A new method for modeling radiative heat transfer based on Bayesian artificial neural networks and Monte Carlo method in participating media

Alex Royer, Olivier Farges, Pascal Boulet, Daria Burot

A new method for modeling radiative heat transfer based on Bayesian artificial neural networks and Monte Carlo method in participating media

Alex Royer[a,b,*], Olivier Farges[a], Pascal Boulet[a], Daria Burot[b]

[a] Université de Lorraine, CNRS, LEMTA, F-54000 Nancy, France
[b] Safran Aircraft Engines, Etablissement de Villaroche, Rond Point René Ravaud, Moissy-Cramayel France

A B S T R A C T

A new method based on predictive capacity of Feedforward Artificial Neural Networks (FANN) is proposed to estimate the divergence of the radiative flux in an axisymmetric domain. Training and validation databases have been built thanks to results given by the SNB-CK model and computed accordingly with a null collision Monte Carlo algorithm. The major aim of this work is to combine advantages of spectral models in terms of accuracy and the computational efficiency of neural networks in order to make possible the accurate modeling of radiative heat transfer. As a result, ANNs are able to model the radiative flux divergence on the basis of training data and some keys to avoid the pitfalls related to ANNs are provided.

Keywords:
Artificial neural network
Monte Carlo method
Machine learning
Radiative heat transfer
Bayesian interpolation

1. Introduction

Radiative heat transfer for combustion applications is a major issue in coupled simulations. In an environment where temperatures and concentrations of radiative species are high, this heat transfer mode is predominant and has to be well-modeled. An accurate computational thermal analysis allows an optimal system design with regards to reduction of weight and fuel consumption, more accurate modeling of pollutant formation, increasing of global engine lifetime or more accurate maintenance plan among others. Unfortunately, accurate radiative heat transfer modeling requires a lot of computational resources whereas, in multiphysics modeling, most of them are used by Computational Fluid Dynamics (CFD) and combustion calculations. Artificial Neural Networks (ANN) have already been investigated in heat transfer applications. This tool has been used for heat transfer data analysis [1], for thermal engineering [2–5], for the calculation of heat transfer [6,7]. In more specific cases, neural networks have been studied several times for radiative transfer calculation. Yarahmadi et al. [8] used Monte Carlo method to generate data to train a neural network in order to model the radiative transfer within an enclosure with homogeneous and heterogeneous wall emissivities. Sun et al. [9] trained ANN to efficiently compute cross sections for SLW model. Mishra and Molinaro [10] used Physic Informed Neural Network to model radiative heat transfer. Neural networks combined with bayesian optimization has also been used to replace efficiently heavy lookup table in aim to predict k-distributions [11]. An original and effective use of neural networks for modeling radiative transfers at high temperature is proposed by André et al. [12]. The authors implement recurrent networks coupled to an l-distribution model. The authors succeed in giving physical meaning to the network parameters and removing the complex phase of defining the topology.

In this work, we use neural networks to directly predict the divergence of the radiative flux. We propose a coupling between a Monte Carlo algorithm for the generation of training points and a Bayesian method to train and determine the network topology. The reduced size of networks allows to reduce the training time and to suppress the topology optimization phase generally necessary when using this type of tool.

Combustion modeling in constrained environments is difficult because of the multi-scale and multi-physical phenomena that need to be taken into account and radiative heat transfer is only one of the critical points. Nevertheless, the search for a reduction in computation time can not be achieve by a loss of accuracy. Indeed, the major issue here is to solve with a high degree of accuracy transfer phenomena while limiting the use of com-

* Corresponding author.
E-mail address: alex.royer@univ-lorraine.fr (A. Royer).

Nomenclature

Greek Symbols

α, β	Regularization hyperparameters
δ	Dirac impulsion
ϵ_j^l	Sensitivity of the cost function towards the jth neuron of the lth layer
ε	Relative error
γ	Number of effective parameters
$\kappa_{abs,\eta}$	Monochromatic absorption coefficient (m^{-1})
$\kappa_{abs,\eta}^{soot}$	Monochromatic absorption coefficient of soot (m^{-1})
$\kappa_{n,\eta}$	Null collision coefficient (m^{-1})
$\widehat{\kappa_\eta}$	Monochromatic extinction coefficient (m^{-1})
η	Wavenumber (cm^{-1})
$\div q$	Divergence of the radiation flux ($W \cdot m^{-3}$)
σ	Activation function of the neuron
σ'	Derivative of the activation function of the neuron
τ_η	Monochromatic transmissivity
μ	Learning rate
ν_i	ith parameter to be minimized

Latin Symbols

a_j^l	Output of the jth neuron of the lth layer after activation function
b	Biais of a neuron
B, S	Coefficients depending on the mole fraction, the pressure and other parameters
C	Cost function
C_x	Cost of one training set
E_D	Mean squared error
E_W	Sum of squares of the network weights
f_v	Soot volume fraction
f(k)	Inverse of the Laplace transformation of the SNB gas transmissivity
g(k)	Cumulative of f(k) function
He	Hessian Matrix of the training set
$\mathbf{I}$	Identity matrix
$\mathbf{J}$	Jacobienne matrix
k	SNB gas transmissivity
$\mathcal{L}$	Isothermal and homogeneous path of a photon
L	Length of the geometry (m)
l	Free path of a photon (m)
I_η	Monochromatic radiation intensity ($W/m^{-3}/sr/cm$)
N	Number of training set
N_{mc}	Finite and large integer number
n	Total number of parameters in the network
p_X	Probability density function
P	Pressure (bar)
q	Wall radiation flux ($W \cdot m^{-2}$)
R	Radius of the geometry (m)
r	Radial coordinate (m)
r_i	Randomly generated number between 0 and 1
T	Temperature (K)
$\mathbf{u}$	Photon direction
$\mathbf{v}$	Error vector
w_j	jth weight of a neuron
$\mathbf{x}$	Photon location
x	Longitudinal coordinate (m)
x_j	jth input of a neuron
x_{H_2O}	H_2O mole fraction
x_{CO_2}	CO_2 mole fraction
x_k	Parameter of the network
z	Weighted sum at neuron output

Superscripts

b	Relative to black body radiative properties
L	Output network layer
l	lth network layer

puting resources in order to be able to use them to lift other important locks. ANN are complex and non-linear approximators and have shown in many other domains their ability to generalize multi-variant functions [13]. Here a Monte Carlo algorithm with a spectral model is coupled with an ANN to predict the radiative heat flux and its divergence over an entire field. The development of Monte Carlo method applied to radiative transfer problems dates back to the 1960s with the work of Howell and Perlmutter [14,15]. Benefiting from methodological and computer advances, the method is well mastered in applications of thermal radiation in participating media [16,17] where it is considered as the reference and validation method. The extension of this method to heterogeneous participating media through the null collision formulation [18–20] allows the efficient consideration of real application cases. However, although the Monte Carlo method is recognized as a reference method for radiative transfer, it is most effective for performing probe calculations. The use of ANNs makes it possible to overcome this limitation to obtain field values of the radiative quantities of interest. The main objective of this work is to demonstrate the multivariate regression capabilities of neural networks in the prediction of a complete field of a value of interest, based on a reduced number of training points. The results communicated in this paper, based on a benchmark case known in the radiative transfer community, are the basis for the construction of new accurate and efficient radiation modeling methodologies.

This paper is organized as follows:

- Section 2 is dedicated to the Radiative Transfer Equation (RTE) and its resolution by the Monte Carlo method in its classical and null-collision formulations.
- In Section 3 the gas model used and how soot has been taken into account are described.
- Section 4 focuses on Feedforward Artificial Neural Networks (FANN) and BackPropagation algorithm, followed by explanations about how this method is interfaced and connected with Monte Carlo algorithm to predict radiative heat flux distributions.
- In Sections 5 and 6, three well-known benchmark validation cases are presented and a validation of the Monte Carlo algorithm is performed.
- In Section 7, results obtained with this new method based on regression abilities of FANN are presented and compared with Monte Carlo reference results for different study cases.
- Finally, Section 7.3 presents a comparison between ANN performance and other interpolation methods.

2. Solving the radiative transfer equation (RTE) with a null collision Monte Carlo algorithm

2.1. Stochastic formulation of the radiative transfer equation (RTE)

Here we consider radiative transfer in participating and non-scattering media. The spectral absorption coefficient is highly dependent on many variables such as temperature, pressure, soot volume fraction or species concentrations. Due to a very different radiative behavior between gas and soot, we usually split the absorption coefficient expression into two distinct contributions. The models used for gas and soot are presented respectively in Section 3.1 and 3.2.

In order to solve the RTE, the choice was made to use the Monte Carlo method on the one hand because it is a reference in the relevant field and on the other hand because it has the advantage of providing information on the accuracy of the result obtained. Indeed, obtaining the statistical error associated with the quantity of interest makes it possible to control the quality of the results which will be provided to the neural networks thereafter. In this case, RTE is usually expressed in its integral form. Thus, the spectral transmissivity $\tau_\eta(\mathbf{x}, \mathbf{u}, l)$ is the probability for a photon at location $\mathbf{x}$ to travel a free path l in the direction $\mathbf{u}$ in the medium without being absorbed. If we consider that this probability is independent of the distance already travelled, in a non-scattering medium, we can define the transmissivity according to Beer-Lambert's law [21] [22, chap. 9.11] as described in Eq. (1).

$$\tau_\eta(\mathbf{x}, \mathbf{u}, l) = \exp\left(-\int_0^l \kappa_{abs,\eta}(\mathbf{x} + l'\mathbf{u})\mathrm{d}l'\right) \tag{1}$$

As the transmissivity is equal to the exponential of the optical thickness, we can express the radiation intensity at a location $\mathbf{x}_0$ in the direction $\mathbf{u}_0$ as the spatial integral over the backward path $]-\infty, \mathbf{x}_0]$ of an emission term attenuated according to the Beer-Lambert law.

$$I_\eta(\mathbf{x}_0, \mathbf{u}_0) = \int_{-\infty}^{\mathbf{x}_0} \kappa_{abs,\eta}(\mathbf{x})\mathrm{d}\mathbf{x} \exp\left(-\int_{\mathbf{x}}^{\mathbf{x}_0} \kappa_{abs,\eta}(\mathbf{x}')\mathrm{d}\mathbf{x}'\right) I_\eta^b(\mathbf{x}) \tag{2}$$

With this formulation, the spectral radiation intensity is written as the expected value of a random variable and can be solved with a Monte Carlo algorithm.

