Artificial Neural Networks in Radiation Heat Transfer Analysis

Mehran Yarahmadi
Department of Mechanical Engineering,
Virginia Tech,
Blacksburg, VA 24061

J. Robert Mahan[1]
Department of Mechanical Engineering,
Virginia Tech,
Blacksburg, VA 24061

Kevin McFall
Department of Mechatronics Engineering,
Kennesaw State University,
Kennesaw, GA 30144

In the Monte Carlo ray-trace (MCRT) method, millions of rays are emitted and traced throughout an enclosure following the laws of geometrical optics. Each ray represents the path of a discrete quantum of energy emitted from surface element i and eventually absorbed by surface element j. The distribution of rays absorbed by the n surface elements making up the enclosure is interpreted in terms of a radiation distribution factor matrix whose elements represent the probability that energy emitted by element i will be absorbed by element j. Once obtained, the distribution factor matrix may be used to compute the net heat flux distribution on the walls of an enclosure corresponding to a specified surface temperature distribution. It is computationally very expensive to obtain high accuracy in the heat transfer calculation when high spatial resolution is required. This is especially true if a manifold of emissivities is to be considered in a parametric study in which each value of surface emissivity requires a new ray-trace to determine the corresponding distribution factor matrix. Artificial neural networks (ANNs) offer an alternative approach whose computational cost is greatly inferior to that of the traditional MCRT method. Significant computational efficiency is realized by eliminating the need to perform a new ray trace for each value of emissivity. The current contribution introduces and demonstrates through case studies estimation of radiation distribution factor matrices using ANNs and their subsequent use in radiation heat transfer calculations.
[DOI: 10.1115/1.4047052]

Keywords: artificial neural network, Monte Carlo ray-trace method, radiation heat transfer

1 Motivation

The Monte Carlo ray-trace (MCRT) method [1–3] has emerged as the dominant tool for formulating high-fidelity models of radiation heat transfer processes. This is because of its universal applicability to problems involving radiant exchange among surfaces and within participating media, the ease with which it conforms to complicated irregular geometries, and its ability to treat directional and wavelength-dependent optical properties. Recently, its value has been significantly enhanced by contributions which establish a statistically meaningful paradigm for estimating the uncertainty, to a stated level of confidence, of predicted heat transfer results [4,5]. A widely lamented disadvantage of the MCRT method is the excessive computational cost associated with achieving high accuracy when fine spatial resolution is required. The fact that rays are mutually independent entities permits massive parallelization, with a proportionate reduction in processor time; however, associated cost, power, volume, and weight penalties exclude massive parallelization in applications where real-time results are required for data interpretation and decision-making on board autonomous space probes [6] and fire-and-forget weapons [7]. The alternative to a slow or computationally ponderous high-fidelity model in such applications would be a reduced-order model that provides comparable accuracy and spatial resolution but in real-time and with significantly reduced hardware requirements [8]. The current contribution describes such an alternative.

2 Brief Review of the Monte Carlo Ray-Trace Method

The MCRT method as commonly practiced by the radiation heat transfer community begins with the creation of a matrix of *radiation distribution factors*, defined here and elsewhere [2,3] as

$$D_{ijk} \equiv Q_{a,ijk}/Q_{e,ik}, \quad 1 \le i \le n, \quad 1 \le j \le n, \quad 1 \le k \le K \quad (1)$$

In Eq. (1), n is the number of surface and volume elements making up the enclosure; K is the number of wavelength intervals, or bands; $Q_{e,ik}$ is the power emitted by surface or volume element i in band k; and $Q_{a,ijk}$ is the power emitted by surface or volume element i that is absorbed by surface or volume element j in band k. Equation (1) is completely general in that it holds whether the radiative interchange is among surface or volume elements or a combination of the two, and for spectral directional radiation as well as for gray diffuse radiation.

It is perhaps worth noting that the distribution factor considered in the current contribution is distinctly different from the *geometrical factor*—also referred to as the *angle factor*, the *view factor*, and the *configuration factor*—which dominated radiation heat transfer pedagogy and practice in the second half of the 20th C [9–17]. Several such factors have been defined and used down through the years to calculate radiant exchange, but the distribution factor defined by Eq. (1) lies at the heart of the MCRT method. The earliest mention of this quantity is attributed to Gebhart, who refers to it as the *absorption factor* in a 1961 article [18]. Gebhart showed that, for the special case of a gray diffuse enclosure, the elements of his absorption factor matrix could be constructed from surface properties and angle factors. In 1968, Howell [19] introduced the term *exchange fraction* for the version of the absorption factor evaluated using the Monte Carlo method. Later, Mahan and Eskin [20,21] refer to this same quantity as the radiation distribution factor because of its essential role in distributing radiation emitted by surface or volume element i to surface or volume element j. While this latter term is in common usage, other authorities refer to the distribution factor as the exchange factor [22–24], although the exchange factor used by Yuen [24] is more directly akin to Gerhart's absorption factor since it is evaluated analytically without recourse to ray tracing. Finally, Larsen and Howell [25] attribute the term exchange factor to a family of auxiliary factors that, when used together, describe radiative exchange in the zonal method. It should also be noted that Lin and Sparrow [26] use the term exchange factor to describe radiant interchange among a mixture of diffuse and specular surfaces.

If the enclosure contains a participating medium and the first N out of n are surface elements and the remaining $n - N$ are volume elements, then the power absorbed by surface or volume element i is

$$Q_{i,a} = \Delta A_i \sum_{k=1}^{K} \left[\varepsilon_{ik} \sum_{j=1}^{N} e_b(\Delta \lambda_k, T_j) D_{ijk} \right]$$
$$+ 4\pi \Delta V_i \sum_{k=1}^{K} \left[\kappa_{ik} \sum_{j=N+1}^{n} i_b(\Delta \lambda_k, T_j) D_{ijk} \right] \quad (2)$$

and the power emitted by surface or volume element i is

$$Q_{i,e} = \begin{cases} \Delta A_i \sum_{k=1}^{K} \varepsilon_{ik} e_b(\Delta \lambda_k, T_i), & 1 \le i \le N \\ 4\pi \Delta V_i \sum_{k=1}^{K} \kappa_{ik} i_b(\Delta \lambda_k, T_i), & N+1 \le i \le n \end{cases} \quad (3)$$

In Eqs. (2) and (3), ΔA_i is the area of surface element i, ΔV_i is the volume of volume element i, ε_{ik} is the emissivity of surface element i in band $\Delta \lambda_k$, κ_{ik} is the absorption coefficient of volume element i in band $\Delta \lambda_k$, $e_b(\Delta \lambda_k, T_j)$ is the blackbody emissive power in band $\Delta \lambda_k$ corresponding to temperature T_j, $i_b(\Delta \lambda_k, T_j)$ is the blackbody intensity in band $\Delta \lambda_k$ corresponding to temperature T_j, and D_{ijk} is the bandwise spectral radiation distribution factor in band $\Delta \lambda_k$. The net power emitted from surface or volume element i is then

$$Q_i = Q_{i,e} - Q_{i,a} \quad (4)$$

In the absence of a participating medium, if the surface temperature distribution T_j is specified across the surfaces of a diffuse gray enclosure, Eqs. (2)–(4) reduce to

$$q_i = \varepsilon_i \sum_{j=1}^{n} \sigma T_j^4 (\delta_{ij} - D_{ij}), \quad 1 \le i \le n \quad (5)$$

where

$$\delta_{ij} \equiv \begin{cases} 1, & i = j \\ 0, & i \ne j \end{cases} \quad (6)$$

