



**HAL**  
open science

# Amélioration des méthodes de calcul thermique par réseau de neurones dans les chambres de combustion aéronautiques

Alex Royer

► **To cite this version:**

Alex Royer. Amélioration des méthodes de calcul thermique par réseau de neurones dans les chambres de combustion aéronautiques. Thermique [physics.class-ph]. Université de Lorraine, 2023. Français. NNT : 2023LORR0141 . tel-04292117

**HAL Id: tel-04292117**

**<https://hal.univ-lorraine.fr/tel-04292117v1>**

Submitted on 17 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**UNIVERSITÉ  
DE LORRAINE**

**BIBLIOTHÈQUES  
UNIVERSITAIRES**

## AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact bibliothèque : [ddoc-theses-contact@univ-lorraine.fr](mailto:ddoc-theses-contact@univ-lorraine.fr)  
*(Cette adresse ne permet pas de contacter les auteurs)*

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

[http://www.cfcopies.com/V2/leg/leg\\_droi.php](http://www.cfcopies.com/V2/leg/leg_droi.php)

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



# Thèse

présentée pour l'obtention du titre de

**Docteur de l'Université de Lorraine**

Mention Énergie et mécanique

par **Alex Royer**

## **Amélioration des méthodes de calcul thermique par réseau de neurones dans les chambres de combustion aéronautiques**

Soutenue le 23 Juin 2023 devant la commission d'examen composée de

<i>Président :</i>	M. Rodolphe VAILLON	Directeur de recherche - CNRS
<i>Rapporteurs :</i>	Mme Mouna EL HAFI	Professeure - IMT Mines Albi
	M. Rodolphe VAILLON	Directeur de recherche - CNRS
<i>Examineurs :</i>	Mme Bénédicte CUÉNOT	Docteure - Cheffe de projet - Cerfacs
	Mme Daria BUROT	Docteure Ingénieure de recherche - Safran Aircraft Engines
<i>Directeur :</i>	M. Pascal BOULET	Professeur - Université de Lorraine
<i>Co-directeur :</i>	M. Olivier FARGES	Maitre de Conférences - Université de Lorraine



# Résumé

Les chambres de combustion aéronautiques sont des environnements contraints fonctionnant à des températures et des pressions en constante augmentation depuis les dernières décennies. Dans ces conditions, les transferts thermiques par rayonnement sont conséquents et leurs influences sur la structure de flamme et sur les températures de parois sont sensibles. Il est donc essentiel d'être capable de modéliser avec précision ce mode de transfert sur des cas industriels très hétérogènes et sur des maillages de plusieurs centaines de millions de cellules. Les méthodes de résolution courantes ne permettent pas d'atteindre une précision satisfaisante compte tenu des besoins croissants. Des méthodes de résolution fines et l'utilisation de modèles de gaz spectraux seraient nécessaires mais impliqueraient une consommation de ressources de calcul prohibitives.

Pour palier ce problème, nous proposons dans ces travaux une nouvelle méthodologie de modélisation du rayonnement basée sur les méthodes de Monte Carlo et les réseaux de neurones bayésiens. Le principe du concept détaillé dans ce manuscrit commence par l'échantillonnage de points de référence tirés aléatoirement dans le milieu d'étude. Un calcul de la grandeur radiative d'intérêt est réalisée en chacun de ces points par méthode de Monte Carlo avec un modèle spectral fin. Les résultats de ces calculs sont ensuite stockés dans une base de données sur laquelle un réseau de neurones est entraîné selon un algorithme de Levenberg-Marquardt. Le paradigme bayésien nous permet de définir la topologie optimale du réseau et d'adapter automatiquement l'ensemble des paramètres inhérents au réseau sans intervention de l'utilisateur. Le réseau entraîné permet de reconstituer les champs de flux ou de divergence de flux au sein du domaine. Après une description théorique des notions mises en œuvre dans ces travaux, des résultats sur cas académiques sont présentés dans différentes configurations. Nous démontrons alors la capacité de la méthode développée à modéliser le rayonnement en milieu participatif avec une précision très élevée et des coûts de calcul faibles.

**Mots clés :** *Rayonnement, réseaux de neurones, transfert thermique, méthode de Monte Carlo*



# Abstract

Aeronautical combustion chambers are constrained environments operating at temperatures and pressures that have been steadily increasing over the last decades. Under these conditions, radiative heat transfer is significant and its influence on flame structure and wall temperatures is noticeable. It is therefore essential to be able to accurately model this transfer mode on very heterogeneous industrial cases and on mesh of several hundred million cells. Common resolution methods do not provide satisfactory accuracy in view of the growing needs. A fine resolution and the use of spectral gas models would be necessary but would imply a prohibitive consumption of computing resources.

To overcome this problem, we propose in this work a new radiation modelling methodology based on Monte Carlo methods and Bayesian neural networks. The principle of the concept detailed in this manuscript starts with the sampling of reference points randomly drawn in the study domain. The radiative quantity of interest is computed at each of these points using a Monte Carlo method and a fine spectral model. The results of these calculations are then stored in a database on which a neural network is trained using a Levenberg-Marquardt algorithm. The Bayesian paradigm allows us to define the optimal topology of the network and to automatically adapt the set of parameters inherent to the network without user intervention. The trained network allows to reconstitute the flux or flux divergence fields within the domain. After a theoretical description of the notions used in this work, results on academic cases are presented in different configurations. We then demonstrate the ability of the developed method to model radiation in a participating environment with very high accuracy and low computational costs.