We can then reformulate Eq. (2) introducing the characteristic probability density function of the Beer-Lambert's law and thus obtaining a statistical formulation of the radiation intensity (Eq. (4)). Equation (5) is a probability density which governs the probability of a photon to be absorbed by the medium. This probability increases with the distance between $\mathbf{x}_0$ and $\mathbf{x}$ and the value of the absorption coefficient.

$$I_\eta(\mathbf{x}_0, \mathbf{u}_0) = \int_{-\infty}^{\mathbf{x}_0} p_X(\mathbf{x})\mathrm{d}\mathbf{x} I_\eta^b(\mathbf{x}) \tag{3}$$

With the probability density function :

$$p_X(\mathbf{x}) = \kappa_{abs,\eta}(\mathbf{x}) \exp\left(-\int_{\mathbf{x}}^{\mathbf{x}_0} \kappa_{abs,\eta}(\mathbf{x}')\mathrm{d}\mathbf{x}'\right) \tag{4}$$

$$I_\eta(\mathbf{x}_0, \mathbf{u}_0) = \mathbb{E}[I_\eta^b(\mathbf{x})] = \lim_{N_{mc} \to \infty} \left(\frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} I_\eta^b(\mathbf{x})\right) \tag{5}$$

In practice, N_{mc} is a finite and large enough number. As mentioned above, the control of uncertainty is probably the main strength of Monte Carlo method because it is always possible to give a confidence interval with a result of the simulation. If the variance is too high, it is possible to reduce it by increasing N_{mc}, or with more sophisticated techniques of variance reduction [22, 20.10]. The wall flux can be defined as the integral of all the spectral radiation intensities from the the incoming hemisphere. Numerically, this integral can be estimated as the sum of all the Monte Carlo realizations, according to Eq. (3):

$$q(\mathbf{x_0}) = \int_0^\eta \int_0^{2\pi} I_{\eta'}(\mathbf{x_0}, \mathbf{u}')d\eta' d\mathbf{u}' \tag{6}$$

The computation of the divergence of the flux $\div q$ corresponds to the integral of the net radiation intensity at point $\mathbf{x_0}$ over the space. Here, we consider an isotropic domain. The divergence of the flux can be expressed as follows:

$$\div q(\mathbf{x}_0) = \int_0^\eta d\eta' \int_0^{4\pi} \kappa_{abs,\eta}(\mathbf{x_0})\left[I_{\eta'}^0(\mathbf{x_0} - I_{\eta'}(\mathbf{x_0}, \mathbf{u}')\right]d\mathbf{u}' \tag{7}$$

2.2. Null-collision formulation

In a homogeneous medium, it is easy to calculate the mean free path l_i of the photons. An analytical solution exists and only depends on the absorption coefficient of the medium:

$$l_i = -\frac{\ln(1 - r_i)}{\kappa_{abs,\eta}} \tag{8}$$

Where r_i is a randomly generated number between 0 and 1. In a heterogeneous medium, it is no longer possible to define a mean free path. Indeed, it is no longer possible to integrate the optical thickness of the medium analytically [18]. In terms of resolution, this characteristic makes things very complex from a numerical point of view. There are techniques to overcome this difficulty, which usually involve either discretizing the space or inversing the optical thickness rather than the free path. In both cases and despite a very high computational cost, the exact nature of Monte Carlo method is lost because a numerical bias appears that is difficult to quantify. To overcome this limitation, a null collision algorithm is used [18]. As highlighted by Howell and Daun [23], this approach leads to computational savings and simplification of the equations due to its meshless nature, enabling highly detailed simulations of heterogeneous media.

While it is used in its purely absorbing form here, there is no additional difficulty in taking scattering phenomenon into account with this method. The principle behind null collision algorithm is based on a theoretical field of a new majoring extinction coefficient $\widehat{\kappa_\eta}$ defined as follows:

$$\widehat{\kappa_\eta} = \kappa_{abs,\eta} + \kappa_{n,\eta} \tag{9}$$

In practice, the null collision coefficient $\kappa_{n,\eta}$ is completely arbitrary and is never expressed. It is simply defined as the difference between $\widehat{\kappa_\eta}$ and $\kappa_{abs,\eta}$. The principle is that $\widehat{\kappa_\eta}$ is a majorant of the absorption coefficient field. The null collision coefficient then becomes an adjustment variable. However, the integration of this new coefficient should not change the physics of the problem. We can thus consider that the events related to null collisions finally correspond to a forward scattering. Mathematically, adding this type of collision to the differential form of the RTE leads to:

$$\begin{aligned}\mathbf{u}.\nabla I_\eta(\mathbf{x}, \mathbf{u}) = &-\left[\kappa_{abs,\eta}(\mathbf{x}) + \kappa_{n,\eta}(\mathbf{x})\right] I_\eta(\mathbf{x}, \mathbf{u}) + \kappa_{abs,\eta}(\mathbf{x}) I_\eta^b \\ &+ \kappa_{n,\eta}(\mathbf{x}) \int_{4\pi} \delta(\mathbf{u} - \mathbf{u}') I_\eta(\mathbf{x}, \mathbf{u}') d\mathbf{u}'\end{aligned} \tag{10}$$

This formulation is therefore rigorously equivalent to the classical formulation of the RTE for a purely absorbing medium. Based on the Eq. (10), a more relevant equation is deduced. Null collision is a forward scattering event represented by a Dirac phase function. We can thus complete Eq. (3) in the following way:

$$\begin{aligned}I_\eta(\mathbf{x_0}, \mathbf{u_0}) = &\int_0^{+\infty} dl' \widehat{p_X}(\mathbf{x}') \\ &\times \left[\frac{\kappa_{abs,\eta}(\mathbf{x}')}{\widehat{\kappa_\eta}(\mathbf{x}')} I_\eta^b(\mathbf{x}') + \frac{\kappa_{n,\eta}(\mathbf{x}')}{\widehat{\kappa_\eta}(\mathbf{x}')} \int_{4\pi} \delta(\mathbf{u_0} - \mathbf{u}') I_\eta(\mathbf{x}', \mathbf{u}') d\mathbf{u}'\right]\end{aligned} \tag{11}$$

Where $\mathbf{x} = \mathbf{x_0} - l$ and $\widehat{p_X}(\mathbf{x}')$ is analogous to $p_X(\mathbf{x})$ (see Eq. (4)):

$$\widehat{p_X}(\mathbf{x}') = \widehat{\kappa_\eta}(\mathbf{x}) \exp\left(-\int_0^l \widehat{\kappa_\eta}(\mathbf{x_0} - l'\mathbf{u_0}) dl'\right) \tag{12}$$

This last expression can easily be reformulated in a recursive form as proposed by Galtier [24]. The new extinction coefficient previously defined allows to work with an artificial but uniform coefficient field. Then, it becomes possible to sample a free length path according to the probability density function $\widehat{p_X}(\mathbf{x}')$ (see Eq. (12)). Then a new coordinate of the next collision point $\mathbf{x}'$ is computed and its nature is then determined, real absorption

or null collision, thanks to a random number uniformly sampled between 0 and 1:

- If this random number is lower than $\frac{\kappa_{abs,\eta}(\mathbf{x}')}{\widehat{\kappa_\eta}(\mathbf{x}')}$, this means that the event at this point is an absorption. The optical path of the photon is then completed and we can calculate the Monte Carlo weight associated with this point, equal to the radiative intensity of the black body $I^b_\eta(\mathbf{x}')$
- Else, the event is a null-collision which is equivalent to a forward scattering. We will then sample a new free path without changing the direction of propagation of the photon in the same way as before and so on until we reach a collision point where the chosen event is an absorption.

Provided that the coefficient $\widehat{\kappa_\eta}$ majors the whole real field, the heterogeneities of the environment does not raise any difficulty. Moreover, there is no approximation here in the formulation which implies that this method guarantees the preservation of the reference character of the Monte Carlo method. In practice, the calculation of the coefficient $\widehat{\kappa_\eta}$ for each spectral band requires a necessary pre-processing phase. Depending on the thermophysical conditions it is possible to define a priori the values of the absorption coefficient field and set the highest values for each band in order to major the field at any point. Recent studies use this approach in an optimized way, by majoring the field of $\kappa_{abs,\eta}$ by zone thanks to a hierarchical grid, within which the values of the absorption coefficient are close [20]. This method greatly reduces the computational time by limiting the probability of occurrence of the null collision event and is particularly recommended in very heterogeneous media.

3. Gas and soot modeling

3.1. Spectral gas model

The approach presented here is not dependent on the model used for gas properties. In order to be close to the benchmark cases used in the Section 5 [25] without adding bias, the SNB-CK is selected. The SNB model of Malkmus [26], commonly considered as a reference due to its accuracy uses a different formalism and describes the radiative properties of the medium using the transmissivity which cannot been implemented with this resolution method. The SNB-CK approach can be considered as a roundabout way to take advantage of an accuracy close to that of a narrow band model by using an absorption coefficient formulation [27]. In the SNB model, the transmissivity over an isothermal and homogeneous path can be written as follows:

$$\overline{\tau}_\nu(\mathcal{L}) = \exp\left[-\frac{\pi B}{2}\left(\sqrt{1+\frac{4S\mathcal{L}}{\pi B}}-1\right)\right] \tag{13}$$

Where B and S are coefficients depending on the mole fraction, the pressure and other parameters which are described in details by Rivière and Soufiani [28]. The transmissivity of a gas mixture can be approximated by the product of the transmissivity of each radiative species (H_2O and CO_2 mainly, for combustion applications). In the SNB-CK version, we take the inverse Laplace transformation of the SNB gas transmissivity $f(k)$ on each of the 367 bands considered and its cumulative $g(k)$ [29,30]. We uniformly discretized the absorption spectrum of the two radiating species into 367 narrow bands of $25\,cm^{-1}$ thickness, over a range from $150\,cm^{-1}$ to $9300\,cm^{-1}$. Beyond these bands, the radiation of the gases considered is negligible. In practice, a wavenumber is sampled on the whole spectrum according to the Planck's law. If the wavenumber is sampled outside the SNB-CK limits, the absorption coefficient of gases is considered as zero ($\kappa_{abs,\eta} = 0$), on the other hand, the radiation of soot are taken into account which can be significant for large wavenumbers as highlighted in Section 3.2. For combustion applications, it is very efficient to deal with radiative species separately [19] which means that a separate k-distribution is introduced for each specie supposed independent [31]. We use a 7 points Gauss- Legendre quadrature to determine the absorption coefficient of each gas at the randomly selected spectral band before summing them. The value of each absorption coefficient at each quadrature point can be determined by the inversion of $g(k)$ function thanks to a Newton-Raphson method with a very good efficiency [30]. Here this calculation step is performed at each location of the optical path without pre- calculation of the absorption coefficient field. Any radiative observable that depends only on the absorption coefficient can then be calculated. We use the Eymet et al. [19] procedure to calculate the flux divergence for all the following cases.