In Eq. (5), ε_i is the emissivity of surface element i and n is the number of surface elements making up the enclosure. In the current contribution the surfaces of the enclosure are gray and diffuse, in which case Eq. (1) reduces to

$$D_{ij} \cong N_{ij}/N_i, \quad 1 \le i \le n, \quad 1 \le j \le n \quad (7)$$

where N_i is the number of equal-strength rays diffusely emitted from surface or volume element i, and N_{ij} is the number of those rays absorbed in surface or volume element j [2,3]. Equation (5) may be thought of as an approximation because Eq. (7) produces an estimate of D_{ij} the accuracy of which increases with the number of rays traced for a given number of surface elements n. The elements of the radiation distribution factor matrix are created by following the life cycles of a large number of rays whose behavior is governed by application of statistical principles to the laws of geometrical optics. The details of ray tracing are widely available in the literature, including in Refs. [2] and [3].

Concerns expressed elsewhere [23] about the perceived need to "smooth" the exchange or distribution factor matrix are unwarranted when its elements are estimated using the MCRT method. The use of Eq. (7) to compute D_{ij} ensures that both conservation of energy and reciprocity are satisfied to a high level of accuracy. Uncertainties in the values of individual elements resulting from a finite number of rays being traced are equivalent to slight local distortions of the enclosure geometry or minor local variations in the surface properties. The related surface heat flux uncertainties resulting from application of Eq. (5) are quantified in Refs. [4] and [5].

When the number of surface elements n is large and high accuracy is required, an exceedingly large number of rays must be traced, as already remarked in Sec. 1. Furthermore, because D_{ij} depends on the emissivity ε_i, computational costs can become excessive for optimization processes in which ε_i is a parameter. This motivates the search for a computationally less intensive approach.

3 Introduction to Artificial Neural Networks

3.1 Background. Artificial neural networks (ANNs) are nonlinear mapping systems with structures based on principles inspired by the human biological nervous system. They provide a fundamentally different approach from other numerical solution methods for forecasting the future. Artificial neural networks can accurately model the inherent relationship between sets of input and output data without reference to the underlying physical system, and yet they are able to consider all the parameters affecting the physical system. Various considerations such as nonlinearity, multiplicity of variables and parameters, and noisy and uncertain input and output values are easily dealt with. Artificial neural networks depend on neither prior knowledge of correlations nor recourse to iterative methods, but rather require only a population of input/output samples. These latter are used to train the neural network which, once trained, is able to produce meaningful outputs in response to the introduction of test inputs not used in training. Artificial neural networks consist of a large number of processing units, which run in parallel to achieve results whose accuracy is comparable to that obtained using computationally more expensive traditional approaches. They are also able to perform dynamic modeling and adaptive control tasks in the presence of abrupt changes in system parameters and imposed control signals. Complexities not easily treated by traditional approaches to thermal system analysis can be accurately modeled with significantly less computing time using an ANN.

Artificial neural networks have been under development for about four decades. They have been widely used in many engineering applications because of their ability to obtain solutions more easily, frequently with an accuracy comparable to that of higher order models [27]. In recent years, ANNs have been used in various thermal applications describing heat transfer in solar energy systems, design of steam generating plants, estimation of heating loads of buildings, waste heat recovery heat exchangers, and related performance prediction and dynamic control applications. Thibault and Grandjean [28] used an ANN for heat transfer data analysis. Parcheco-Vega et al. [29] applied ANNs for modeling the heat transfer phenomena in fin-tube refrigerating heat exchanger systems. An ANN algorithm was used by Bechtler et al. [30] to model the steady-state performance of a vapor–compression liquid heat pump. Lazrak et al. [31] modeled a dynamic absorption chiller using artificial neural networks. An ANN model was developed to predict the convective heat transfer coefficient during condensation of R134 in inclined tubes [32]. Chang et al. [33] predicted heat transfer of supercritical water using ANNs. Ye et al. [34] proposed a novel ANN model for predicting convective heat transfer in sCO$_2$. Kaya and Hajimirza designed a two-layer ANN surrogate model to estimate the optical absorptivity of the solar ultrathin organic cells [35,36]. Additional investigations of heat transfer using ANNs have also been reported [37,38].

The cited applications demonstrate that ANNs are often well suited to thermal analysis of engineering systems. This is especially true when performing a parametric study involving repetitive solution of a complex model, in which case it is desirable to accelerate the analysis without comprising the underlying physics.

Although a variety analytical and numerical approaches have been employed in radiation heat transfer analysis, to the best of our knowledge, ANN methods have yet to be applied in this area. This further motivates the present work, which investigates the

applicability of ANNs to the radiation heat transfer analysis of diffuse gray enclosures. As an introductory exercise, the computationally intensive MCRT method is used to compute the radiation distribution factors among the surface elements of a two-dimensional diffuse gray enclosure for a range of surface emissivity. Then a back-propagation algorithm is used to train an ANN based on these limited results. The ability of the much faster artificial neural network to accurately predict the distribution factor matrices corresponding to values of emissivity not used in the training cases is then evaluated. Various network configurations are investigated in a search for the optimal network. Once introduced, the method is then extended to increasingly complex problems.

3.2 Description of the Artificial Neural Network. An artificial neural network is an information processing paradigm consisting of a large number of simple processing elements called neurons, or nodes, organized in layers [39]. The node layers are organized into three groups: the input layer, one or more hidden layers, and an output layer. Each layer is occupied by a number of nodes, as illustrated in Fig. 1. All the nodes of each hidden layer are connected to all nodes of the previous and following layers by means of synaptic connectors. Each connector is characterized by a synaptic weight. The input layer is used to designate the parameters for the problem under consideration, while the output layer corresponds to the unknown variables characterizing the performance of the system. The weights of the connectors determine the relative importance of the signals from all the nodes in the previous layer. At each hidden-layer node, the node input consists of a sum of all the outputs of the nodes in the previous layer, each modified by an individual interconnector weight. At each hidden node, the node output is determined by an activation function, which performs nonlinear input–output transformations. The information treated by the connector and node operations is introduced at the input layer, and this propagates forward toward the output layer [40]. Such ANNs are known as feed-forward networks, which is the type used in this study. Figure 1 is a schematic representation of typical feed-forward architecture. The configuration shown has one input layer, two hidden layers, and one output layer.

