correlations for $j$, under wet-surface conditions, and the errors are given in Table 1 under “Minimizing $S_Q$.” The direct correlation of dimensional heat rate gives predictions that have 56.9 percent less error compared to the j-factor correlation for dropwise condensation and 58.6 percent for film condensation. Thus, predicting the heat rate rather than j-factors gives more accurate predictions. This is somewhat expected since the j-factor assumes the existence of the heat transfer coefficient, while $Q$ does not.

5 Artificial Neural Networks

ANNs offer an attractive alternative to the correlation method discussed in the previous sections to predict the performance of heat exchangers. Although there are many different types of neural networks, the feedforward configuration has become the most widely used in engineering applications [19]. This consists of a series of layers, each with a number of nodes, the first and last layers being the input and output layers while the remaining are hidden layers. The nodes of each layer are connected only to those of the layer before and the one after. The connections are associated with weights and the nodes with biases. These can be adjusted during the training procedure using known data: for a given input the actual output is compared with the target output, and the weights and biases are repeatedly adjusted using the backpropagation algorithm [20] until the actual differs little from the target output. All variables are normalized to be within the [0.15, 0.85] range. Further details are in Sen and Yang [5].

5.1 Separation of Data for Training and Testing. The ANN structure chosen for the present analysis consists of four layers: the input layer at the left, the output layer at the right and two hidden layers. This 5-5-3-1 configuration, where the numbers stand for the number of nodes in each layer, is similar to the schematic shown in Fig. 4 which is used in the next section for the heat exchanger analysis, the only difference being that the fourth layer has only one output, the heat rate. The inputs to the network correspond to the air-flow Reynolds number $Re_p$, inlet air dry-bulb temperature $T_{in, dry}$, inlet air wet-bulb temperature $T_{in, wb}$, inlet water temperature $T_{in, w}$, and fin spacing $\delta$. The output is the total heat rate $Q$. For testing the trained ANN, the variables $Re_p$, $T_{in, dry}$, $T_{in, wb}$, $T_{in, w}$, and $\delta$ are input and the corresponding $Q'$ are predicted.

A total of $M = 327$ experimental runs were reported by McQuiston [15] for three different surface conditions. The data can be separated in different ways into training and testing data. Diaz et al. [13], for example, used 75 percent of the total data sets available for training and the rest for testing. This has two disadvantages: first, the data set available for training is smaller than the total amount of information available so that the predictions are not the best possible, and second, if the training data do not include the extreme values the predictions fall in the extrapolated range and are hence less reliable. The issue of separating the complete data into training and testing sets is further analyzed in the following way.

![Fig. 3 ANN prediction errors versus percentage of data used for training; –×– error $E_a$ using training data; –○– error $E_b$ using data not used for training; –○– error $E$ using complete data. Error bars indicate standard deviations. The $E_a$ and $E$ curves have been shifted horizontally by −1 percent and 1 percent, respectively, for clarity.](image)

The $M$ available sets of experimental data are first randomly sorted to avoid introducing any bias in the selection process and their order is then fixed. Only the first $M_a$ of these are chosen for training and the rest $M_b = M - M_a$ kept aside for the moment. The fraction used for training is thus $P_s = M_a/M$. The rms values of the relative output errors

$$S_{Q'} = \left[ \frac{1}{M_a} \sum_{i=1}^{M_a} \left( \frac{Q'_i - Q_i}{Q_i} \right)^2 \right]^{1/2}$$

are calculated at each cycle of the training process in order to evaluate the performance of the network and to update the weights. Here $i = 1, ..., M_a$, where $Q''$ are the predictions, and $Q'$ are the experimental values of the heat rates. A reasonably low level of error, $S_{Q'}$, in the training process is obtained with 250,000 cycles, which is the same for the rest of the procedure.

After the training is finished, three different data sets are tested: (a) the same $M_a$ data that were used for training are tested, (b) the $M_b$ data left out of the training process are tested, and (c) the complete $M$ data sets are tested. In each case the percentage error between the predicted and experimental values are calculated, being $E_a$, $E_b$, and $E$, respectively. Without reordering the $M$ data sets the procedure described above is repeated for different values of the percentage of splitting, i.e., $P_s = 10$ percent, 20 percent,..., 90 percent, 95 percent, and 99 percent. The exact shape of the error versus $P_s$ curve depends on the initial order of the data sets, but some general features can be identified.