**Keywords :** *Radiative heat transfer, artificial neural network, Monte Carlo method*





# Non-Remerciements

Enfin le temps béni des digressions est arrivé! Il paraîtrait qu'un manuscrit de thèse ne soit pas le lieu pour les métaphores filées illisibles comme je les aime. C'est donc bien malgré moi que je dus réfréner mes élans volubiles d'une limpidité parfois nébuleuse au bénéfice de phrases plus courtes et directes. «Sujet, verbe, complément» qu'on me disait et répétait au fur et à mesure des relectures. Quel ennui! Depuis quand un manuscrit de thèse doit-il être compréhensible par tout le monde? Je dois bien admettre Pascal que j'admire ta nature opiniâtre. Jusqu'à la dernière relecture, tu t'es afféré à souligner mes formulations avec des petites vaguelettes rouges, voire des points d'exclamation quand une impression de déjà-vu – mais non corrigée – te sautait au visage. Parfois néanmoins, je vainquais et réussissais à t'arracher, la mort dans l'âme, cette si douce mélodie : «Après tout, c'est ton style». Certes ton intervention sera sans aucun doute plébiscitée par l'ensemble des lecteurs de ce travail mais je peux dire au revoir au Goncourt pour cette année! Je ne te remercie pas pour ça.

Ta méthode Olivier, était autrement plus radicale, il faut le dire. Elle se traduisait assez simplement par le fait d'appuyer sur la touche «suppr» de ton clavier en plein dans le fichier Latex. Pragmatique et efficace. Ces intrusions étaient d'ailleurs loin d'être isolées! Combien de packages, de méthodes inconnues par le commun des mortels et de commandes personnalisées as-tu ajoutées pour que la compilation fonctionne enfin? A tel point que par ta faute, le manuscrit est devenu agréable et ergonomique, presque flatteur. Toi non plus je ne te remercie pas pour ça. Votre sérieux et votre engagement ont fait d'un brouillon austère et illisible, un manuscrit clair, fluide et pertinent pour les lecteurs s'intéressant au sujet traité. Par vos conseils, vos relectures successives et votre expertise, je peux maintenant présenter et archiver, un travail que je suis fier d'avoir rédigé; un travail ponctuant trois années de recherche et bien plus encore d'apprentissage. Vous comprenez ce que ça signifie? Que des gens vont sûrement le lire maintenant! Franchement Bravo!

Mais je n'en ai pas fini avec vous. Ni avec toi Daria, d'ailleurs. Vous trois qui par votre confiance, votre bienveillance et votre sympathie, avez eu l'outrecuidance de faire de ces trois dernières années, une excursion épanouissante et réjouissante au pays de la recherche industrielle. La porte du club - pourtant assurément trop permissif - des doctorants traumatisés m'est maintenant définitivement fermée. Rien ne m'a été épargné. A tous les badauds qui me demandaient avec un tressaillement de compassion dans la voix : «alors ta thèse, ça va, ce n'est pas trop dur?», j'étais contraint de répondre que tout allait bien au contraire avec un sourire sincère par-dessus le marché! La faute originelle te revient assurément Daria. Depuis quand une tutrice de stage suit un étudiant dans ses délires et le laisse construire lui-même son sujet de thèse sur ce qu'il a exactement envie de faire? Quand bien même le stage en question se serait bien passé, depuis quand une ingénieure d'un grand groupe mondiale s'engage aussi franchement dans un projet pareil? Depuis quand permet-on à un candidat de choisir son laboratoire lui assurant à la fois un encadrement exemplaire et un équilibre de vie idéal? C'est vraiment du grand n'importe quoi.

Sérieusement. Qu'est ce qui vous a pris? Normalement, une histoire pareille est racontée un peu comme une légende urbaine. «J'ai connu un jour un doctorant qui» mais ça n'arrive jamais à personne. Vous avez été formidables du début à la fin. Sachez que je suis tout à fait conscient de la chance que j'ai eue. Ces remerciements aussi informels soient-ils, sont tout à fait sincères. On dit parfois qu'une thèse est un moment particulier pour un ingénieur et dont il faut profiter. Dans mon cas et au-delà de toutes

les connaissances théoriques acquises, c'est sans aucun doute l'aventure humaine à vos côtés qui me laissera le souvenir le plus impérissable. Merci pour tous.

Passons maintenant à mon copain Thomas. Le deuxième comparse du duo « tic & tac ». Dans une jolie histoire comme celle-ci, Robinson a toujours son Vendredi sans qui les journées paraissent fades et sans saveur. Tantôt brillant scientifique, tantôt larron à l'humour caustique, je dois dire que l'aventure n'aurait pas été la même sans toi. Tu es un vrai emmerdeur mais j'aurais certainement fini par manger tous les cailloux du plateau de Saint-Front sans toi. On est passé à deux doigts du CAC40 mais ce n'est que partie remise. Le récit ne s'arrête pas là, j'en suis certain. La suite va nous surprendre si tu veux mon avis. Tu n'es pas le seul à avoir contribué à tout ça. Je peux aussi citer tous les vieux de la vieille, présents depuis un certain nombre d'années maintenant et qui, malgré un éloignement géographique certain, m'ont soutenu moralement quand il le fallait. Antoine, Laurent, Benate, Florent, Benjamin. A cause de vos conneries, j'ai parfois bien eu du mal à rédiger mon manuscrit au bénéfice d'une ou deux parties de WZ. Ou trois. J'ai de la chance de vous avoir.