The error at each output node can be determined by comparing the calculated feed-forward result with the results obtained from conducting the original MCRT-based numerical experiments. Training of the network adjusts its weights to minimize the errors between the ANN result and known output. The training procedure for feed-forward networks is known as the supervised back propagation (BP) learning scheme, where the weights and biases are adjusted layer by layer from th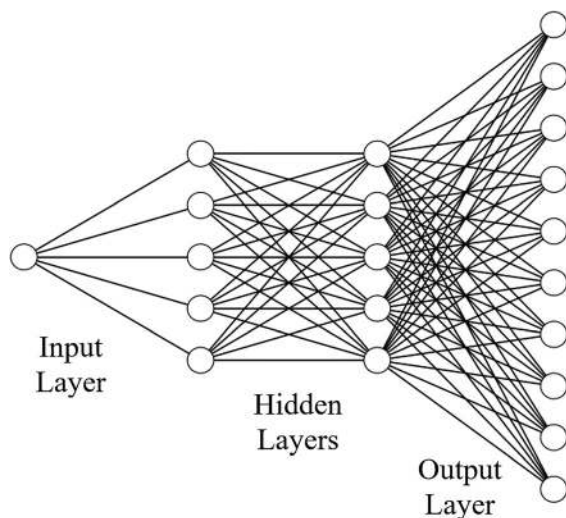e output layer toward the input layer [41]. The mathematical basis, the procedures for training and testing the ANNs, and more descriptions of the BP algorithm can be found elsewhere [42].

An a priori selection of ANN hyperparameters such as network topology, training algorithm, and network size is usually made based on experience. After training, the final sets of weights and biases trained by the network can be used for prediction purposes, and the corresponding ANN becomes a model of the input/output relation of the given problem. Because the ANN is to be trained to interpret the relationship between input and output data, the data used for training must be sufficient to capture the dynamics of the process being modeled. The MCRT method described in Sec. 2 is used to generate the training and test data needed to create and validate the ANN.

4 Implementation and Results

As a demonstration of the approach advanced in this contribution, we consider three case studies of increasing complexity. All three cases involve radiant exchange within an enclosure consisting of gray diffuse surfaces in the absence of a participating medium; that is, radiant exchange is governed by Eqs. (5)–(7). However, once the distribution factors have been computed using the MCRT method, the ANN approach advanced here is expected to work equally well in the presence of a participating medium and with directional spectral surface models.

4.1 Case Study 1: A Long Box Channel With Uniform Emissivity. Figure 2 represents a long square-cross-section box channel having uniform wall emissivity and prescribed wall temperatures. The walls have been subdivided into 40 equal-area segments in anticipation of an MCRT analysis. The corresponding ANN will have a single input node representing the emissivity ε, and 1600 output nodes representing the 1600 elements of the 40-by-40 radiation distribution factor matrix.

The first author has created a convenient windows application [43] that uses the MCRT method to compute the radiation distribution factors among any number of surface elements making up any two-dimensional diffuse gray enclosure representing a long duct or channel. The MCRT method for two-dimensional geometries is demonstrated in Ref. [44], and the uncertainties associated with the method are thoroughly established in Refs. [4] and [5]. In the current effort, we have used the application described in Ref. [43] to compute the distribution factors for the long square-cross section duct illustrated in Fig. 2. The duct walls are maintained at uniform temperatures of 300 and 500 K as shown in the figure, and the corresponding net heat flux distribution on the walls is sought. The duct has been subdivided into $n = 40$ longitudinal surface elements, and 100 numerical experiments were carried out covering the emissivity range $0.01 \leq \varepsilon \leq 1$. For each value of emissivity, two million rays were traced per surface element to obtain estimates of the corresponding distribution factor matrices

Fig. 1 Configuration of a 1-5-5-10 artificial neural network
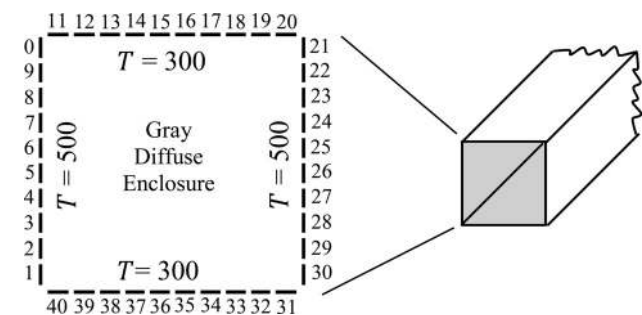

Fig. 2 The long square-cross-section duct having gray diffuse walls considered in case studies 1 and 2 (dashes represent individually numbered wall surface elements shown separated by nonexistent gaps for clarity)

D_{ij}, where $1 \leq i \leq 40$ and $1 \leq j \leq 40$. The resulting dataset was then randomly divided into training and test datasets. The training dataset used to regulate the weights on the ANN contained only 10% of the available data. The test dataset, consisting of the remaining 90% of the data, was used to evaluate the predictive ability of the ANN. While we recognize that it is more common to use the majority of the available data for training and a minority for testing, the ratio used here was found to give excellent results across the test data.

An Adam optimization algorithm is used in this study to converge the ANN output with the target data during the training process. This stochastic optimization method is straightforward to implement, is computationally efficient, has little memory requirements, is invariant to diagonal rescaling of the gradients, and is well suited for problems that are large in terms of data. Further details about Adam optimization can be found in Ref. [45].

Mean-squared error is used as the objective loss function for the ANN optimization. An inherent weakness of the BP algorithm is that it can converge to a local minimum. One way to avoid this tendency is to change the learning rate during the network training process. "Learning rate" refers to the rate of change of the neural network weights during optimization. The learning rate was ultimately set to 0.0002. Training of the neural network is terminated when a predetermined maximum number of training cycles have been completed. Selection of the maximum number is a trial-and-error process in which the number may be changed if the performance of the neural network during the training falls short of expectations. In this study 25,000 iterations were found to produce good results. The relative error of every predicted output is defined by

$$\mathrm{RE}_{ij} = \frac{|D_{ij}^{\mathrm{True}} - D_{ij}^{\mathrm{Pred}}|}{D_{ij}^{\mathrm{True}}} \tag{8}$$

where D_{ij}^{Pred} is the predicted result (that is, the output of the ANN) and D_{ij}^{True} is the result from the Monte Carlo ray-trace (that is, the target output). During the neural network training process, performance was evaluated by calculating the mean value of the relative error

$$\mathrm{MRE} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{RE}_{ij} \tag{9}$$

Reciprocity and conservation of energy are two important restrictions imposed on the radiation distribution factor [3]. To ensure that these two restrictions are reflected in the predicted result, values of mean reciprocity and mean summation, defined

$$\mathrm{MR} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\varepsilon_i A_i D_{ij}^{\mathrm{Pred}} - \varepsilon_j A_j D_{ji}^{\mathrm{Pred}}) \tag{10}$$

and

$$\mathrm{MS} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} D_{ij}^{\mathrm{Pred}} \tag{11}$$

are computed. For a good prediction, MR should be near zero and MS should be near unity. The relative difference between the calculated net heat flux recovered from the predicted and true radiation distribution factor matrix

$$\mathrm{MFD} = \frac{1}{n} \sum_{i=1}^{n} \frac{q_i^{\mathrm{True}} - q_i^{\mathrm{Pred}}}{q_i^{\mathrm{True}}} \tag{12}$$

is also evaluated.

Generalization is a term used to describe the ability of an ANN to provide accurate output results when input data that have not been used for training are introduced into the trained network. Generalization is an essential property of any ANN. The network topology and size, as determined by the number of hidden layers and the number of hidden nodes, will affect the predicted performance. The performance of the trained network is evaluated by comparing its predicted results with data set aside for testing. In this study, in order to facilitate the search for a configuration producing relatively good prediction, the ten different ANN configurations listed in Table 1 were considered.