3.2. Soot radiation modeling

The radiative behaviour of soot is relatively straightforward as it is assumed that there is no scattering due to these particles. In sooting media, the RTE is unchanged. The soot radiation contributes to the expression of the absorption coefficient as expressed in Eq. (14).

We assume here that soot is unagglomerated and its optical properties follow the Rayleight small particle limit [22].It leads to consider that soot radiation is linearly proportional to the total soot volume fraction as proposed by Solovjov and Webb [32]. Thus, this leads to:

$$\kappa_{abs} = \kappa^{gas}_{abs}(T, P, [H_2O], [CO_2], \eta) + \kappa^{soot}_{abs}(f_v, \eta) \tag{14}$$

$$\kappa^{soot}_{abs,\eta} = 5.5 f_v \eta \tag{15}$$

Where f_v is the soot volume fraction. The value of the constant of this model is empiric and can be discussed. It can possibly be modified in some more specific applications. This constant depends on the type of flames for combustion applications. This constant could take another value depending on the fuel and the oxidizer. We have chosen for these academic cases a commonly used value [33].

4. Artificial neural network and radiative transfer

Artificial Neural Networks (ANN) are part of the tools used in applications related to Artificial Intelligence. Radiative transfer is a great field of application to introduce this type of tools in modeling science and engineering issues. The critical point in the radiative transfer is the computational resources needed to solve the RTE a large number of times to finally compute a single source term embedded in other complex equation systems. In the field of engineering sciences, thermal modeling is in most cases part of a multi-physics problem, applied to a complex geometry. The paradox lies in the fact that a fine and accurate modeling of radiative transfer is often fundamental in the design of systems subjected to high temperatures and therefore needs to be given great attention. The objective of this work is to demonstrate the predictive ability of an ANN to model radiative transfer with the accuracy of a spectral model at a low cost in terms of computing power required.

4.1. Definition: Neurons, activation functions and networks

The main idea of ANN is to copy the behavior, connections and abilities of our brain to solve complex problems such as image recognition, language comprehension, etc. Even if these tasks seem very natural and without any real difficulties, they actually call

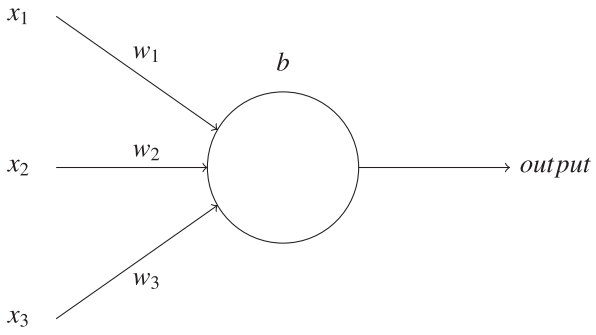


Fig. 1. Simple neuron representation.

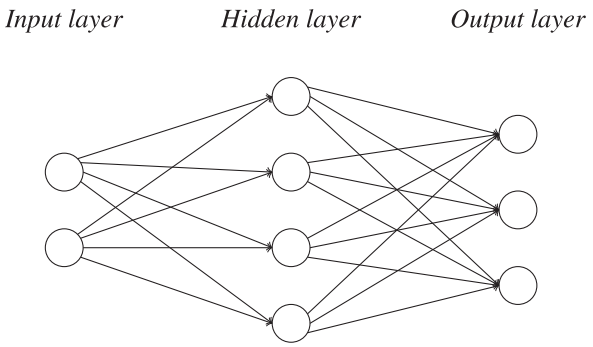


Fig. 2. An example of a simple FANN.

upon laws that are extremely difficult to reconstruct. Tools such as artificial neural networks are not based on deterministic or statistical laws like most models. We can consider them as approximators of complex non-linear functions that try to reproduce a behavior observed from known tests in order to deduce more general laws, applicable to "unseen" tests. The ANN principle was was originally proposed by McCulloch and Pitts [34], shortly after the development of the Turing machine. The elementary component of an ANN is the neuron which has been introduced by Rosenblatt [35] namely a Perceptron. This type of neurons is a binary object with a threshold which defines the value of the output as a function of the input :

$$output = \begin{cases} 0 & if \ \sum_j w_j x_j + b \leq threshold \\ 1 & if \ \sum_j w_j x_j + b > threshold \end{cases} \tag{16}$$

Where x_j is the jth input, weighted by the coefficient w_j (see Fig. 1). Perceptrons are quite obsolete today because of their binary character. The Heaviside function which is called the activation function of the Perceptron, has been replaced by continuous functions. In this work, we used a sigmoöd transfer function:

$$output = \sigma\left(\sum_j w_j x_j + b\right) = \sigma(z) = \frac{2}{1+\exp(-2z)} - 1 \tag{17}$$

Where b is the bias of the neuron which is simply an offset of the sigmoid response. This function has the advantage to be bounded, symmetric and continuous. In practice, others functions with these characteristics can be used for regression tasks. For some specific applications such as image recognition, unbounded function as ReLU ("Rectified Linear Unit") are possible but they will not be discussed here. An ANN consist in an interconnected panel of neurons, which are able to transmit information with each other. There are many different possible architectures depending on the task at hand. We will focus on one of the simplest, namely the single-hidden layer FeedForward Network which will be designated by FANN in all this paper. As shown in Fig. 2 this type of network is composed of:

- An input layer which is basically the parameters of the modeled function that the network take into account,
- Only one "hidden" layer which corresponds to the computational layer. Each neuron of this layer is linked to all the input neurons. Weights of each branch and biases of each neuron are the parameters of the network and change during the training phase to determine the best combination of them and the extrapolation and interpolation abilities of the FANN,
- An output layer which can possibly contain several neurons depending on the problem to be treated. Each output neuron is also connected to all the neurons of the hidden layer and supports modifiable weights and biases during the training phase.

The choice of this basic architecture has been made for several reasons. First, it has been shown that this type of networks possesses the universal approximation property with arbitrary accuracy. It means that theoretically, a single hidden layer FANN is able to approximate any continuous function on its interval of definition under conditions on the activation function in particular and without limitation on the number of neurons [36]. Moreover, adding other hidden layers can lead to a difficult training step [37]. In addition, this choice allows a certain control over the calculations performed compared to more complex topologies as it results in a simplified set of equations. Finally, the bayesian paradigm (see Section 4.3) allows to define easily the optimum topology.

4.2. Backpropagation and Levenberg-Marquardt algorithm

The science of neural networks is an immense field of possibilities and in that way, there is plenty of manners to train and test these tools. Backpropagation algorithm and all its derivatives are probably one of the most common and efficient method today, particularly for FANN as introduced by Rumelhart et al. [38]. By "*learning*" we mean modifying the weights and biases of the network allowing the approximation of the function to be modeled. We use a cost function to determine whether a modification of these parameters is favorable or not to the modeling. We seek to minimize this cost function through learning. In practice, this function depends on a very large number of parameters and the search for an extremum for each of them is utopian. The goal is to find a global set of parameters satisfying an accuracy criterion fixed upstream and depending on the function to be modeled. Backpropagation algorithm is based on the gradient descent as explained by Nielsen [37] and Keller et al. [39, chap. 3.2].

Levenberg Marquardt algorithm is used for the optimization of the network parameters [40]. This variant of the classical backpropagation methods allows a faster convergence but for a higher computational cost. This algorithm is therefore particularly well suited for training using small databases (from a few hundred training sets to a few thousand) mostly due to consequent matrix inversion operations. This method also requires the use of a cost function of the type MSE. The main steps of this algorithm are as follows:

- First, a *forward pass* is necessary to calculate the error vectors and the associated MSE
- Secondly, the Jacobian matrix $J(x_k)$ of the whole training database is built. For that, we compute first the *Marquardt sensivities* on the last layer and they are back-propagated into the entire network.
- We then solve eq. (18) in order to deduce the new value of the network parameters. We notice that by applying the Gauss-Newton method, the Hessian matrix He is calculated by the

product of the Jacobian matrix and its transpose. This theorem allows us not to have to calculate second derivatives. Nevertheless, this matrix may not be invertible, that is why we approximate it by adding a $\mu\mathbf{I}$ diagonal matrix which ensures its invertibility. The term μ is the equivalent of the hyperparameter called learning rate in classical backpropagation algorithms.
- We finally recalculate the value of the error vector after changing the value of the parameters. If the new value of the MSE is lower than the old one, these modifications are kept and μ is divided by 10 then we start a new iteration of the algorithm until convergence. Otherwise, μ is multiplied by 10 and we start the algorithm again at the previous step.

$$\Delta\mathbf{x_k} = -\left[\mathbf{J^T}(\mathbf{x_k})\mathbf{J}(\mathbf{x_k}) + \mu\mathbf{I}\right]^{-1}\mathbf{J^T}(\mathbf{x_k})\mathbf{v}(\mathbf{x_k}) \tag{18}$$

This algorithm is rather straightforward nevertheless, as all backpropagation algorithm, its performance depends a lot on the value of the learning rate μ. Usually, the best value of this hyperparameter is never known in advance. It exists some heuristic principles to know approximately the range in which μ belongs but nothing more. In addition, its value must change over the training for accuracy to continue to increase and once again, the law of motion of this parameter can not be known in advance. With this algorithm and according to Hagan et al. [40], the value of μ and its evolution is given at all iteration. This is one other avantage of Levenberg Marquardt training method.