Terminons à présent par là où tout a commencé. D'ailleurs à ce sujet, si l'un d'entre vous a des doléances à soumettre sur votre serviteur, ils gèrent le service après-vente. Je parle forcément de mes parents, de ma famille et ma compagne. Ils sont responsables de tout. Du pire, comme du meilleur. Les premiers m'ont élevé de leur mieux et m'ont notamment inculqué le goût de l'effort, l'importance de la curiosité et du dépassement de soi. Il ne fait aucun doute que cette réussite et avant tout la vôtre. J'espère que vous en êtes fiers autant que je le suis d'être votre fils. Merci pour tout. J'ai globalement eu la chance d'être excessivement bien entouré, il faut le savoir. Enfin, Ambre, toi qui partages ma vie au quotidien, tu es celle qui mérite sans doute le plus d'être citée ici. En fait, je dirais même que ce manuscrit ne serait toujours pas terminé sans toi. Sans ton intervention, je serais sûrement encore en train de remettre à plus tard telle correction ou la rédaction de tel paragraphe. Ta ténacité est remarquable parce que tous ceux qui me connaissent savent à quel point il est difficile de faire bouger la bourrique que je suis. Mais en réalité tu es beaucoup plus que cela. Tu es celle qui équilibre la balance de mon existence. Le remède à mes travers et l'onguent de mes douleurs. Tu es celle qui rend mes rêves réalisables, celle qui tempère mais aussi celle qui encourage et soutien. J'ai beaucoup de chance de t'avoir bien que je ne te le dise que trop peu. Maintenant, c'est écrit. Merci d'être toi. Je t'aime.

# Table des matières

Résumé	i
Abstract	iii
Non-Remerciements	v
Table des figures	ix
Liste des tableaux	xiii
Introduction	1
<b>1 Enjeux, notions et modèles du rayonnement pour l'industrie aéronautique</b>	<b>5</b>
1.1 Fonctionnement d'une turbomachine et définition du besoin	7
1.1.1 Un rapide historique	7
1.1.2 Architectures de base d'un turboréacteur et rendement thermique	7
1.2 Grandeurs radiatives fondamentales	10
1.3 Rayonnement thermique en chambre de combustion : état de l'art	12
1.3.1 Formulation différentielle de l'équation de transfert radiatif	13
1.3.2 Résolution numérique de l'ETR sous forme différentielle par la méthode des Ordonnées Discrètes	14
1.3.3 Les modèles globaux de rayonnement des gaz	15
1.4 Approche stochastique par méthode de Monte Carlo et modèles spectraux de gaz	17
1.4.1 Formulation intégrale de l'Équation de Transfert Radiatif	17
1.4.2 Résolution de l'ETR sous forme intégrale par Méthode de Monte Carlo en milieu hétérogène : introduction des collisions nulles	18
1.4.3 Expression des grandeurs radiatives d'intérêt	21
1.4.4 Modèles de gaz spectraux en milieu participant	22
1.4.5 Choix des modèles utilisés et solution nouvelle de modélisation du rayonnement	31
Références	31
<b>2 Une solution de rupture : l'intelligence artificielle pour la modélisation physique</b>	<b>37</b>
2.1 Du cerveau aux réseaux de neurones artificiels : origines et modèle humain	38
2.2 Les réseaux de neurones artificiels modernes	40
2.2.1 Un premier modèle de neurone moderne « on-off » : le perceptron	40
2.2.2 Grands principes heuristiques de construction	40
2.2.3 Les réseaux modernes : architecture et composition générale	41
2.2.4 Algorithme par descente de gradient : description, intuitifs et explication d'une méthode d'entraînement incontournable	47
2.2.5 Le sur-apprentissage : le mal de l'algorithme par descente de gradient	53
2.2.6 Régularisation des poids : principes et techniques	55

2.3	Un réseau adaptée aux contraintes industrielles . . . . .	56
2.3.1	Les contraintes d'un outil industriel . . . . .	56
2.3.2	Une optimisation de la descente de gradient : algorithme de Levenberg-Marquardt . . . . .	57
2.3.3	Paradigme et régularisation bayésienne . . . . .	60
	Références . . . . .	67
<b>3</b>	<b>Modélisation du rayonnement par réseau de neurones bayésien et algorithme de Monte Carlo</b>	<b>69</b>
3.1	Rappel du constat et proposition . . . . .	70
3.2	Interfaçage des méthodes : Monte Carlo pour la base de données et IA pour interpolation . . . . .	71
3.2.1	Un modèle réduit statistique : quelques analogies . . . . .	71
3.2.2	Procédure : du calcul stochastique par point à l'interpolation en champ . . . . .	72
3.2.3	Tirage des points de référence . . . . .	74
3.3	Synthèse de la méthode . . . . .	76
	Références . . . . .	78
<b>4</b>	<b>Evaluation de la méthode sur cas académiques</b>	<b>81</b>
4.1	Choix des modèles physiques et du réseau de neurones . . . . .	82
4.1.1	Modélisation du coefficient d'absorption . . . . .	82
4.1.2	Réseaux de neurones utilisés . . . . .	83
4.2	Etude d'un cas axisymétrique homogène de Benchmark . . . . .	84
4.2.1	Validation du code radiatif en milieu homogène . . . . .	84
4.2.2	Validation du concept méthodologique : entraînement et interpolation par réseau de neurones . . . . .	84
4.3	Cas hétérogène : étude d'une flamme de diffusion simplifiée . . . . .	86
4.3.1	Paramétrisation . . . . .	86
4.3.2	Validation du code de Monte Carlo sur cas hétérogène . . . . .	88
4.3.3	Influence de la méthode de tirage des points de référence . . . . .	90
4.3.4	Comparaison avec d'autres méthodes d'interpolation . . . . .	94
4.3.5	Etude du cas de référence à haute pression . . . . .	96
4.4	Prise en compte de perturbations . . . . .	98
4.4.1	Point chaud . . . . .	100
4.4.2	Gestion du bruit numérique . . . . .	103
4.5	Ouverture et perspective . . . . .	106
4.5.1	Passage aux applications industrielles et couplage . . . . .	106
4.5.2	Un interpolateur complexe global . . . . .	108
	Références . . . . .	109
	<b>Conclusion et Perspectives</b>	<b>111</b>
	<b>A Cycle thermodynamique d'un turboréacteur</b>	<b>115</b>
	<b>B Champs soustraits : comparaisons des méthodes d'interpolation</b>	<b>119</b>
	<b>C Tableaux de valeurs</b>	<b>121</b>
	<b>D Script du papier présenté à la conférence IHTC-17</b>	<b>127</b>
	Références . . . . .	138
	<b>Bibliographie</b>	<b>139</b>