Note that in Table 1, mean MR and mean MS are averaged over all the 10 training datasets and 90 test datasets, with different random weight initialization for each input. Both quantities are important for an assessment of the relative success of the ANN analysis. We can see from inspection of the table that almost any configuration produces adequate results; however, some of them result in poor generalization. For example, the 1-10-1600 configuration produces a mean MRE error of about 15% for the test data despite the low error of 2.5% for the training data. All of the configurations yield the required reciprocity and conservation of energy properties of radiation distribution factors. For the three-layer ANN, when the number of hidden nodes is increased from five to ten, improvements in mean MRE and mean MFD are insignificant, indicating that increasing the number of nodes does not necessarily lead to better performance. For selecting the best configuration, the mean MFD for the test data in conjunction with the mean MRE for the training and test data are both taken into consideration, leading to selection in the current example of the 1-20-20-1600 configuration (shown in bold type in the table).

The predicted radiation distribution factor matrix for a sample of the test data corresponding to $\varepsilon = 0.75$ is represented in Fig. 3. The printed values of D_{ij} are too small to read in the image, but the color shading, for which bright red indicates the maximum value ($D_{1,40} = D_{40,1} = D_{10,11} = D_{11,10} = D_{20,21} = D_{21,20} = D_{30,31} = D_{31,30} \approx 0.2256$) and dark green represents the minimum value ($D_{1,10} = D_{10,1} = D_{11,20} = D_{20,11} = D_{21,30} = D_{30,21} = D_{31,40} = D_{40,31} \approx 0.0049$), very clearly reveals the expected symmetry in the matrix.

Table 1   Comparison of errors associated with various ANN configurations for case study 1

ANN configuration	Train error				Test error			
	Mean MRE (%)	Mean MR	Mean MS	Mean MFD (%)	Mean MRE (%)	Mean MR	Mean MS	Mean MFD (%)
1-5-1600	2.498	9.3×10^{-5}	1.0000	0.1921	5.399	7.3×10^{-5}	0.9999	0.1923
1-10-1600	2.531	8.3×10^{-5}	1.0000	0.1826	15.405	5.6×10^{-4}	1.0011	0.4651
1-20-1600	1.944	4.0×10^{-5}	1.0000	0.1369	4.716	4.0×10^{-5}	1.0000	0.1738
1-50-1600	0.375	6.6×10^{-5}	1.0000	0.0487	9.035	4.3×10^{-4}	1.0003	0.3110
1-5-5-1600	2.524	1.0×10^{-5}	1.0000	0.1752	5.445	1.0×10^{-5}	1.0000	0.1809
1-5-10-1600	2.526	1.5×10^{-6}	1.0000	0.1749	9.484	2.4×10^{-4}	0.9991	0.3634
1-10-20-1600	0.508	5.5×10^{-5}	1.0000	0.0763	2.863	7.6×10^{-5}	1.0000	0.1196
1-20-20-1600	**0.707**	**5.3×10^{-5}**	**1.0000**	**0.0693**	**1.831**	**5.8×10^{-5}**	**0.9999**	**0.0951**
1-50-50-1600	0.0264	1.2×10^{-5}	1.0000	0.0481	4.628	1.3×10^{-4}	1.0008	0.1631

Note: For selecting the best configuration, the mean MFD for the test data in conjunction with the mean MRE for the training and test data are both taken into consideration, leading to selection in the current example of the 1-20-20-1600 configuration (shown in bold type in the table).
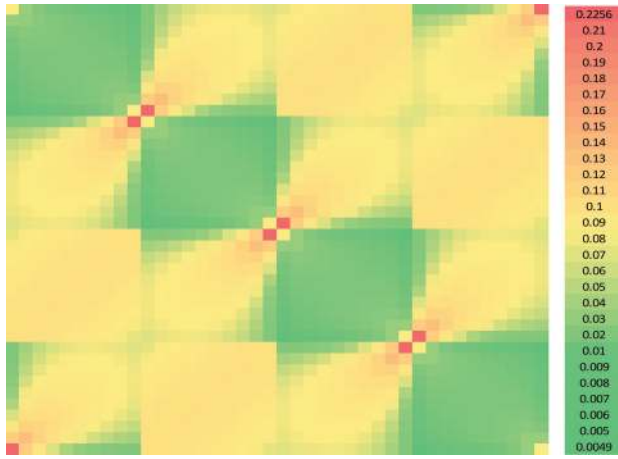
Fig. 3 Radiation distribution factor matrix produced by the ANN approach for a sample of the test data corresponding to $\varepsilon = 0.75$ (case study 1)

Finally, Fig. 4 compares the MCRT-based and ANN-based net heat flux distributions on the four surfaces of the enclosure depicted in Fig. 2 corresponding to the same ANN test case whose distribution factor matrix is shown in Fig. 3. The expected symmetry in the net heat flux distribution is evident, and excellent agreement is exhibited between the two approaches, with the
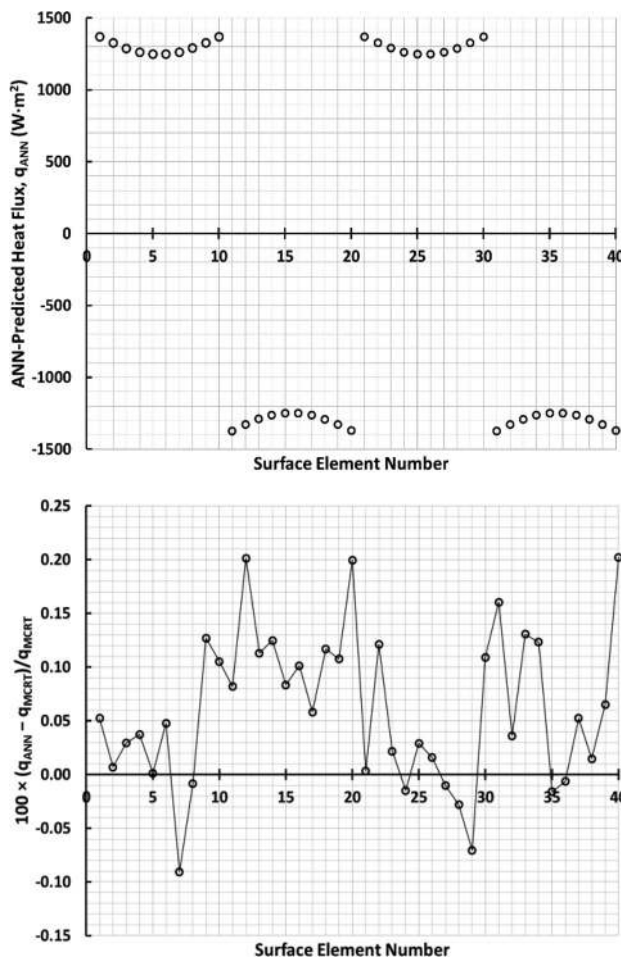
Fig. 4 Comparison of the MCRT-based and ANN-based heat flux predictions (case study 1). The upper graph verifies the expected symmetry and the lower graph shows the percentage error between ANN and MCRT.

relative difference between them typically of the order of 0.1%. It is clear that the ANN approach is a potentially powerful alternative to costly ray tracing in radiation heat transfer analysis requiring a parametric study of surface emissivity. For example, once the investment in creating and training the ANN model has been made, the time required to create the data in Fig. 3 is measured in seconds as opposed to hours on a typical desktop computer using the MCRT method.