4.3. Overfitting and Bayesian regularization

One of the major problems with ANN in regression tasks is overfitting. The phenomenon can happen for many reasons and has to be controlled during the training phase. BackPropagation algorithms are particularly prone to this type of problem which occurs when the training is not stopped in time. If a network is trained too much, it loses its ability to generalize: the accuracy of the network will be extremely high compared to training point, but it will not be able to interpolate or extrapolate results from "unseen" data sets. This phenomenon is not always easy to avoid because the cost functions are based on the correspondence between the expected output and the actual output of the network. However, in case of overfitting, this correspondence will always be very good with the data used to train the network. ANNs with many free parameters (weight and bias) are much more prone to this phenomenon than simpler ANNs (Single Layers FANN for example), which is why it is important not to use oversized topologies if it is not necessary. There are some techniques to limit the impact of overfitting :

- Increase the number of test data which also increases the training time,
- Reduce the size of the network,
- Use a regularization technique.

Regularization techniques are almost always used with BackPropagation. They are based on a simple principle: keep network weights low which means that network behavior will change little when inputs change little. In other words, an entry that is very similar to another will result in a similar response, so the ANN will not learn the "noise" that may be contained in the learning data. The techniques of regularization are not sufficient on their own and will not be able to replace an incomplete training database, however they allow most of the time to limit the overfitting and to strongly increase the capacities of generalization of a neural network. In this work, we used the Bayesian regularization. Although it requires more time and resources to train networks, it is very efficient with small or noisy datasets. We note C the cost function error and E_D the mean squared error (MSE). Without any regularization, we have $C = E_D$. Here, we introduce E_W, the sum of squares of the network weights and two hyperparameters α and β.

$$C = \alpha E_D + \beta E_W \tag{19}$$

These parameters represent the competition between the minimization of the MSE and the weights. If $\alpha \ll \beta$, then the training algorithm will promote small weights whereas, if $\beta \ll \alpha$, a small MSE value will be preferred. Actually, as well as for the learning rate, these hyperparameters can not be *a priori* known because they depend on the modeled function : if α is too small, the network will ignore training data and if it is too high it could become overfitted. MacKay [41] proposed to apply the laws of Bayes, based on conditional probabilities, to determine these parameters during the training. It has been shown that BackPropagation algorithm combined with Bayesian regularization is very effective for regression tasks with good abilities to generalize, even when the modeled function fluctuates a lot [13]. The application of the Bayesian paradigm in the case of neural networks implies a certain number of assumptions on the training data which, in the majority of cases, are not rigorously adapted. Actually, to determine the set of motion laws of the hyperparameters of a Bayesian neural network, we consider that the training output data are defined as the sum of an exact value and a Gaussian noise centered in zero. This assertion is usually false but in this work, the theoretical output data are generated by a Monte Carlo algorithm. The computed values are therefore naturally of the form *exact value + standard deviation*. Bayesian neural networks are therefore particularly adapted. Moreover, this method has the great advantage of intrinsically allowing the application of a strong and essential philosophical principle in engineering called Occam's Razor which can be understood as an injunction to parsimony [42]. The main idea behind this principle is to always make sure to exploit as much as possible the hypotheses already used before posing new ones. Thus, it is not the simplicity of *the assumptions* that counts but rather the low complexity of *the set of assumptions* made to solve a problem. In practice here, Bayes' laws allow us to determine, thanks to the Hessian matrix calculated during the application of the Levenberg-Marquardt algorithm, an effective number of parameters γ used by the neural network [40]. Thanks to this parameter, it is possible to adjust the topology of the neural network during the computation in order not to oversize it and hence, to optimize the training time and strongly limit overfitting. This characteristic of Bayesian networks also allows us to eliminate the time-consuming phase of determining the topology. The algorithm as a whole is then automated to adapt to new cases of study without difficulty or redundant task. In practice, all these hyperparameters can be computed iteratively during the execution of the algorithm according to the following relations:

$$\alpha = \frac{\gamma}{2E_W} \tag{20}$$

$$\beta = \frac{N_{out} - \gamma}{2E_D} \tag{21}$$

$$\gamma = n - 2\alpha \mathrm{Tr} He^{-1} \tag{22}$$

Where n is the total number of parameters, N is equal to the number of training output data and He, the hessian matrix of the training set in the Levenberg-Marquardt algorithm [40]. At the first iteration, γ is initialized to n then α and β can be deduced easily. In practice and to ensure that the network is large enough we slightly increase the number of effective parameters by 20% compare to the calculated value of γ [13]. If the number of free parameters is too close to the γ value, the training may not be optimal

because γ may increase beyond its final value before reaching a converged solution.

4.4. Train an ANN with MC reference probe points

As mentioned in previous sections, the major aim of this method is to combine advantages of Monte Carlo algorithms and ANNs. Statistical methods allow a control of uncertainty at precise points. However, these methods are not well suited to compute an entire field of variables. With ANNs, it becomes possible to generalize Monte Carlo results efficiently and accurately in terms of computational resources required and to switch from probe calculation to field calculation. This needs to happen under the condition of a good and complete ANN training. Thanks to the cost function, we have a quality criterion characterising the confidence we can place in the ANN results. The main idea is to sample a number of reference points in the domain and to calculate the value of the radiative quantity of interest, for instance, the flux divergence thanks to a Monte Carlo algorithm solving the radiative transfer equation in its integrated form (see Eqs. (2) to (7)). These points will constitute the training base of a neural network thanks to which prediction of this quantity in the rest of the domain. The construction of the databases and the neural networks used depend on the complexity of the study cases:

- For the homogeneous cases, the training points are sampled randomly in the domain according to a uniform distribution. The input data of the neural networks are the spatial coordinates of the training points. The output data is the divergence of the radiative flux.
- For the heterogeneous case, the training points are sampled according to a low-discrepancy Halton sequence in two dimensions [43]. It leads to a significant gain on the performance of the neural network thanks to this type of sampling. It allows to limit the appearance of white areas in the training points of the network and to improve the quality of predictions for a zero additional computation cost. The neural network used in this case is also different. In a highly heterogeneous case, a simple spatial interpolation based on the coordinates of the training points is not satisfactory. Indeed, similar performances could be obtained with more classical and faster interpolation methods (see Section 7.3). We then added three additional in-

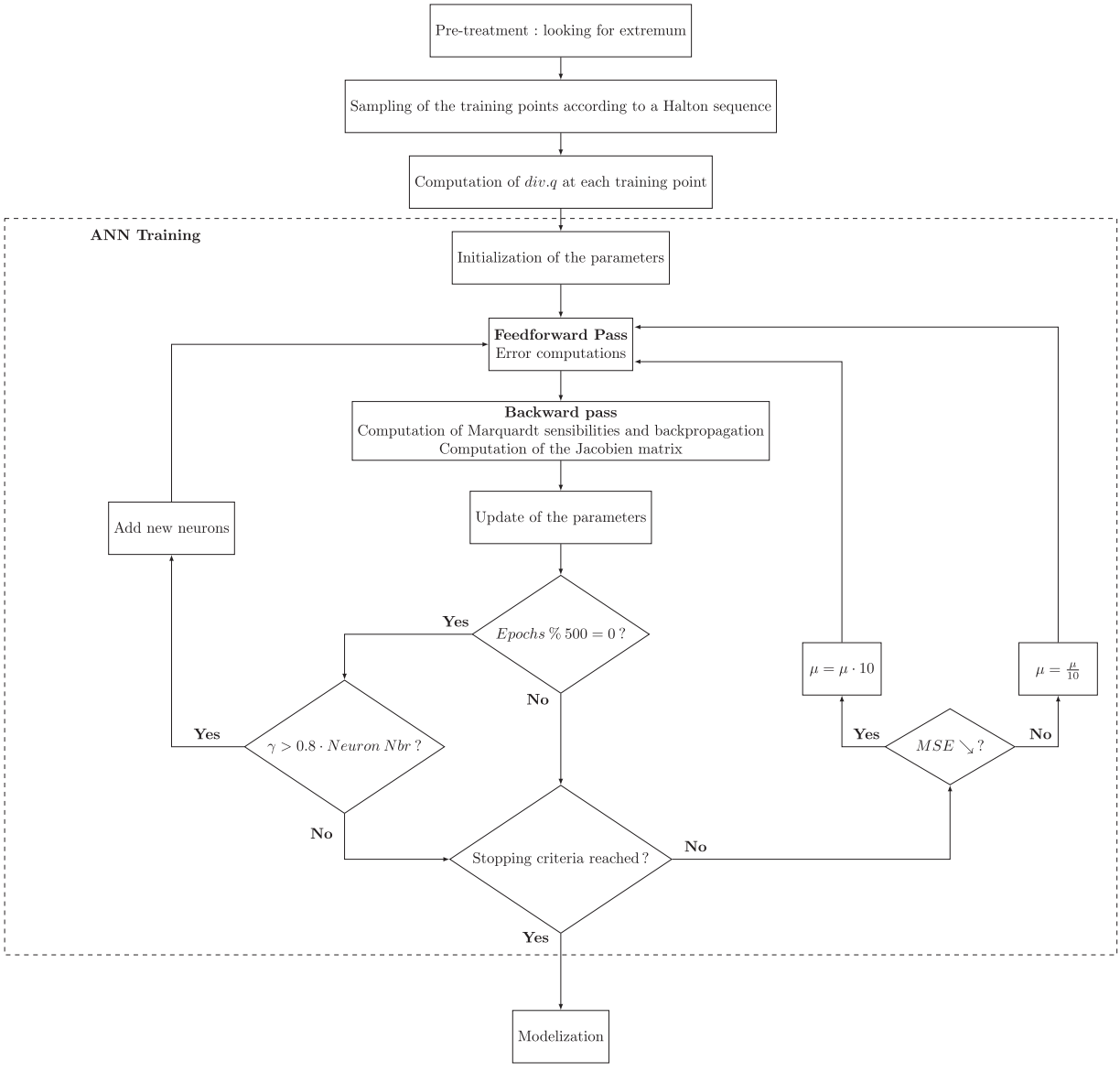


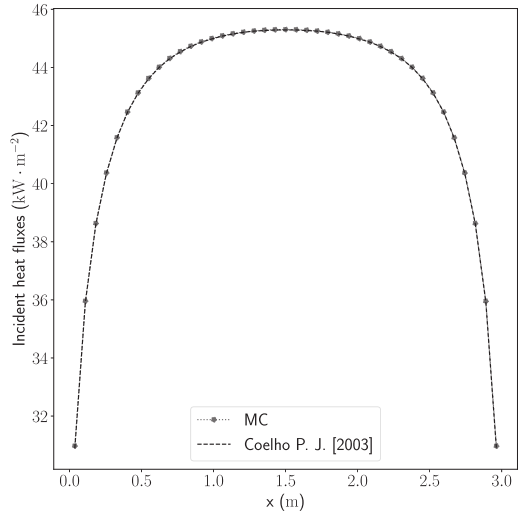
Fig. 3. Description of the algorithm.

puts containing thermophysical data at the training point: temperature, H_2O and CO_2 concentrations. The addition of this data allows the network to implicitly model the emissive contribution at the point of interest and, in particular, allows the divergence of the flux at the points with the highest temperatures to be modeled with much better accuracy in interpolation and even in extrapolation.