# Table des figures

<b>1</b>	<b>Enjeux, notions et modèles du rayonnement pour l'industrie aéronautique</b>	<b>5</b>
1.1	Schéma simplifié d'une turbomachine	8
1.2	Schéma simplifié d'une turbomachine <i>double corps</i>	8
1.3	Cycle idéal de Brayton d'une turbomachine simplifiée	9
1.4	Résolution de proche en proche de la luminance à travers un volume de contrôle selon la direction $\mathbf{u}$	15
1.5	Homogénéisation des coefficients d'absorption par un majorant	19
1.6	Exemple de chemin parcouru par un photon dans un champ majoré et uniformisé	21
1.7	Exemple de réorganisation d'une portion de spectre d'absorption pour un mélange de gaz $H_2O - N_2$ à pression atmosphérique (issu de PIERROT [1])	24
<b>2</b>	<b>Une solution de rupture : l'intelligence artificielle pour la modélisation physique</b>	<b>37</b>
2.1	Analogie entre neurone biologique et artificiel	39
2.2	Le perceptron	40
2.3	Un exemple de FANN à une couche cachée	42
2.4	Exemples de fonctions d'activation utilisées pour la regression	43
2.5	Notations des poids et des biais	45
2.6	Évolution recherchée de $\eta$	50
2.7	Cas problématiques en début (gauche) ou fin d'algorithme (droite)	50
2.8	Présence d'un minimum local	50
2.9	<i>Feedforward Pass</i> : calcul des sorties de chaque neurone et des erreurs	51
2.10	<i>Backward Pass</i> : Rétro-propagation des sensibilités	52
2.11	Logigramme de l'algorithme d'entraînement par descente de gradient	54
2.12	Exemple de sortie d'un réseau de neurone correctement entraîné (gauche) et en sur-apprentissage (droite). Les points rouges représentent la réponse du réseau à des données d'entraînement et les points verts, celle à des données inconnues	54
2.13	Représentation de la MSE sur les points d'entraînement (bleu) et sur les points de validation (vert) en cas d'overfitting	55
2.14	Logigramme de l'algorithme de Levenberg-Marquardt	61
2.15	Logigramme de l'algorithme <i>LM</i> couplé à la régularisation bayésienne	66
<b>3</b>	<b>Modélisation du rayonnement par réseau de neurones bayésien et algorithme de Monte Carlo</b>	<b>69</b>
3.1	Comparaison de la disparité des points dans un carré unité tirés aléatoirement ou suivant une séquence de Halton	75
3.2	Logigramme complet de l'algorithme développé	77
<b>4</b>	<b>Evaluation de la méthode sur cas académiques</b>	<b>81</b>
4.1	Géométrie étudiée - Cas homogène	84
4.2	Validation de l'algorithme de calcul par Monte Carlo et SNB-CK sur cas homogène à une température de 1200 K- Comparaison avec les données de référence [2]	85

4.3	Prédiction de la divergence du flux radiatif sur l'axe de rotation du domaine par réseau de neurones - Cas homogène à $T = 1200\text{ K}$ . . . . .	87
4.4	Géométrie étudiée - Cas hétérogène . . . . .	88
4.5	Représentation des champs thermophysiques du cas de benchmark hétérogène [2] . . . . .	89
4.6	Validation du code Monte Carlo en collision nulle - Divergence du flux radiatif sur l'axe de révolution du domaine ( $r = 0$ ) en $kW.m^{-3}$ . . . . .	90
4.7	Construction du champ de divergence de flux par réseaux de neurones - Cas hétérogène, tirage aléatoire de 500 points de références . . . . .	92
4.8	Soustraction du champ interpolé à celui de référence et visualisation des 500 points d'entraînement aléatoires (échelle en $kW.m^{-3}$ ) . . . . .	93
4.9	Soustraction du champ interpolé à celui de référence et visualisation des 500 points d'entraînement selon une suite Halton (Echelle en $kW.m^{-3}$ ) . . . . .	94
4.10	Comparaison entre les données de benchmark et les interpolations par réseaux de neurones sur la ligne centrale ( $r = 0$ ) - Base de données de 500 points . . . . .	95
4.11	Comparaison entre les différentes méthodes d'interpolation - Résultats sur la ligne centrale du cylindre ( $r = 0$ ) . . . . .	97
4.12	Divergence du flux radiatif sur la ligne centrale ( $r = 0$ ) à haute pression ( $P = 22\text{ bars}$ ) et en présence de $CO$ . . . . .	99
4.13	Champs thermophysiques - Cas « Point Chaud » . . . . .	101
4.14	Saturation du réseau en l'absence de pré-traitement du champ thermophysique - Tracé de la divergence de flux ( $kW.m^{-3}$ ) sur la ligne axiale ( $r = 0.15$ ) . . . . .	102
4.15	Comparaison entre les interpolations du réseau de neurones et les valeurs de référence sur les lignes de contrôle - Cas « Point chaud » . . . . .	104
4.16	Champs thermophysiques - Cas bruité . . . . .	105
4.17	Divergence du flux radiatif sur la ligne centrale ( $r = 0$ ) en cas de bruit numérique gaussien . . . . .	107
<b>A</b>	<b>Cycle thermodynamique d'un turboréacteur</b> . . . . .	<b>115</b>
A.1	Cycle idéal de Brayton d'une turbomachine simplifiée . . . . .	116
<b>B</b>	<b>Champs soustraits : comparaisons des méthodes d'interpolation</b> . . . . .	<b>119</b>
B.1	Champs de divergence de flux en $kW.m^{-3}$ calculés par méthodes d'interpolation et champs soustraits . . . . .	120
<b>C</b>	<b>Tableaux de valeurs</b> . . . . .	<b>121</b>
<b>D</b>	<b>Script du papier présenté à la conférence IHTC-17</b> . . . . .	<b>127</b>
D.1	Building geometry and rooftop photovoltaic field. . . . .	129
D.2	Building geometry and conducto-convective-radiative transfers at interfaces . . . . .	133
D.3	Neuron representation and two layer neural network. . . . .	133
D.4	Weather data (global horizontal irradiance, air temperature and wind speed) for 2019 for Las Vegas, Cape Town and Nancy. For the radiation data the daily maximum is displayed and for the air temperature the daily average is plotted. . . . .	135
D.5	Monte Carlo simulation points and FANN interpolation for Las Vegas and Cape Town in 2019 summer periods for the concrete roof configuration with photovoltaic field. . . . .	136
D.6	Monte Carlo simulation points and FANN interpolation for Las Vegas and Cape Town in 2019 summer periods for the concrete roof configuration without photovoltaic field. . . . .	137