4.2 Case Study 2: A Long Box Channel With Nonuniform Emissivity. In many radiation heat transfer applications of practical interest, surface emissivities vary with both position and instance due to the strong heterogeneity of surface properties associated with topography, surface chemistry, contamination, and aging. Here, "instance" refers both to changes that take place over time for a given enclosure and to differences from one enclosure to the next associated with manufacturing tolerances. In practice, it is unlikely that the surfaces comprising any two enclosures will have exactly the same emissivities. Methods for monitoring the temporal variations of emissivity are discussed elsewhere [46,47]. Here, as a practical demonstration, we allow the surface emissivity of the four walls of the enclosure geometry shown in Fig. 1 to vary from one instance to the next. It is assumed for each instance that each wall has a different but uniform emissivity. Therefore, the emissivity of the enclosure for any instance is represented as a vector whose elements are the emissivities of the four walls. In order to use the ANN to predict the evolution of the radiation distribution factor matrix either with time or from one enclosure to the next, training data are produced using the MCRT method with a 5% random perturbation of the emissivity of each wall about a mean value. In other words, in each instance, the emissivities of the four walls are randomly perturbed according to

$$\varepsilon_i = \varepsilon_d(1 \pm 0.05 \times \text{rand}), \quad i = 1, 2, 3, 4 \quad (13)$$

where ε_d is the design wall emissivity and rand is a uniformly distributed random number between zero and unity. In this case, four ANN input nodes are used corresponding to the four emissivities ε_i, while the 1600 output nodes still correspond to the 1600 elements of the 40-by-40 radiation distribution factor matrix. Again, 100 numerical experiments were carried out to produce data. Twenty percent of the data was used to train the neural network. The remaining 80% of the data was used as the test data to validate the predictive power of the network.

Table 2 shows the ANN results for the 4-100-100-1600 configuration selected for the case under consideration. The results are not as satisfactory as in the case with only one emissivity as the input feature. We can see that the mean test error associated with predicting the radiation distribution factors is quite high; however, they still allow accurate prediction of the net heat fluxes. Also, a mean MR of 3.2×10^{-4} and a mean MS of 1.0006 show that the model is yielding the required reciprocity and conservation of energy. The relatively high mean MRE error in the test data is due to a small number of high errors in relatively few pixels for a small number of samples.

Figure 5 compares the percent absolute relative error in the radiation distribution factors, defined

$$\frac{|D_{ij}^{\text{MCRT}} - D_{ij}^{\text{ANN}}|}{D_{ij}^{\text{MCRT}}} \times 100\%$$

for relatively "good" and relatively "bad" distribution factor matrix predictions using the ANN. Both results are drawn from the test dataset used in constructing Table 2. The red-tinted cells in the bottom (bad) matrix of Fig. 5, which correspond to errors exceeding 2%, reveal that some elements of the radiation distribution factor matrix are predicted with relatively poor accuracy. However, these large relative errors correspond to small values of D_{ij} (see Fig. 3) so that the relative errors are disproportionately

Table 2 Errors associated with 4-100-100-1600 ANN configuration for case study 2

Train error				Test error			
Mean MRE (%)	Mean MR	Mean MS	Mean MFD (%)	Mean MRE (%)	Mean MR	Mean MS	Mean MFD (%)
1.119	1.5×10^{-4}	1.0004	0.5493	34.867	3.2×10^{-4}	1.0006	0.7489
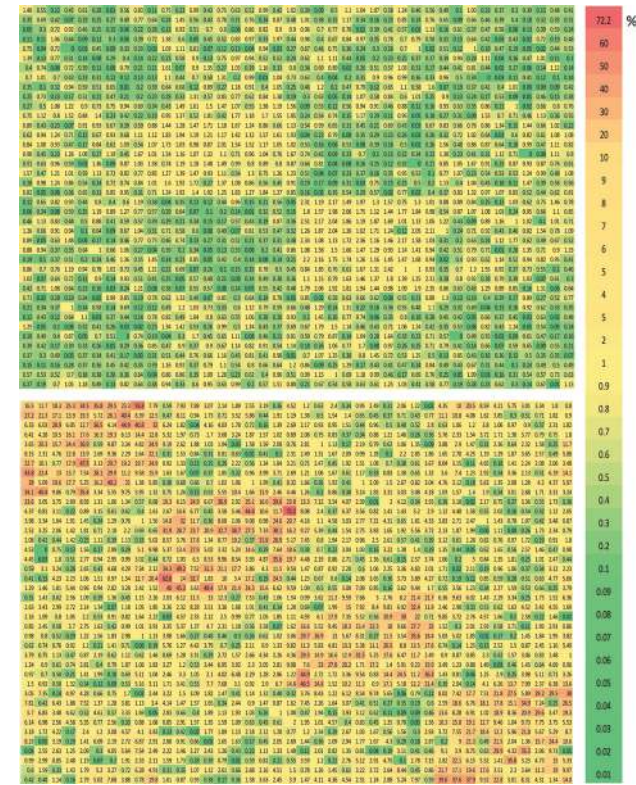
Fig. 5 Comparison of the percent absolute relative error in the radiation distribution factors between a relatively "good" (top) and a relatively "bad" (bottom) distribution factor matrix predicted using the ANN (case study 2)
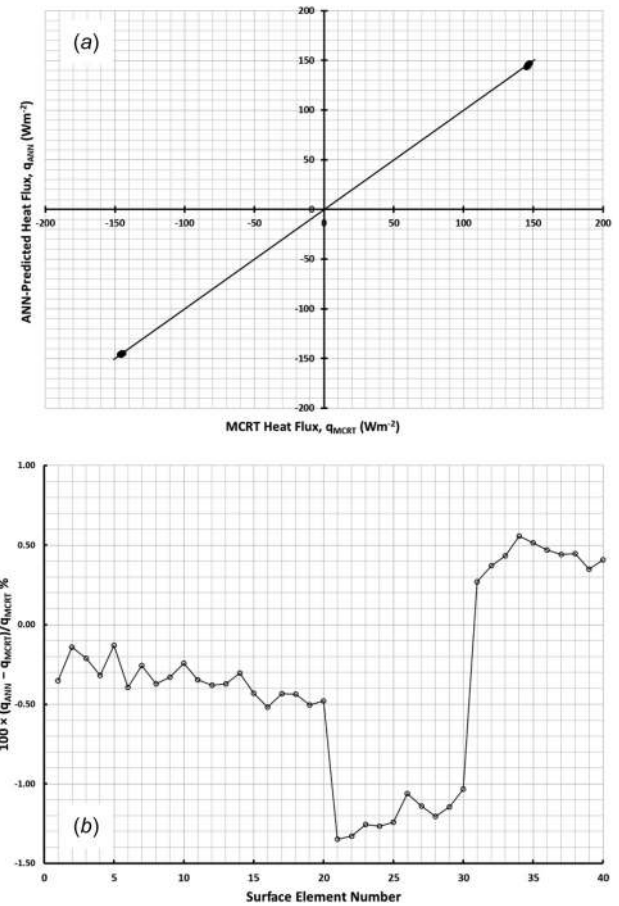

Fig. 6 (a) The ANN-based net heat flux distribution plotted against the MCRT-based distribution, and (b) the local percentage differences between the two distributions corresponding to the "good" test dataset result of case study 2

amplified due to division by small numbers. Although this produces a large value of mean MRE for the test dataset error in Table 2, the small values of these distribution factors themselves minimize their effect on the heat flux analysis, thereby yielding a small value of mean MFD.