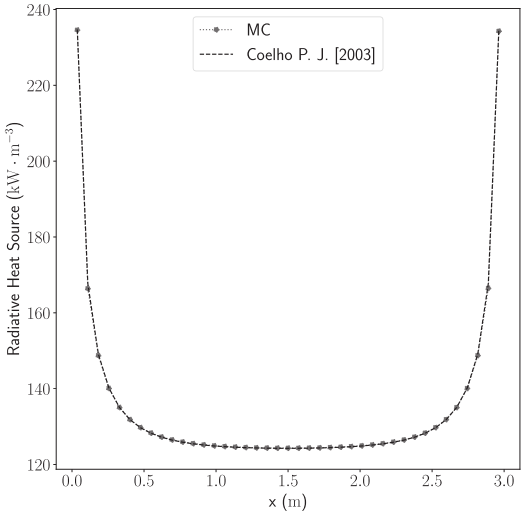
Once neural networks are trained, the predictions of the AI are compared against the Monte Carlo calculation on "unknown" points and the average and maximum relative or absolute errors are calculated, depending on the case. It is quite unrealistic to imagine a very large network that will be able to predict the radiative heat transfer in all situations as well as there is no advantage in using ANNs if learning requires many expensive numerical simulations. Thus the method proposed here consists in training small ANNs during the simulation with a reduced but very targeted database. For this method to be feasible, it is necessary to keep the number of neurons as small as possible in order to train the networks in a short time. As mentioned earlier, the network topology is fixed according to the value of the number of effective parameters γ for each case of study and networks used include only one hidden layer. The entire algorithm is describe in details in Fig. 3.

5. Benchmark academic cases

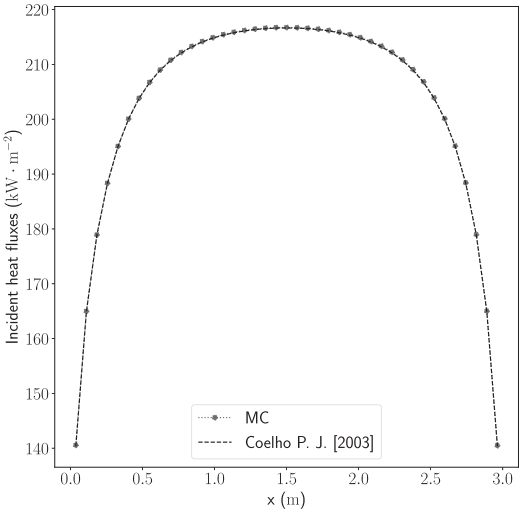
To highlight the performance of the code presented here for radiative heat transfer modeling, academic benchmark cases are studied. Attention is focused on the divergence of the radiative flux, noted $\div q$ and expressed in $\mathrm{kW \cdot m^{-3}}$ within the domain. The geometry used in the three cases is an axisymmetric enclosure. For the first two cases of study, we consider a cylinder of 3m length and 0.5m radius with black walls. The temperature of the walls is 300 K and a homogeneous mixture of radiative species is considered. The volumetric composition is 20% H_2O, 10% CO_2 and 70% N_2. The soot volumetric fraction is set to 10^{-7} and the temperature of the medium is equal to 1200 K in the first case and 1800 K in the



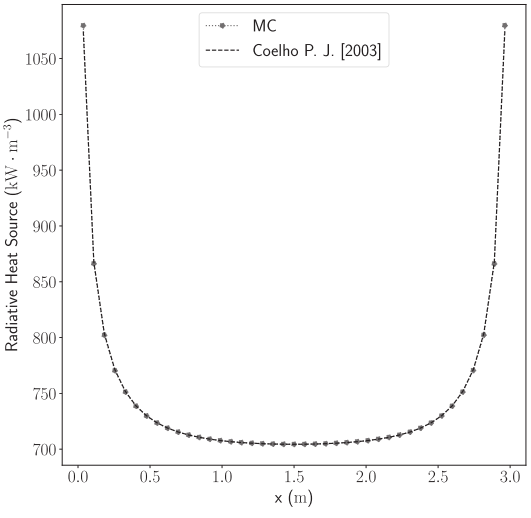
(a) Wall flux at the lateral wall for T = 1200 K



(b) Flux divergence at the centerline for T = 1200 K



(c) Wall flux at the lateral wall for T = 1800 K



(d) Flux divergence at the centerline for T = 1800 K

Fig. 4. Comparison of our results with referenced data [25].

second one. In the third situation, we consider a cylinder of 1.2m length and 0.3m radius with black walls. The wall temperature is set at 800 K except the right wall, maintained at 300 K. In this more complex case, we deal with an heterogeneous distribution in temperature and in molar fraction of species following :

$$T(x,r) = 800 + 1200\left(1-\frac{r}{R}\right)\left(\frac{x}{L}\right) \tag{23}$$

$$x_{H_2O}(x,r) = 0.05\left[1-2\left(\frac{x}{L}-0.5\right)\right]\left(2-\frac{r}{R}\right) \tag{24}$$

$$x_{CO_2}(x,r) = 0.04\left[1-3\left(\frac{x}{L}-0.5\right)\right]\left(2.5-\frac{r}{R}\right) \tag{25}$$

These cases were the subject of a benchmark between several reference methods in radiative transfer, namely Ray Tracing Method, Monte Carlo Method-Net Exchange Formulation and Discrete Ordinates Method [25]. These results are used to compare and validate our simulations. A modified version of HTRDR code [44], developed by Meso-Star and based on Star-Engine open-source library is used to compute the radiative quantities of interest.

6. Validation of Monte Carlo algorithm

In this section, the Monte Carlo algorithm used to build the training database for the ANN is validated. Each simulation is made for a standard deviation of less than 1% of the quantity being evaluated. The first case addresses a homogeneous environment. As presented on Fig. 4 comparison between the study of Coelho et al. [25] and the data calculated by our code shows a good agreement for both the wall flux and the divergence of the flux.

For the heterogeneous case, Fig. 5 also shows a very good agreement between the results. Each point takes about 4 min and 45 s in the first case and around 30 s in the second one to calculate with four cores of a standard desktop computer equipped with an Intel®Core™ i5-8265U processor and 15.5 GB of RAM.

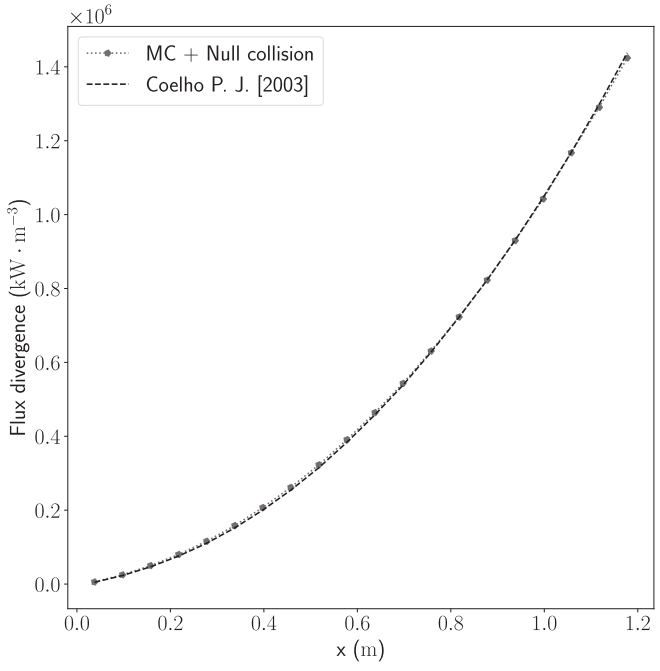


Fig. 5. Flux divergence at the centerline of the heterogeneous case.

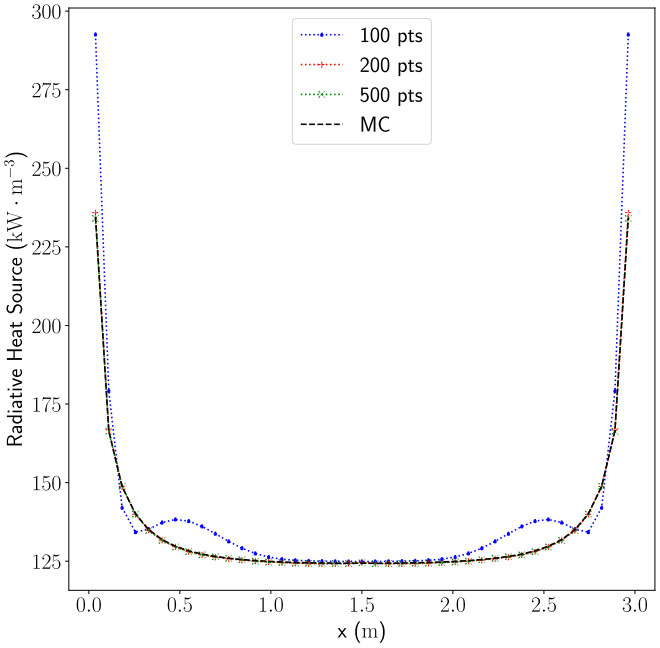


Fig. 6. Predictions of a trained ANN using three databases of different sizes (100, 200 and 500 points) for T = 1800 K.