To get the overall characteristics of the error, the curves were calculated ten times and the results averaged to remove the influence of the initial random ordering of the data sets. Figure 3 shows the average error in prediction calculated in the three different ways as a function of the training fraction $P_s$. The error bars indicate the standard deviations, $\sigma_a$, $\sigma_b$, and $\sigma$ corresponding to $E_a$, $E_b$, and $E$, respectively, that resulted from the ten different curves. From this figure, it can be seen that as $P_s$ increases the prediction errors for all three cases asymptote, on the average, to approximately the same values. For any $P_s$, $E_b$ is always small since the training and testing data are identical. At small values of $P_s$, $E_a$ and $E$ are both very large indicating that an insufficient fraction of the data has been used for training. As we increase $M_a$, better predictions are obtained. Beyond $P_s = 60$ percent approximately the differences in the prediction errors
for all data sets are small. The same trend is observed for the $s$ which become smaller as $P_s \rightarrow 100$ percent, indicating that the prediction is somewhat insensitive to the initial random ordering of the data. Near the end, however, at $P_s = 99$ percent for example, $\sigma_b$ becomes large again because now $M_b$ is very small and the error depends greatly on the data set that is picked for testing.

From this exercise one can deduce that, in this case at least, if more than 60 percent of the available data is used for training, the results will be essentially the same. In the rest of the paper we use all of the available data for training, and use the same for testing. This gives us the best prediction over the widest parameter range. For the correlations also, as is commonly done, the entire data set was used for finding the correlation and testing its predictions.

5.2 Heat Exchanger Analysis. From the total of $M = 327$ experimental runs were reported by McQuiston [15], $M_1 = 91$ corresponded to dry-surface conditions, $M_2 = 117$ corresponded to dropwise condensation, and $M_3 = 119$ to film condensation. Figure 4 shows a schematic of the feedforward neural network configuration used for the present analysis. There are four layers in this configuration. The inputs to the network correspond to the same physical variables described in the previous section. The outputs correspond to $j_s$, $j_t$, and the total heat rate $\dot{Q}$. Thus, $j_s$, $j_t$, and $\dot{Q}$ are functions of $\text{Re}_D$, $T_{\text{a,db}}$, $T_{\text{a,wb}}$, $T_{\text{w,ib}}$, and $\delta$. For testing the trained ANN, the physical variables are input and the corresponding $j_s^p$ and $j_t^p$ and $\dot{Q}^p$ are predicted. The $j$-factors are not necessary for the heat rate predictions, but are calculated here merely for the purpose of comparison with the other methods.

As mentioned before, the performance of the network is evaluated by computing the rms values of the output errors

$$S_e = \left[ \frac{1}{M_k} \sum_{i=1}^{M_k} \left( \frac{O_{i,k}^p - O_{i,k}^e}{O_{i,k}^e} \right)^2 \right]^{1/2}$$

(17)

at each stage of the training. Now $i = 1, \ldots, M_k$, $k = 1, 2, 3$, where $O_{i,k}^p = [j_s, j_t, \dot{Q}]$ are the set of predictions, and $O_{i,k}^c = [j_s, j_t, \dot{Q}]$ are the experimental output values. Several ANN configurations were tested, as in Diaz et al. [13], and the best results were given by the fully-connected 5-5-3-3 configuration which was chosen. A number of 800,000 cycles was selected for training to assure a reasonably low level of error $S_e$.

6 Artificial Neural Network Results

The ANN results are also shown in Table 1 along with those of correlations previously discussed. For all three surfaces, the ANN predictions are much better than any of the correlations. It is of
interest to note that the ANN gives better prediction for dry surfaces than for wet. This is expected since the physical phenomena associated with condensation are more complex.

To determine whether the network based on training data separated by some physical condition would perform better than another trained with the combined data set, the entire set of runs $M$ was combined to train a single ANN. The error, shown in Table 1 under “Combined,” is larger than the ANN predictions for individual cases, indicating that this ANN has more difficulty in differentiating between the different physics involved. However, even then the predictions of the total heat rate have errors of the order of only 2.7 percent.