Figure 6(a) shows good agreement between the net heat fluxes predicted using the ANN-based and MCRT-based distribution factor matrices corresponding to the upper panel of Figs. 5, and 6(b) confirms that the local net heat flux errors are less than 1.5% in this case. Figure 7(a) also reveals good agreement between the ANN-based and MCRT-based net heat flux distributions even though the distribution factor matrix prediction is relatively bad, and the local net heat flux errors shown in Fig. 7(b), though somewhat larger than those in Fig. 6(b), are generally well under 2% in this case. We may conclude that the ANN approach works well for the case of a nonuniform emissivity distribution. Furthermore, while minimizing mean MRE is a valid strategy for defining the ANN hyperparameters, its value should not be interpreted as a measure of the ability of the ANN-produced distribution factor matrix to predict local net heat flux.

4.3 Case Study 3: A Long Box Channel With an Interior Obstruction. In cases 1 and 2, we considered a geometry in which all wall segments have a direct view of all other wall segments. We now consider a more complex geometry involving an

interior obstruction, which partially blocks the direct view of some surfaces from other surfaces. In such cases, the MCRT method is the only practical approach for analyzing the radiation heat transfer. Howell was among the first to predict the emerging dominance of the Monte Carlo method for treating radiative heat transfer [1] in such cases. Figure 8 represents a benchmark two-dimensional enclosure that has been used in previous radiation heat transfer studies [48,49]. In this study, we have divided it into 40 equal-area longitudinal surface elements.

The flexibility of the Monte Carlo method to accommodate complex geometries comes at a significant computational cost when the code must be executed many times in the context of a parametric study; e.g., when searching for an optimum value of emissivity for a given application. This cost can be significantly reduced by replacing the high-fidelity MCRT model with a reduced-order ANN model of comparable accuracy in the search algorithm.

Once again assuming that the emissivity is uniform across all the walls of the enclosure, the ANN has only a single input node, corresponding to the emissivity, while 1600 output nodes are required to represent the 1600 elements of the radiation
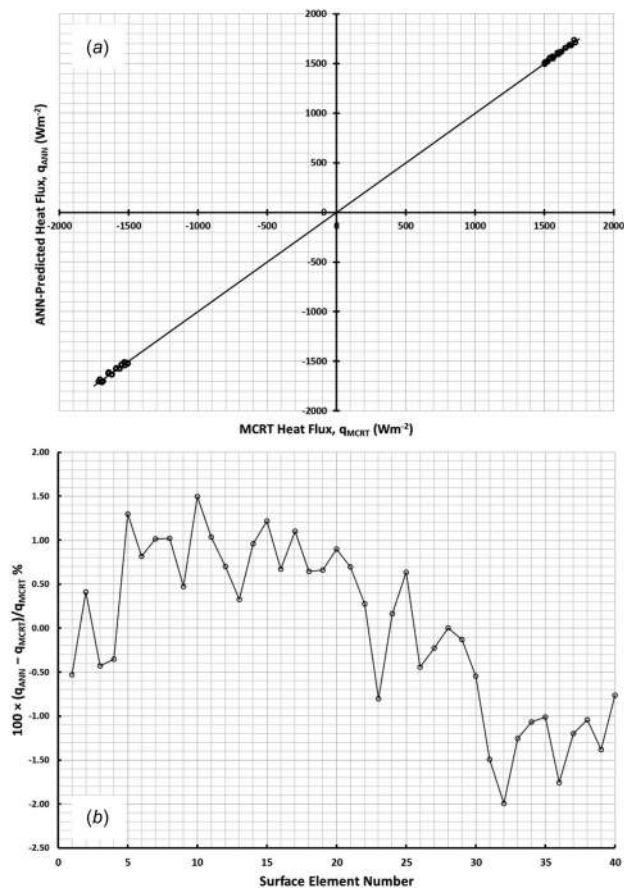
Fig. 7 (*a*) The ANN-based net heat flux distribution plotted against the MCRT-based distribution, and (*b*) the local percentage differences between the two distributions corresponding to the "bad" test dataset result of case study 2
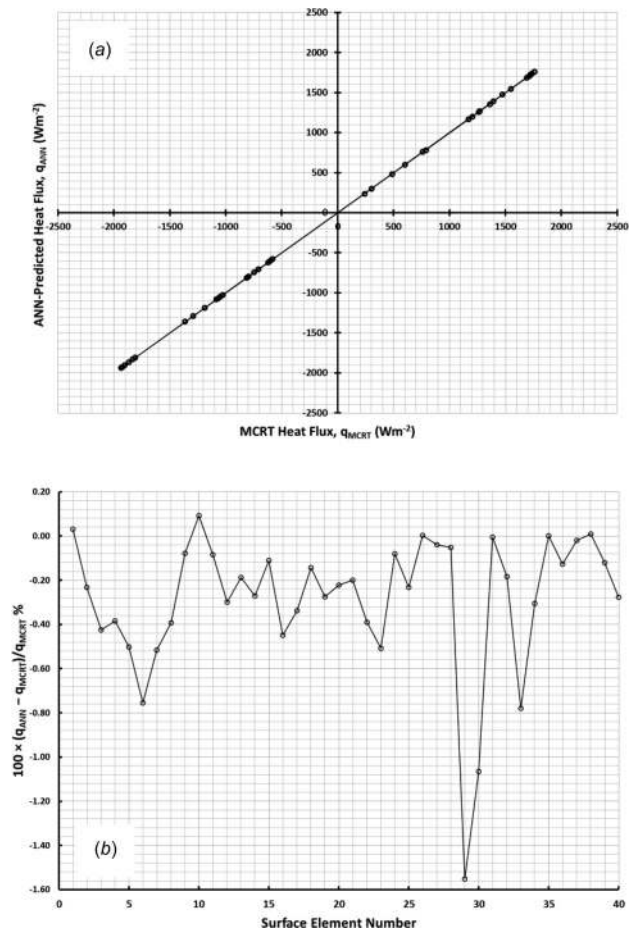

Fig. 9 (*a*) The ANN-based net heat flux distribution plotted against the MCRT-based distribution, and (*b*) the local percentage differences between the two distributions (case study 3, $\varepsilon = 0.75$)
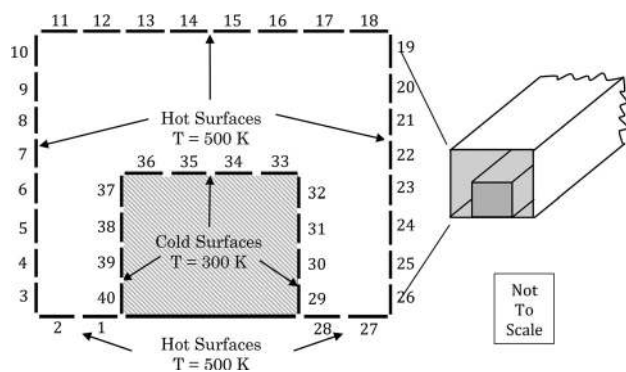

Fig. 8 The long two-dimensional duct having gray diffuse walls (dashes represent individually numbered wall surface elements shown separated by nonexistent gaps for clarity) and an interior obstruction considered in case study 3

distribution factor matrix. One hundred numerical experiments were carried out to produce training and test datasets and, as before, 10% of the data were used to train the neural network, with the remaining 90% used as the test data to validate the predictive power and generality of the ANN.