7. Predictive abilities of ANNs for radiative heat transfer modeling: Results and discussions

7.1. Homogeneous cases

For homogeneous cases, training points are randomly sampled in the domain according to a uniform distribution. Inputs of the ANN correspond to the spacial coordinates of the probe points. Tables 1 and 2 compile the average and maximum relative errors for flux divergence at these points obtained with Monte Carlo method and predicted by ANNs. Three different databases sizes (100, 200 and 500 points) are tested. All results presented in this

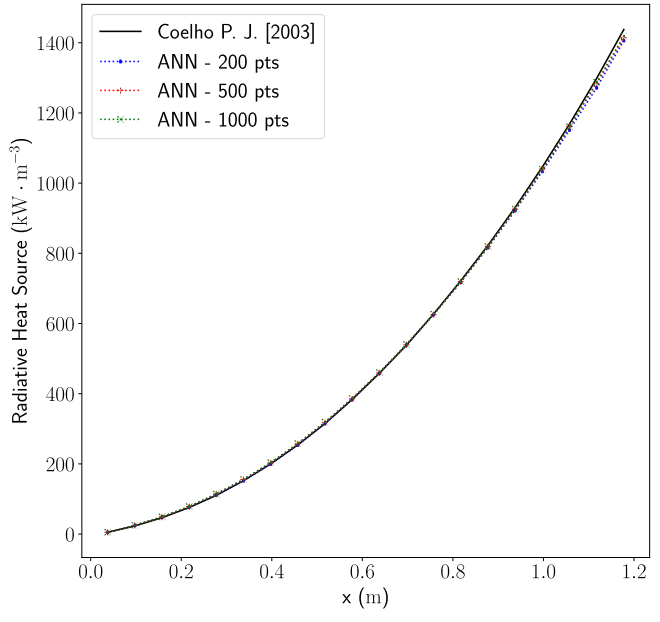


Fig. 7. Flux divergence at the centerline of the heterogeneous case - Comparison between Coelho's data [2003] and ANN predictions for different databases.

Table 1
ANN's predictions when T = 1200 K.

Training points	Neurons	Parameters	γ	$\varepsilon_{mean}\%$	$\varepsilon_{max}\%$
100	11	45	36	0.86%	39.74%
200	11	45	37	1.53%	23.51%
500	17	69	55	0.23%	1.69%

Table 2
ANN's predictions when T = 1800 K.

Training points	Neurons	Parameters	γ	$\varepsilon_{mean}\%$	$\varepsilon_{max}\%$
100	13	53	44	0.29%	5.25%
200	14	57	45	0.32%	7.05%
500	17	69	57	0.07%	1.54%

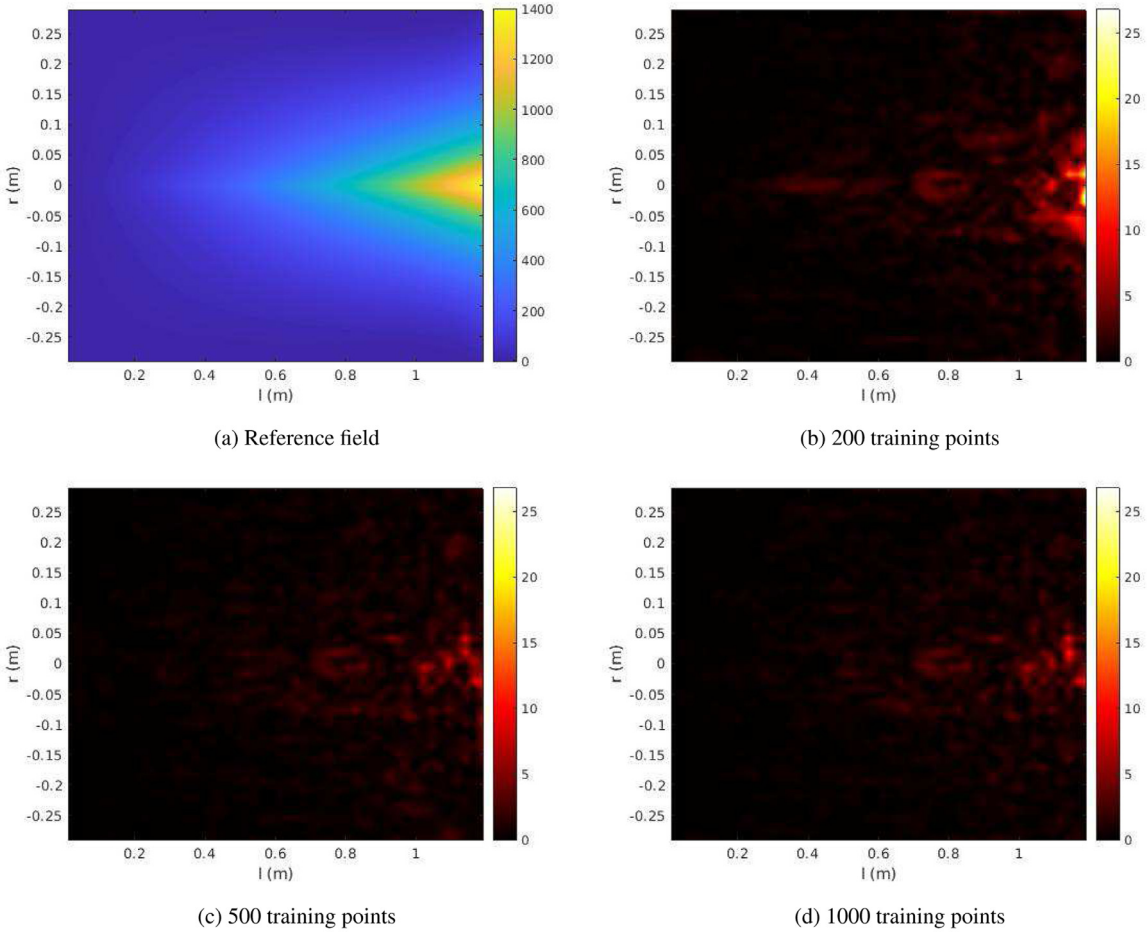
paper ensure a relative deviation on the training data of less than 1%. As mentioned before, the topology of the neural network and more specifically, the number of hidden neurons in the hidden layer has been determined according to the value of γ computed during the training phase. We can notice that the number of neurons needed increases with the number of training points. This means that the computational time will rapidly increase with the number of training points and concomitantly of neurons. At least 500 probe points are needed to reach a certain confidence in ANN results everywhere according to Tables 1 and 2. Actually, even if the mean relative error for 100 or 200 probe points is close to 1%, there are some points which are far from the reference values with a maximal relative error close to 40% in certain cases. However, we noticed that these high values of relative error were reached near the walls where the value of the flux divergence is very low. Therefore, a very small absolute error can lead to a large relative error without the result being harmful from a more macroscopic point of view for the simulation as a whole.

Actually, if the training database is too small, the ANN performances are highly dependent of the initialization of the weights and biases of the networks. According to Tables 1 and 2, we recommend to use 500 probe points here to ensure a good modeling of the flux divergence, including close to the walls or this type of case and for this geometric dimensions.

Figure 6 shows an ANN trained with three very converged databases. Then predictions of the flux divergence on the centerline for different database sizes are plotted. We can notice that the training with 100 probe points is clearly not sufficient. Results are better with 200 points but we can notice a small gap on the points very close to the wall. This difference does not exist anymore with a database of 500 points. The curve overlaps almost perfectly with the reference. It should be noted that for this type of network with a database of 500 points, the training phase is almost instantaneous. The time required is of the order of a second without parallelization and with the same desktop as mentioned before.

7.2. Heterogeneous case

The last case treated here is heterogeneous in species and temperature [25]. This case presents the interest to test the neural net-



(a) Reference field

(b) 200 training points

(c) 500 training points

(d) 1000 training points

Fig. 8. Reference fields of radiative source term ($kW \cdot m^{-3}$) calculated by Monte Carlo method and subtraction between the field predicted by the ANN's - Comparison between the different databases.

Table 3
ANN's predictions in the heterogeneous case.

Training points	Neurons	Parameters	γ	Mean AE kW · m^{-3}	Max AEkW · m^{-3}
200	15	106	86	0.784kW · m^{-3}	26.86 kW · m^{-3}
500	15	106	87	0.577kW · m^{-3}	9.433 kW · m^{-3}
1000	16	113	94	0.544kW · m^{-3}	9.432kW · m^{-3}

works on complex distributions of thermophysical conditions. In this more complexe case, the sampling of training points is carried out through a low-discrepancy sequence as mentionned by Mishra and Molinaro [10]. We used the Halton method in base 2 and 3 to determine the coordinate of each training points [43]. For this case an ANN with 5 inputs is implemented, i.e. the spatial coordinates, the concentrations of CO_2 and H_2O and the temperature at the point considered. The soot volume fraction and the pressure being constant, these parameter is not used as an input. In this situation, the indicators used in the previous study are no longer appropriate. Indeed, the values of the divergence of the radiative flux can be particularly low here. Thus, small absolute error on this quantity can lead to very large relative errors without any impact on the other quantities of interest in a problem of this type (calculation of the wall temperature, the wall flux, etc.). We preferred to use indicators like the mean absolute error and the maximum absolute error. On this case, databases containing 200, 500 and 1000 points are used.

In the Table 3, the mean and maximum absolute error are particularly low for each database and in particular for 500 and 1000 points. The average value of the flux divergence in the whole domain is equal to 155.34kW · m^{-3}. Thus the mean absolute errors represent respectively 0.51%, 0.37% and 0.35% compared to this value for 200, 500 and 1000 point databases.