D.7 Las Vegas, 2019 Summer week : Monte Carlo simulation in black color (reference solution), interpolation using TMY with the FANN in red color. Configuration with photovoltaic field, green roof layer. . . . . 137





# Liste des tableaux

<b>1</b>	<b>Enjeux, notions et modèles du rayonnement pour l'industrie aéronautique</b>	<b>5</b>
1.2	Evolution des caractéristiques de la turbine Westinghouse 501 - Extrait de SARAVANAMUTTOO et al. [3]	10
1.3	Quadrature de Gauss Legendre à 16 points	27
1.4	Définition des bandes spectrales dans le modèle HPBM26	30
<b>2</b>	<b>Une solution de rupture : l'intelligence artificielle pour la modélisation physique</b>	<b>37</b>
<b>3</b>	<b>Modélisation du rayonnement par réseau de neurones bayésien et algorithme de Monte Carlo</b>	<b>69</b>
<b>4</b>	<b>Evaluation de la méthode sur cas académiques</b>	<b>81</b>
4.2	Prédiction des réseaux de neurones à $T = 1200$ K	85
4.3	Qualité des interpolations sur l'ensemble du domaine de la divergence de flux par réseaux de neurones en milieu hétérogène - Tirage uniforme	91
4.4	Qualité des interpolations par réseaux de neurones en milieu hétérogène - Tirage de Halton	92
4.5	Comparaison de la valeur de la RMSE calculée pour les différentes méthodes d'interpolation	96
4.6	Paramètres et caractéristiques de l'entraînement par ANN - Cas à haute pression	98
4.7	Caractéristiques du réseau de neurones et de l'entraînement - Cas «Point chaud»	101
4.8	Caractéristiques du réseau de neurones et de l'entraînement - Cas bruité	106
<b>A</b>	<b>Cycle thermodynamique d'un turboréacteur</b>	<b>115</b>
<b>B</b>	<b>Champs soustraits : comparaisons des méthodes d'interpolation</b>	<b>119</b>
<b>C</b>	<b>Tableaux de valeurs</b>	<b>121</b>
C.1	Valeurs de la divergence du flux sur la ligne centrale $r = 0$ - Cas à haute pression	122
C.2	Divergence du flux sur la ligne centrale $r = 0$ - Cas «Point chaud»	123
C.3	Divergence du flux sur la ligne longitudinale $r = 0.15$ - Cas «Point Chaud»	124
C.4	Divergence du flux sur la ligne radiale $x = 0.1$ - Cas «Point Chaud»	125
<b>D</b>	<b>Script du papier présenté à la conférence IHTC-17</b>	<b>127</b>
D.1	Thermal properties of the green roof structure.	129
D.2	Energy evaluation of the building's energy needs for the concrete configuration without photovoltaic field in 2019 and comparison between the Monte Carlo algorithm ( $10^6$ realizations) and the FANN.	136



# Introduction

## Modélisation du rayonnement thermique en contexte industriel : les besoins grandissants de l'industrie aéronautique

Dans le contexte environnemental et sociétal actuel, l'industrie aéronautique doit se réinventer. La diminution des émissions de gaz à effet de serre et de polluants par les turbomoteurs est un enjeu fondamental pour le groupe Safran. En tant que premier motoriste aéronautique du monde<sup>1</sup>, l'entreprise investit depuis longtemps pour l'augmentation des rendements et la maîtrise de l'empreinte de ses réacteurs sur l'environnement. Parallèlement à l'augmentation des performances globales des moteurs, il est nécessaire d'assurer les plus hauts niveaux de sécurité aux voyageurs et la plus grande opérabilité possible des systèmes aux aviateurs. Cette difficile équation implique pour Safran de se doter des meilleurs outils possibles pour la conception de ses produits et d'investir dans l'amélioration constante de ses procédés de modélisation. C'est dans ce contexte que s'inscrit l'étude proposée dans ce manuscrit.