Comparisons of $j_s$ and $j_t$ under dry and wet-surface conditions between the previously published correlations and the ANN results are shown in Fig. 5–7. For the dry surface there is only one $j$-factor, but under condensing conditions both $j_s$ and $j_t$ are shown. In all the figures, the straight lines indicate equality between prediction and experiment. The accuracy and precision given by the ANN is remarkable. There are some data points in Fig. 5–7 that are outliers and can clearly be eliminated to improve the predictions, if desired.

The experimental measurements reported by McQuistion [15] were classified according to a visual procedure. For dry surface conditions this method is relatively reliable, but under wet conditions it is very difficult to distinguish between different forms of condensation by simple observation. One can ask if this visual information is really necessary, or if the data can be separated using some pattern recognition procedure. One of the simplest is numerical clustering using a k-means algorithm [21]. We use this method to classify the data into three groups, with all variables normalized in the [0, 1] range.

Table 2 shows the percentage of data sets of all three surface conditions that belong to a particular cluster I, II, or III. These results show that cluster I clearly corresponds to dry surface conditions, while the situation is less crisp for clusters II and III. The difference between the dry and wet wall data would have been sharper if only two clusters (dry and wet wall) had been used. Each of the three clusters is analyzed separately to train three ANNs using their respective data sets, $M_I=81$ for cluster I, $M_{II}=66$ for cluster II and $M_{III}=180$ for cluster III. The total heat rate errors given by the ANN procedure are 0.866 percent for cluster I, 1.242 percent for cluster II and 1.626 percent for cluster III which are of the same order for data separated by visual observation. Visual information is thus not really necessary and the clustering technique is a viable alternative that produces equally good predictions.

### 7 Conclusions

Correlations found from heat exchanger experimental data using a regression procedure are often non-unique. In this case a global error minimization algorithm, rather than a local procedure, should be used. We have shown that there can be small but significant differences between the two. Furthermore, the correlation that is obtained can only be as good as its assumed functional form. It must also be pointed out that heat transfer coefficients are defined upon the assumption of a characteristic temperature difference and their usefulness depends on the similarity of the temperature profiles. In other words, if the temperature distributions are always identical except for a constant, this constant is the only quantity needed to determine the heat flux at the wall. With property variations, condensation, laminar and turbulent flows, and many other phenomena present, this is often not the case. The situation is worse for mass transfer during condensation, since its rate is clearly not proportional to the enthalpy difference. Thus it is better to develop correlations to predict the heat rates directly, rather than intermediate quantities like the $j$-factors that assume the existence of transfer coefficients. In addition, the user of the information is usually interested only in the overall heat rate and not in intermediate variables like the heat transfer coefficients.

Artificial neural networks do not have the drawbacks of correlations, and are an attractive alternative that can accurately model the thermal behavior of heat exchangers without need to assume a functional form for the correlation nor an accurate mathematical model of the details of the process. We have shown that they can predict the behavior of heat exchangers under dry and wet conditions and their predictions are remarkably more accurate than those from correlations. The network is able to catch the complex physics in a heat exchanger very well. In the end, the error in the predictions of the neural network is probably of the same order as the uncertainty in the measurements, which is the best that can be expected.

ANNs also have some limitations. They do not provide any physical insight into the phenomena in which they are used, but then neither do correlations. The training of the network may be computationally expensive though its subsequent use for predic-

---

**Table 2** Percentage of data sets of different surface conditions that fall in different clusters

<table>
<thead>
<tr>
<th>Surface</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry</td>
<td>89.01</td>
<td>10.99</td>
<td>0.00</td>
</tr>
<tr>
<td>Dropwise</td>
<td>0.00</td>
<td>29.06</td>
<td>70.94</td>
</tr>
<tr>
<td>Filmwise</td>
<td>0.00</td>
<td>18.49</td>
<td>81.51</td>
</tr>
</tbody>
</table>

---

Fig. 7 Experiments versus predictions for heat exchanger with film condensation; +ANN; × McQuistion [16]. Straight line is the perfect prediction; (a) sensible heat $j_s$; (b) total heat $j_t$. 

---

**Journal of Heat Transfer**

A. P.-V. is the recipient of a CONACyT-Fulbright Fellowship from Mexico for which we are grateful. We also acknowledge the support of Mr. D. K. Dorini of BRGD-TNDNDR for this and related projects in the Hydronics Laboratory.

References