The final objective of our tool coupling Monte Carlo method and neural networks is to be able to compute entire fields of a radiative quantity independently of the number of cells. We have tested this functionality by discretizing the study domain very finely and by computing the radiative source term in each cell. The study area is 1.2m long and 0.6m high (from -0.3m to 0.3m). We discretized the whole area according to a uniform mesh of 2 500 cells (50 in the x direction and 50 in the y direction). Figure 8 shows a reference field of the radiative source term computed thanks to Monte Carlo method in its null collision formulation. 3 field subtractions between the reference and the predictions of the neural networks trained on the different databases are presented. The CPU time for each picture of Fig. 8 depends of the number of training points. The Monte Carlo algorithm used here is highly parallelizable. The calculation time could thus be strongly decreased on more consequent machines. The ANN training time is never higher than 30 s on a standard laptop computer. Then, the regression in each cell is almost instantaneous.

Figure 7 and 8 show that a database of 200 points guarantees a good interpolation of the value of interest globally. However, we note a larger underestimation of the flux divergence in the hottest areas on the central line between 1m and 1.2m. Table 3 highlights a significantly larger maximum absolute error than for the two databases. Very concretely, we can thus say that if a database of 200 points can be sufficient for a quick estimation of the flux divergence, it is nevertheless recommended to use at least 500 reference points for applications requiring more precision on this type of study case.

7.3. Comparison with other interpolation methods

The proposed method is based on the powerful interpolation and generalization capabilities of artificial neural networks. Beyond the classical terminology of *Artificial Intelligence*, these tools are more precisely advanced statistical tools which, thanks to supervised learning, model non-explicit laws drawn from available data. In this section, we compare the interpolation capabilities of neural networks on the complex heterogeneous case against more classical spatial interpolation techniques: cubic spline interpolation (CS), natural neighbour radial point interpolation method (NNRP) and bi-linear interpolation (BL) [45,46]. Comparisons between these different spatial interpolation methods are numerous in the literature, particularly in cartography, geology and imaging [47–49]. However, depending on the case, the efficiency of neural networks for simple spatial interpolation is questionable compared to more direct methods. Their use becomes especially relevant in the presence of noisy or error-prone data. The training algorithms of neural networks allow a tolerance on the correspondence between the theoretical training data and the predicted data when the other interpolation methods are based exactly on the values between which the interpolation is done. In our case, the neural networks are not only supplied with spatial coordinates: thermophysical data at the point of interest are also provided. In this way, the network interpolates the divergence of the flux based on the known surrounding points but also, implicitly, on the radiation intensity of the black body emitted at the point of interest depending on these data. The neural network then becomes a much more robust method because it is able to handle more unpredictable variations in the physical properties of the field of study even without immediate proximity to a reference point. In this part, a database made of 500 highly converged points following a Halton sequence is used. As in the previous section, this database will be used as a training tool for a 5 inputs neural network but also as a basis for the different interpolation methods.

Figure 9 presents the value of the radiative flux divergence on the centerline in the previously described heterogeneous case. We

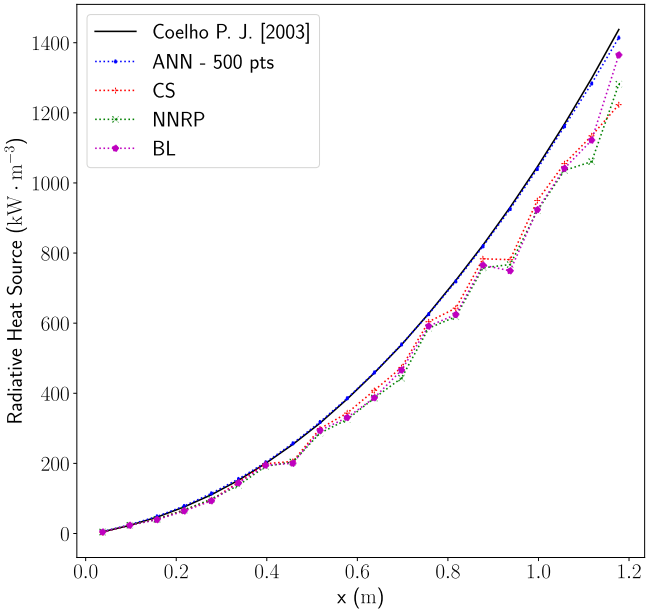
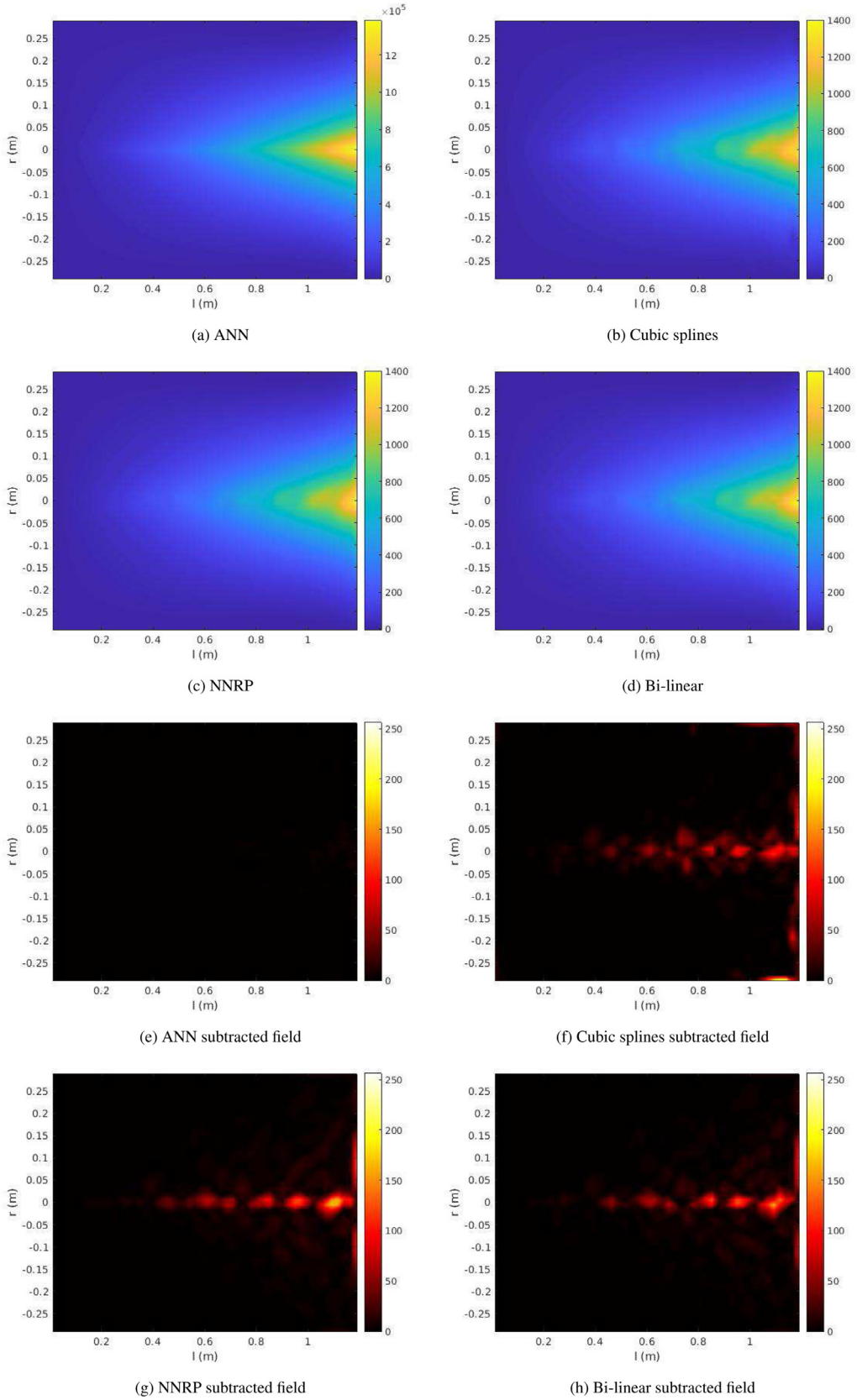


Fig. 9. Flux divergence at the centerline of the heterogeneous case - Comparison between ANN prediction and interpolation methods.



(a) ANN

(b) Cubic splines

(c) NNRP

(d) Bi-linear

(e) ANN subtracted field

(f) Cubic splines subtracted field

(g) NNRP subtracted field

(h) Bi-linear subtracted field

Fig. 10. Reference fields of radiative source term ($kW \cdot m^{-3}$) calculated by Monte Carlo method and subtraction between the field predicted by the ANN's - Comparison between the different databases.

Table 4
Comparison of RMSE value between different interpolation methods.

Method	RMSE
ANN	$1.007 kW \cdot m^{-3}$
Cubic Splines	$15.79 kW \cdot m^{-3}$
NNRP	$15.21 kW \cdot m^{-3}$
Bi-linear	$14.05\ kW \cdot m^{-3}$

notice a very clear difference between the standard interpolation methods and the multivariate neural network interpolation.

The interpolation of a full field of radiative source term is also tested. Figure 10 shows that the trend observed on the central line can be generalized to the whole field. Finally, Table 4 displays an order of magnitude difference between the root mean square error values computed from the neural network results and the other interpolation methods. The Bayesian paradigm also allows to keep the computation time for training the neural network low and to maintain an important competitive advantage compared to other interpolation methods.

8. Conclusion

In this work, several neural networks are trained to predict an entire field of the radiative heat flux divergence, in homogeneous and heterogeneous conditions. The results show the ability of ANNs to predict very efficiently and accurately a multivariant quantity such as radiative heat flux divergence. Concerning the topology of the network, we show that the five input neural network requires far fewer neurons in the hidden layer and thus significantly improves interpolation capabilities in heterogeneous cases. The Bayesian framework allows us to eliminate the phase of choosing the topology and the potentially arbitrary choices that can result from it. Moreover, this regularization technique allows to strongly limit the overfitting and to automatically optimize the arbitration between the different hyperparameters according to the available training data. Finally, we show that this type of training algorithm is adapted to small databases and exploits the maximum information.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Alex Royer: Conceptualization, Data curation, Formal analysis, Writing – original draft. **Olivier Farges:** Conceptualization, Data curation, Formal analysis, Writing – original draft. **Pascal Boulet:** Conceptualization, Data curation, Formal analysis, Writing – original draft. **Daria Burot:** Conceptualization, Data curation, Formal analysis, Writing – original draft.