Plus particulièrement, l'augmentation des rendements implique de récupérer autant d'énergie que possible dans le processus de combustion du carburant. Une étude thermodynamique d'une turbomachine simplifiée permet d'exprimer le rendement thermique du cycle de ce type de système et démontre notamment l'importance de la température de flamme [3]. On constate ainsi sur les dernières décennies une augmentation constante des températures atteintes au sein des chambres de combustion jusqu'à approcher les températures de flamme adiabatique. Par conséquent, les contraintes sur les matériaux sont de plus en plus élevées à tel point qu'aujourd'hui, une grande partie de la recherche industrielle se porte sur la résistance thermique des organes sensibles des moteurs. Les transformations chimiques des gaz de combustion sont également affectées par l'augmentation de la température de flamme. La gestion de la production des NOx thermiques devient par exemple un enjeu majeur pour la conception des chambres.

Afin de pouvoir saisir avec précision l'impact des transferts de chaleur entre les gaz de combustion et la chambre, il devient donc primordial d'être capable de caractériser et de modéliser finement les phénomènes en action. Plus particulièrement, à mesure que la température augmente, les transferts par rayonnement croissent et leur impact sur les températures de parois devient particulièrement sensible. Jusqu'à présent, ce mode de transfert était approximé par des méthodes efficaces d'un point de vue numérique mais occasionnant des erreurs parfois importantes et surtout difficilement quantifiables sur les différents flux. Les méthodes natives des logiciels comme Fluent par exemple sont adaptées au travail sur maillage et peuvent être très facilement couplées aux méthodes de résolution CFD. En particulier, le rayonnement était jusqu'à présent estimé par la méthode des Ordonnées Discrètes et un modèle de somme de gaz gris [4].

Bien que très utilisés, cette méthode de résolution de l'équation de transfert radiatif (*ETR*) et ce modèle de gaz ne permettent pas la prise en compte précise du rayonnement dans un contexte industriel. En

<sup>1</sup>Pour les avions de + de 100 places, <https://www.safran-group.com/fr/societes/safran-aircraft-engines>

effet, il est important de noter que les cas d'étude dont il est question ici sont particulièrement complexes à bien des égards. Tout d'abord, la résolution des transferts radiatifs est couplée à la résolution de la mécanique des fluides, de la chimie et des autres modes de transfert thermique. De plus, les maillages utilisés sur les modèles de chambres contiennent en général plus de 250 millions de cellules du fait de la complexité géométrique de ces éléments. Les cas d'études sont donc particulièrement lourds et difficiles à traiter numériquement.

Les difficultés intrinsèques des cas d'étude envisagés imposent ainsi de sérieuses limites à l'utilisation des méthodes par Ordonnées Discrètes et de modèles de gaz plus précis. En effet, chaque nouveau gaz gris ou bande spectrale supplémentaire implique avec cette méthode une résolution et le stockage de l'information calculée pour chaque cellule du maillage. Il en est de même avec la discrétisation angulaire de l'ETR. Autrement dit et malgré le vrai potentiel des méthodes par Ordonnées Discrètes pour le calcul en champ, la complexité des cas d'études ne permet pas techniquement de modéliser avec une précision satisfaisante et pour un coût acceptable les transferts par rayonnement. C'est dans ce contexte que s'inscrivent les travaux réalisés pendant cette thèse menée en association avec le Laboratoire d'Énergies & Mécanique Théorique et Appliquée (LEMETA) de Nancy dans le cadre d'un contrat CIFRE. Afin de palier aux problématiques exposées précédemment, nous proposons une nouvelle méthode de modélisation en rupture avec les procédés existants.

## **Les réseaux de neurones en complément de la physique : un interpolateur universel puissant pour la généralisation de solutions ponctuelles**

La démarche proposée dans ce manuscrit vise à combiner les avantages des méthodes de Monte Carlo et des réseaux de neurones. Les premières sont des méthodes stochastiques basées sur la théorie des grands nombres. Appliquées au rayonnement, elles permettent de calculer des grandeurs radiatives intégrées comme la luminance totale, le flux à la paroi ou encore la divergence du flux radiatif, sans approximation dans leur formulation [5, 6]. Leur application à des conditions hétérogènes a été étendue dans les travaux de GALTIER et al. [7]. L'expression en collisions nulles qu'il propose pour le rayonnement thermique permet en effet de traiter efficacement des champs complexes sans concession sur la rigueur mathématique de la formulation de l'ETR. Les méthodes de Monte Carlo fournissent également une estimation de l'erreur sur la valeur calculée de l'observable et assure donc un certain contrôle sur la qualité de la modélisation. Enfin, leur insensibilité à la complexité géométrique [8, 9] et à la complexité du modèle de gaz [7] permet leur utilisation sur des cas complexes tout en assurant un degré de précision élevé. Néanmoins et malgré tous ces points forts, les méthodes de Monte Carlo donnent des résultats ponctuels. En d'autres termes, elles permettent d'estimer des grandeurs avec une grande précision en des points précis du domaine. Leur utilisation pour la génération de champs complets et continus de grandeurs radiatives est donc inenvisageable vis à vis du temps de calcul nécessaire sur des cas aussi complexes que ceux envisagés.

C'est à ce moment que les réseaux de neurones interviennent. Ces outils informatiques sont communément classés dans le domaine de l'Intelligence Artificielle, à tort ou à raison d'ailleurs. En réalité, ou tout du moins tels qu'ils sont utilisés ici, les réseaux de neurones doivent plutôt être considérés comme des outils de statistiques avancées permettant de construire des corrélations implicites entre des jeux de données connus. De ce point de vue, et bien que leurs capacités ne s'arrêtent pas à cette tâche, ils jouent le rôle d'interpolateurs complexes et universels [66, 67].