Table 3 shows the ANN results for the 1-100-100-1600 configuration selected for this case study. We again see that the MRE error for the test dataset, about 16.6% in this case, is not a measure of the ability of the distribution factors to predict the net heat flux distribution. The ANN model yields the required reciprocity and obeys conservation of energy to a high degree of accuracy.

Figure 9(*a*) compares the MCRT-based and ANN-based net heat flux distributions on the surfaces of the enclosure depicted in Fig. 8 for a uniform emissivity of 0.75, and Fig. 9(*b*) shows the relative difference between the values calculated for the net heat fluxes by the two methods. The accuracy—generally better than 1%—is quite acceptable.

Table 3 Errors associated with 1-100-100-1600 ANN configuration for case study 3

Train error				Test error			
Mean MRE (%)	Mean MR	Mean MS	Mean MFD (%)	Mean MRE (%)	Mean MR	Mean MS	Mean MFD (%)
1.784	6.0×10^{-4}	1.0000	0.0969	16.627	2.4×10^{-4}	1.0004	0.3284

5 Conclusions and Recommendations

Artificial neural networks are investigated as an alternative to ray tracing in radiation heat transfer applications involving diffuse gray enclosures in the absence of a participating medium. Specifically, they are used to predict the radiation distribution factor matrix and corresponding net heat flux distribution on the walls of long box structures. In each case, a feed-forward back-propagation algorithm is used to train and test the ANN. Net heat flux results obtained using the ANN approach are shown to agree well with those obtained using the standard Monte Carlo ray-trace method for the three cases studied: (1) uniform emissivities on all walls of a square-cross section duct, (2) differing emissivities from wall to wall perturbed about a design value for the same unobstructed duct, and (3) a rectangular duct containing a rectangular obstruction with uniform emissivity on all walls. The authors recommend the approach introduced here when a parametric study is required to determine the optimum value of emissivity for a given application. For example, the results for case study 2, obtained with much less computational effort than would have been required using the MCRT method alone, could be used in a quality-control scheme to determine the variability in the net wall heat flux corresponding to a 5% manufacturing tolerance in wall emissivity. The ANN approach would be the same for the case of a nondiffuse, nongray enclosure filled with a participating medium as for the case of a diffuse gray enclosure in the absence of a participating medium demonstrated in the current effort. This encourages the idea that the approach advanced here would be equally applicable—and even more useful—in these far more complex situations.

Nomenclature

A = area (m^2)
D = radiation distribution factor
e = emissive power (W m^{-2})
i = intensity (W m^{-2} sr^{-1})
K = number of wavelength intervals
MFD = mean relative flux difference
MR = mean reciprocity
MRE = mean relative error
MS = mean summation
n = number of surface and volume elements making up the enclosure
N = number of surface elements in Eqs. (2) and (3), number of equal-strength rays traced in Eq. (7)
q = flux (W m^{-2})
Q = power (W)
rand = random number
RE = relative error
T = temperature (K)
V = volume (m^3)

Greek Symbols

δ = Kronecker delta function
ε = emissivity
λ = wavelength (μm)
κ = volumetric absorption coefficient (m^{-1})

Subscripts

a = absorbed
b = blackbody
d = design
e = emitted
i,j = surface or volume element indices
k = wavelength band index

Superscripts

ANN = artificial neural network method
MCRT = Monte Carlo ray-trace method
Pred = predicted value
True = true value

References

[1] Howell, J. R., 1998, "The Monte Carlo Method in Radiative Heat Transfer," ASME J. Heat Transfer, **120**(3), pp. 547–560.
[2] Mahan, J. R., 2002, *Radiation Heat Transfer: A Statistical Approach*, Wiley, New York.
[3] Mahan, J. R., 2019, *The Monte Carlo Ray-Trace Method in Radiation Heat Transfer and Applied Optics*, Wiley-ASME Press, Hoboken, NJ.
[4] Yarahmadi, M., Mahan, J. R., and Priestley, K. J., 2019, "Uncertainty Analysis and Experimental Design in the Monte Carlo Ray-Trace Environment," ASME J. Heat Transfer, **141**(3), p. 032701.
[5] Yarahmadi, M., Mahan, J. R., and Priestley, K. J., 2019, "Estimation and Use of the Radiation Distribution Factor Median for Predicting Uncertainty in the Monte Carlo Ray-Trace Method," ASME J. Heat Transfer, **141**(6), p. 062701.
[6] Yarahmadi, M., Mahan, J. R., McFall, K., and Ashraf, A. B., 2020, "Numerical Focusing of a Wide-Field-Angle Earth Radiation Budget Imager Using an Artificial Neural Network," Remote Sens., **12**(1), p. 176.
[7] Lloyd, P., 2019, "U.S. Army is Working on AI-Guided Missiles That 'Pick Their OWN Targets' to Blow Up After Scanning Ground Below," Daily Mail, London, UK, accessed Jan. 7, 2020, https://www.dailymail.co.uk/sciencetech/article-7363823/US-Army-working-AI-guided-missiles-pick-targets.html
[8] Irick, K., 2019, "In-Situ Thermal ROM-Based Optimization Using Borg MOEA: A Preliminary Study," ASME Paper No. HT2019:3483.
[9] McAdams, W. H., 1954, *Heat Transmission*, McGraw-Hill Book Company, New York.
[10] Kreith, F., 1958, *Radiation Heat Transfer for Spacecraft and Solar Plant Design*, International Textbook Company, Scranton, PA.
[11] Jakob, M., 1957, *Heat Transfer*, Vol. II, Wiley, New York.
[12] Eckert, E. R. G., and Drake, R. M., 1959, *Heat and Mass Transfer*, McGraw-Hill Book Company, New York.
[13] Sparrow, E. M., and Cess, R. D., 1966, *Radiation Heat Transfer*, Brooks/Cole Publishing Company, Belmont, CA.
[14] Wiebelt, J. A., 1966, *Engineering Radiation Heat Transfer, Holt*, Rinehart and Winston, New York.
[15] Hottel, H. C., and Sarofim, A. F., 1967, *Radiative Transfer*, McGraw-Hill Book Company, New York.
[16] Love, T. J., 1968, *Radiative Heat Transfer*, Charles E. Merrill Publishing Company, Columbus, OH.
[17] Gebhart, B., 1971, *Heat Transfer*, 2nd ed., McGraw-Hill, New York.
[18] Gebhart, B., 1961, "Surface Temperature Calculations in Radiant Surroundings of Arbitrary Complexity—For Gray, Diffuse Radiation," Int. J. Heat Mass Transfer, **3**(4), pp. 341–346.
[19] Howell, J. R., 1968, "Application of Monte Carlo to Heat Transfer Problems," *Advances in Heat Transfer*, Vol. 5, Academic Press, New York.
[20] Mahan, J. R., and Eskin, L. D., 1981, "Application of Monte Carlo Techniques to Transient Thermal Modeling of Cavity Radiometers Having Diffuse-Specular Surfaces," Fourth American Meteorological Association Conference on Atmospheric Radiation, Toronto, ON, Canada, June 16–18, pp. 181–186.
[21] Mahan, J. R., and Eskin, L. D., 1984, "The Radiation Distribution Factor—Its Calculation Using Monte Carlo Techniques and an Example of Its Application," First UK National Heat Transfer Conference, Leeds, Yorkshire, UK, July 3–5, pp. 1001–1012.
[22] Loehrke, R. I., Dolaghan, J. S., and Burns, P. J., 1995, "Smoothing Monte Carlo Exchange Factors," ASME J. Heat Transfer, **117**(2), pp. 524–526.
[23] Daun, K. J., Morton, D. P., and Howell, J. R., 2005, "Smoothing Monte Carlo Exchange Factors Through Constrained Maximum Likelihood Estimation," ASME J. Heat Transfer, **127**(10), pp. 1124–1128.
[24] Yuen, W., 2006, "The Multiple Absorption Coefficient Zonal Method (MACZM), an Efficient Computational Approach for the Analysis of Radiative Heat Transfer in Multidimensional Inhomogeneous Nongray Media," Numer. Heat Transfer, Part B, **49**(2), pp. 89–103.
[25] Larsen, M. E., and Howell, J. R., 1985, "The Exchange Factor Method: An Alternative Basis for Zonal Analysis of Radiating Enclosures," ASME J. Heat Transfer, **107**(4), pp. 936–942.
[26] Lin, S. H., and Sparrow, E. M., 1965, "Radiant Interchange Among Curved Specularly Reflecting Surfaces—Application to Cylindrical and Conical Cavities," ASME J. Heat Transfer, **87**(2), pp. 299–307.