Data availability

Data will be made available on request.

Acknowledgment

The authors acknowledge financial support from Safran Aircraft Engines.

References

[1] J. Thibault, B.P.A. Granjean, A neural network methodology for heat transfer data analysis, Int. J. Heat Mass Transf. 34 (8) (1991) 2063–2070.
[2] A. Pacheco-Vega, M. Sen, K.T. Yang, R.L. McClain, Neural network analysis of fin-tube refrigerating heat exchanger with limited experimental data, Int. J. Heat Mass Transf. 44 (2001) 763–770.
[3] H. Bechtler, M.W. Browne, P.K. Bansal, V. Kecman, Neural networks - a new approach to model vapour-compression heat pumps, Int. J. Energy Res. 25 (2001) 591–599.
[4] A. Lazrak, F. Boudehenn, S. Bonnot, G. Fraisse, A. Leconte, P. Papillon, B. Souyri, Development of a dynamic artificial neural network model of an absorption chiller and its experimental validation, Renew. Energy 86 (2016) 1009–1022.
[5] S. Azizi, E. Ahmadloo, Prediction of heat transfer coefficient during condensation of R134a in inclined tubes using artificial neural network, Appl. Therm. Eng. 106 (2016) 203–210.
[6] W. Chang, X. Chu, A.F.B.S. Fareed, S. Pandey, J. Luo, B. Weigand, E. Laurien, Heat transfer prediction of supercritical water with artificial neural networks, Appl. Therm. Eng. 131 (2018) 815–824.
[7] K. Ye, Y. Zhang, L. Yang, Y. Zhao, N. Li, C. Xie, Modeling convective heat transfer of supercritical carbon dioxide using an artificial neural network, Appl. Therm. Eng. 150 (2019) 686–695.
[8] M. Yarahmadi, J.R. Mahan, K. McFall, Artificial neural networks in radiation heat transfer analysis, J. Heat Transf. 142 (2020).
[9] Y. Sun, J. Ma, Y. Yu, B. Ye, C. Gao, Efficient SLW models for water vapor and carbon dioxide based on neural network method, J. Quant. Spectrosc. Radiat. Transf. 236 (2019).
[10] S. Mishra, S. Molinaro, Physics informed neural networks for simulating radiative transfer, J. Quant. Spectrosc. Radiat. Transf. 270 (2021).
[11] Y. Zhou, C. Wang, T. Ren, C. Zhao, A machine learning based full-spectrum correlated k-distribution model for nonhomogeneous gas-soot mixtures, J. Quant. Spectrosc. Radiat. Transf. 268 (2021).
[12] F. André, C. Cornet, C. Delage, P. Dubuisson, M. Galtier, On the use of recurrent neural networks for fast and accurate non-uniform gas radiation modeling, J. Quant. Spectroc. Radiat.Transf. 293 (2022).
[13] D. Foresee, M.T. Hagan, Gauss-newton approximation to Bayesian learning, in: Proceeding of International Conference on Neural Networks, Houston, Texas, USA, IEEE, 1997, pp. 1930–1935.
[14] J.R. Howell, M. Perlmutter, Radiant transfer through a gray gas between concentric cylinders using monte carlo method, Trans. ASME Series C 87 (1965) 388–396.
[15] J.R. Howell, Application of Monte Carlo to heat transfer problems, in: Advances in Heat Transfer, vol. 5, Elsevier, 1969, pp. 1–54.
[16] J.R. Howell, The Monte Carlo method in radiative heat transfer(1998).
[17] M.F. Modest, Narrow-band and full-spectrum k-distributions for radiative heat transfer-correlated-k vs. scaling approximation, J. Quant. Spectrosc. Radiat. Transf. 76 (1) (2003) 69–83.
[18] M. Galtier, S. Blanco, C. Caliot, C. Coustet, J. Dauchet, M. El-Hafi, V. Eymet, R.A. Fournier, J. Gautrais, A. Khuong, B. Piaud, G. Terrée, Integral formulation of null-collision Monte Carlo algorithms, J. Quant. Spectrosc. Radiat. Transf. 125 (2013) 57–68, doi:10.1016/j.jqsrt.2013.04.001. https://hal.archives-ouvertes.fr/hal-01688110
[19] V. Eymet, D. Poitou, M. Galtier, M. El Hafi, G. Terrée, R. Fournier, Null-collision meshless monte-carlo - application to the validation of fast radiative transfer solvers embedded in combustion simulators, J. Quant. Spectrosc. Radiat. Transf. 129 (2013) 145–157.
[20] M. El Hafi, S. Blanco, J. Dauchet, R. Fournier, M. Galtier, L. Ibarrart, J.M. Tregan, N. Villefranque, Three viewpoints on null-collision monte carlo algorithms, J. Quant. Spectrosc. Radiat. Transf. 260 (2021).
[21] L. Ibarrart, Description en espace de chemins et méthode de Monte Carlo pour les transferts thermiques couplés dans les structures fluides et solides, une approche compatible avec l'informatique graphique, Ecole Nationale Supérieure des Mines d'Albi-Carmaux, Université de Toulouse, 2020 Phd thesis.
[22] M.F. Modest, Radiative heat transfer, 2th edition, 2003.
[23] J.R. Howell, K.J. Daun, The past and future of the Monte Carlo method in thermal radiation transfer, J. Heat Transf. 143 (10) (2021).
[24] M. Galtier, Approche statistique du rayonnement dans les milieux gazeux hétérogènes : de l'échantillonage des transitions moléculaires au calcul de grandeurs radiatives, Ecole des Mines d'Albi-Carnaux, Université de Toulouse, 2015 Phd thesis.
[25] P.J. Coelho, P. Perez, M. El Hafi, Benchmark numerical solutions for radiative heat transfer in two-dimensional axisymmetric enclosures with nongray sooting media, Numer. Heat Transf. Part B Fundamentals 43 (5) (2003) 425–444.
[26] W. Malkmus, Random lorentz band model with exponential-tailed S^{-1} line intensity distribution function, J. Opt. Soc. Am. 57 (3) (1967) 323–329.
[27] P. Rivière, A. Soufiani, J. Taine, Correlated-k and fictitious gas methods for H20 near 2.7 µm, J. Quant. Spectrosc. Radiat. Transf. 48 (2) (1992) 187–203.
[28] P. Rivière, A. Soufiani, Updated band model parameters for H2O, CO2, CH4 and CO radiation at high temperature, Int. J. Heat Mass Transf. 55 (13–14) (2012) 3349–3358.
[29] G.A. Domoto, Frequency integration for radiative transfer problems involving homogeneous non-gray gases: the inverse transmission function, J. Quant. Spectrosc. Radiat. Transf. 14 (9) (1974) 935–942.

[30] F. Liu, G.J. Smallwood, L. Gülder O., Application of the statistical narrow-band correlated-k meethod to low-resolution spectral intensity and radiative heat transfer calculations - effects of the quadrature scheme, Int. J. Heat Mass Transf. (43) (2000) 3119–3135.
[31] J. Taine, A. Soufiani, Gas IR radiative properties: from spectroscopic data to approximate models, Adv. Heat Transf. 33 (1999) 295–414.
[32] V.P. Solovjov, B.W. Webb, An efficient method for modeling radiative transfer in multicomponent gas mixture with soot, J. Heat Transf. 123 (3) (2001) 450–457.
[33] C.A. Hoerlle, F.H.R. França, P.R. Pagot, F.M. Pereira, Effects of radiation modeling on non-premixed sooting flames simulations under oxyfuel conditions, Combust. Flame 217 (2020) 294–305.
[34] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys. 5 (1943) 115–133.
[35] F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain, Psychol. Rev. 65 (6) (1958) 386–408.
[36] M. Leshno, V.Y. Lin, A. Pinkus, S. Schocken, Multilayer feedforward networks with a nonpolynomial activation function can approximate any function, Neural Netw. 6 (1993) 861–867.
[37] M.A. Nielsen, Neural networks and deep learning, 2015.
[38] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323 (9) (1986) 533–536.
[39] J.M. Keller, D. Liu, D.B. Fogel, Fundamentals of Computational Intelligence, 1 ed., 2016.
[40] M.T. Hagan, H.B. Demuth, M.H. Beale, O. De Jesus, Neural network design, 2th ed., 1996.
[41] D.J.C. MacKay, Bayesian interpolation, Neural Comput. 4 (3) (1992) 415–447.
[42] W.M. Thorburn, The myth of occam's razor, Mind 27 (3) (1918) 345–353.
[43] L. Kocis, W.J. Whiten, Computational investigations of low-discrepancy sequences, ACM Trans. Math. Softw. 23 (2) (1997) 266–294.
[44] N. Villefranque, R. Fournier, F. Couvreux, S. Blanco, C. Cornet, V. Eymet, V. Forest, J.-M. Tregan, A path-tracing monte carlo library for 3-d radiative transfer in highly resolved cloudy atmospheres, J. Adv. Model. Earth Syst. 11 (8) (2019) 2449–2473, doi:10.1029/2018MS001602.
[45] N.S.N. Lam, Spatial interpolation methods: a review, Am. Cartogr. 10 (2) (1983) 129–150.
[46] L.M.J.S. Dinis, R.M. Natal Jorge, J. Belinha, Analysis of 3D solids using the natural neighbour radial point interpolation method, Comput. Methods Appl. Mech. Eng. 196 (13–16) (2006) 2009–2028.
[47] J. Go, K. Sohn, C. Lee, Interpolation using neural networks for digital still cameras, IEEE Trans. Consum. Electron. 46 (3) (2000) 610–616.
[48] N. Plaziac, Image interpolation using neural networks, IEEE Trans. Image Process. 8 (11) (1999) 1647–1651.
[49] J.P. Rigol, C.H. Jarvis, N. Stuart, Artificial neural networks as a tool for spatial interpolation, Int. J. Geogr. Inf. Sci. 15 (4) (2001) 323–343.