Dans le cadre de ces travaux, les réseaux de neurones sont utilisés pour généraliser les résultats obtenus par méthode de Monte Carlo. L'idée est avant tout de générer une base de données solide grâce à cette méthode de résolution, combinée avec un modèle de gaz spectral de référence. Des points sont

ainsi échantillonnés dans le domaine d'étude où la grandeur radiative d'intérêt est calculée en fonction des conditions thermophysiques locales. Une base contenant les variables ayant une influence sur le rayonnement et la grandeur radiative correspondante est alors construite. Un réseau de neurones est enfin entraîné sur ces jeux de données. L'outil est alors assimilable à un modèle réduit de la physique du rayonnement ou à une fonction de transfert spécifique au cas d'étude. Il est capable d'estimer en tout point du domaine la grandeur d'intérêt par interpolation sur l'ensemble de son espace des phases.

Au delà de la précision de la méthode développée, il est nécessaire qu'elle respecte les contraintes liées à son contexte d'utilisation. L'outil doit en particulier être *clés en main* et ne doit pas nécessiter de compétence particulière en intelligence artificielle pour être utilisé par le plus grand nombre. Le choix du paradigme bayésien appliqué aux réseaux de neurones a donc été fait [68]. Cette méthode particulière permet d'automatiser à la fois l'entraînement, la gestion de la topologie et de l'ensemble des hyperparamètres du réseaux, si bien que l'outil est capable de s'adapter aux données qu'il reçoit sans intervention de l'utilisateur et sans tendance au surapprentissage [69]. En fonction du cas d'étude, le réseau peut ainsi modifier le nombre de neurones qu'il utilise et garantir un compromis optimal entre temps de calcul et précision de l'interpolation.

## Preuve de concept et validation méthodologique sur cas académiques

Au delà des choix et de la définition conceptuelle de l'outil proposé, les travaux menés au cours de cette thèse ont abouti au développement de codes de calcul permettant la mise à l'épreuve pratique de la méthode. Concernant le code de Monte Carlo, nous nous sommes appuyés sur la librairie open-source *star-engine* développée par la société *Méso-star* et plus particulièrement, sur le code HTRDR [10] initialement développé pour l'étude du rayonnement dans l'atmosphère et la visualisation de nuages. Un important travail a été fourni pour adapter ces outils aux applications de combustion qui nous intéressent avec notamment l'intégration des modèles de gaz choisis pour ces travaux.

Une librairie spécifiquement dédiée aux réseaux de neurones a également été entièrement développée. Elle permet notamment de gérer les données de la base (normalisation, remise à l'échelle, etc.), de construire, d'adapter et d'entraîner des réseaux bayésiens, ainsi que d'interpoler les résultats recherchés. Ce code a été pensé pour pouvoir être utilisé sur n'importe quel cas d'étude sans se limiter aux applications traitées ici.

Avant de déployer la solution proposée, un travail complet de validation sur des cas académiques a été mené. Ce manuscrit compile l'ensemble des cas significatifs traités. Nous démontrons notamment que les réseaux de neurones, à condition d'être bien entraînés, sont capables d'estimer avec une grande précision des grandeurs radiatives comme le flux et sa divergence. Une preuve de concept forte est donc apportée par ces travaux. Bien que le déploiement sur cas industriel n'ait pas pu aboutir dans la limite du temps disponible, il ne fait pas de doute que l'utilisation des réseaux de neurones possède un potentiel important pour lever les verrous techniques parfois rencontrés en modélisation.

## Structure et plan du manuscrit

Ce manuscrit a pour objectif principal de présenter la méthode développée avec précision et détails, ainsi que les résultats obtenus sur cas académiques. Au delà de la mise en œuvre purement technique des outils utilisés, un soin particulier a été porté à la transmission et à la compréhension des concepts liés en particulier aux réseaux de neurones. Nous avons tenté de donner dans ces pages, des images et des intuitifs compréhensibles aux lecteurs afin qu'ils puissent assimiler conceptuellement ces outils.

Le travail présenté ici s'articule autour de quatre parties complémentaires. Le chapitre 1 propose tout d'abord de poser le contexte de la présente étude et d'en décrire les enjeux. Cette section est suivie d'une introduction aux grandeurs fondamentales de la physique du rayonnement, puis d'un rapide état de l'art des méthodes et modèles actuellement utilisés dans l'entreprise. Une partie plus détaillée relative à la formulation intégrale de l'équation de transfert radiatif, aux méthodes de Monte Carlo et à l'expression des grandeurs associées vient compléter cette présentation méthodologique. Enfin, un descriptif non exhaustif des principaux modèles de gaz spectraux en lien avec notre étude vient clore cette première partie.

Le chapitre 2 est quant à lui réservé exclusivement aux réseaux de neurones. Dans la mesure où ces outils ne sont encore que peu utilisés en modélisation physique, nous avons souhaité proposer une présentation assez complète et didactique. L'accent a été mis dans un premier temps sur la présentation de concepts et d'images permettant avant tout de comprendre comment fonctionnent qualitativement les réseaux de neurones au travers de grands principes. Une description plus concrète et quantitative est ensuite détaillée. Les principales caractéristiques des réseaux et de leurs méthodes d'entraînement sont alors évoquées. Plus précisément, le paradigme Bayésien et toutes ses conséquences sur la formulation des algorithmes utilisés sont développés aussi bien conceptuellement que mathématiquement.

Dans le chapitre 3, nous présentons le principe de la méthode ainsi que sa mise en pratique concrète. Les choix des modèles et méthodes utilisés sont notamment explicités et justifiés. Un logigramme reprend également de manière graphique le déroulement de tout le processus de modélisation d'un champ de grandeur radiative.

Enfin, tous les résultats obtenus sur cas académiques sont listés dans le chapitre 4. Nous présentons notamment les étapes de validation des différents codes par comparaison avec des données de référence de la littérature. Les résultats obtenus par notre méthode sont également détaillés, analysés et commentés à la fois sur cas homogène et hétérogène. Des perturbations dans les champs thermophysiques ont été étudiées dont un cas à haute pression se rapprochant des conditions d'opération observées dans les chambres de combustion aéronautiques. L'objectif était de mettre à l'épreuve la robustesse et l'adaptabilité de l'outil.