[27] Lin, T. Y., and Tseng, C. H., 2000, "Optimum Design for Artificial Neural Networks: An Example in a Bicycle Derailleur System," Eng. Appl. Artif. Intell., **13**(1), pp. 3–14.

[28] Thibault, J., and Grandjean, B. P. A., 1991, "A Neural Network Methodology for Heat Transfer Data Analysis," Int. J. Heat Mass Transfer, **34**(8), pp. 2063–2070.

[29] Pacheco-Vega, A., Sen, M., Yang, K. T., and McClain, R. L., 2001, "Neural Network Analysis of Fin-Tube Refrigerating Heat Exchanger With Limited Experimental Data," Int. J. Heat Mass Transfer, **44**(4), pp. 763–770.

[30] Bechtler, H., Browne, M. W., Bansal, P. K., and Kecman, V., 2001, "Neural Networks—A New Approach to Model Vapour-Compression Heat Pumps," Int. J. Energy Res., **25**(7), pp. 591–599.

[31] Lazrak, A., Boudehenn, F., Bonnot, S., Fraisse, G., Leconte, A., Papillon, P., and Souyri, B., 2016, "Development of a Dynamic Artificial Neural Network Model of an Absorption Chiller and Its Experimental Validation," Renewable Energy, **86**, pp. 1009–1022.

[32] Azizi, S., and Ahmadloo, E., 2016, "Prediction of Heat Transfer Coefficient During Condensation of R134a in Inclined Tubes Using Artificial Neural Network," Appl. Therm. Eng., **106**, pp. 203–210.

[33] Chang, W., Chu, X., Fareed, A. F. B. S., Pandey, S., Luo, J., Weigand, B., and Laurien, E., 2018, "Heat Transfer Prediction of Supercritical Water With Artificial Neural Networks," Appl. Therm. Eng., **131**, pp. 815–824.

[34] Ye, K., Zhang, Y., Yang, L., Zhao, Y., Li, N., and Xie, C., 2019, "Modeling Convective Heat Transfer of Supercritical Carbon Dioxide Using an Artificial Neural Network," Appl. Therm. Eng., **150**, pp. 686–695.

[35] Kaya, M., and Hajimirza, S., 2018, "Surrogate Based Modeling and Optimization of Plasmonic Thin Film Organic Solar Cells," Int. J. Heat Mass Transfer, **118**, pp. 1128–1142.

[36] Kaya, M., and Hajimirza, S., 2018, "Application of Artificial Neural Network for Accelerated Optimization of Ultra-Thin Organic Solar Cells," Sol. Energy, **165**, pp. 159–166.

[37] Mohanraj, M., Jayaraj, S., and Muraleedharan, C., 2015, "Applications of Artificial Neural Networks for Thermal Analysis of Heat Exchangers—A Review," Int. J. Therm. Sci., **90**, pp. 150–172.

[38] Kamble, L., Pangavhane, D., and Singh, T., 2014, "Heat Transfer Studies Using Artificial Neural Network—A Review," Int. Energy J., **14**(1), pp. 25–42.

[39] Piuri, V., and Alippi, C., 1998, "Artificial Neural Networks," J. Syst. Archit., **44**(8), pp. 565–567.

[40] Basheer, I. A., and Hajmeer, M., 2000, "Artificial Neural Networks: Fundamentals, Computing, Design, and Application," J. Microbiol. Methods, **43**(1), pp. 3–31.

[41] Teng, F. C., 2001, *Application of Neural Networks to Adaptive Control of Nonlinear Systems, by GW Ng*, Research Studies Press, Taunton, UK, p. 198. 1997, Int. J. Robust Nonlinear Control: IFAC-Affil. J., **11**(9), pp. 879–880.

[42] Haykin, S., 1994, *Neural Networks: A Comprehensive Foundation*, Prentice Hall PTR, Upper Saddle River, NJ.

[43] Yarahmadi, M., 2018, *An Engine for Computing Radiation Distribution Factors for Two-Dimensional Enclosures*, Zenodo, Geneva, Switzerland.

[44] Yarahmadi, M., and Mahan, J. R., 2018, "Verification of Two-Dimensional Monte Carlo Ray-Trace Methodology in Radiation Heat Transfer Analysis," arXiv preprint arXiv:1810.05204.

[45] Kingma, D. P., and Ba, J., 2014, "Adam: A Method for Stochastic Optimization," arXiv preprint arXiv:1412.6980.

[46] Raj, V. C., and Prabhu, S. V., 2017, "A Refined Methodology to Determine the Spatial and Temporal Variation in the Emissivity of Diffusion Flames," Int. J. Therm. Sci., **115**, pp. 89–103.

[47] Li, Z. L., Becker, F., Stoll, M. P., and Wan, Z., 1999, "A Method for Monitoring the Temporal Variation of Surface Spectral Emissivity: Application to Thermal Infrared Multispectral Scanner (TIMS) Data in HAPEX-Sahel," J. Geophys. Res.: Atmos., **104**(D24), pp. 31217–31224.

[48] Mosavati, M., Kowsary, F., and Mosavati, B., 2013, "A Novel, Noniterative Inverse Boundary Design Regularized Solution Technique Using the Backward Monte Carlo Method," ASME J. Heat Transfer, **135**(4), p. 042701.

[49] Yarahmadi, M. J., Mahan, R., and Kowsary, F., 2019, "A New Approach to Inverse Boundary Design in Radiation Heat Transfer," Proceedings of the 16th UK Heat Transfer Conference (UKHTC2019), Nottingham, UK, Sept. 8–9, pp. UKHTC2019–73.