# Enjeux, notions et modèles du rayonnement pour l'industrie aéronautique

## Sommaire du chapitre

1.1	Fonctionnement d'une turbomachine et définition du besoin . . . . .	7
1.2	Grandeurs radiatives fondamentales . . . . .	10
1.3	Rayonnement thermique en chambre de combustion : état de l'art . . . . .	12
1.4	Approche stochastique par méthode de Monte Carlo et modèles spectraux de gaz . . . . .	17
	Références . . . . .	31

## Notations

Symbole	Description	Unité
$a_j$	Poids des gaz gris	-
$c$	Célérité de la lumière dans le milieu	$\text{m s}^{-1}$
$c_0$	Célérité de la lumière dans le vide	$\text{m s}^{-1}$
$\bar{\delta}$	Espacement moyen entre les raies	-
$d\dot{Q}_\eta$	Puissance élémentaire rayonnée	W
$\partial_t$	Dérivée temporelle	-
$\epsilon$	Emissivité totale	-
$\mathbb{E}$	Espérance	-
$f_v$	Fraction volumique de suie	$\text{m}^3 \text{m}^{-3}$
$g$	Distribution statistique cumulée des coefficients d'absorption	-
$g_i$	Point de quadrature de Gauss Legendre	-
$\bar{\gamma}$	Demi-largeur moyenne de raie	-
$I$	Nombre de points de quadrature de Gauss Legendre	-
$\kappa_{abs,\eta}$	Coefficient d'absorption monochromatique	$\text{m}^{-1}$
$\widehat{\kappa}_\eta$	Coefficient d'absorption majorant	$\text{m}^{-1}$
$\widehat{\kappa}_{n,\eta}$	Coefficient d'absorption de collision nulle	$\text{m}^{-1}$
$\kappa_{abs,\eta}^{suié}$	Coefficient d'absorption monochromatique des suies	$\text{m}^{-1}$
$k_B$	Constante de Boltzmann	$\text{J K}^{-1}$

Symbole	Description	Unité
$\bar{k}$	Intensité moyenne de raie	-
$l_c$	Longueur de la colonne de gaz	m
$L_\eta^0$	Luminance du corps noir monochromatique	$\text{W m}^{-2} \text{sr}^{-1} \text{cm}$
$L_\eta$	Luminance monochromatique	$\text{W m}^{-2} \text{sr}^{-1} \text{cm}$
$L$	Luminance totale	$\text{W m}^{-2} \text{sr}^{-1}$
$M$	Emittance	$\text{W m}^{-2}$
$M^0$	Emittance du corps noir	$\text{W m}^{-2}$
$n$	Indice optique	-
$\nabla_{\mathbf{x}}$	Gradient spatial	-
$\eta$	Nombre d'onde	cm
$\hat{n}$	Normale sortante à une surface	-
$N_{mc}$	Nombre de réalisations de Monte Carlo	-
$\eta_{th}$	Rendement thermique	-
$d\Omega$	Angle solide élémentaire	-
$p$	Pression totale	Pa
$p_X(\mathbf{x})$	Densité de probabilité d'absorption	-
$\overline{p_X(\mathbf{x})}$	Densité de probabilité d'absorption majorante	-
$q$	Flux radiatif	$\text{W m}^{-2}$
$\nabla \cdot q$	Divergence du flux radiatif	$\text{W m}^{-3}$
$r_i$	Nombre aléatoire entre 0 et 1	-
$S_{abs}$	Terme puits d'absorption	$\text{W m}^{-3} \text{sr}^{-1} \text{cm}$
$s$	Abscisse curviligne	-
$S_e$	Terme source d'émission	$\text{W m}^{-3} \text{sr}^{-1} \text{cm}$
$\sigma$	Constante de Stefan-Boltzmann	$\text{W m}^{-2} \text{K}^{-4}$
$S$	Élément de surface	$\text{m}^2$
$T$	Température	K
$t$	Temps	s
$T_1$	Température des gaz en entrée du moteur	K
$T_2$	Température des gaz en sortie de compresseur	K
$T_3$	Température des gaz en entrée de turbine	K
$T_4$	Température des gaz d'éjection	K
$\theta$	Angle formé entre la normale sortante $\hat{n}$ et la direction $\mathbf{u}$	rad
$\tau_\eta$	Transmissivité spectrale	-
$\bar{\tau}_\eta$	Transmissivité moyenne à travers un colonne de gaz homogène et isotherme	-
$\bar{\tau}_{\eta_{eq}}$	Transmissivité moyenne équivalente d'un mélange gazeux	-
$\mathbf{u}$	Vecteur direction	-
$\omega_i^{GL}$	Poids de quadrature de Gauss Legendre	-
$\mathbf{x}$	Vecteur de coordonnées spatiales	m



## 1.1 Fonctionnement d'une turbomachine et définition du besoin

---

### 1.1.1 Un rapide historique

---

Les turbomachines sont incontestablement une des inventions les plus importantes du XX<sup>ème</sup> siècle. Originellement développés avant la Seconde Guerre Mondiale pour la génération d'électricité, ces moteurs connurent un développement massif pendant ce conflit armé pour équiper les avions de chasse de l'Axe et des Alliés. Ces moteurs qui étaient alors assourdissants, à très faible rendement et dangereux, sont malgré tout devenus peu à peu le standard de l'industrie aéronautique civile et militaire.

A partir des années 1970, les moteurs simple flux ont laissé place aux réacteurs à double flux pour l'industrie civile (voir section 1.1.2). Les avions qui avaient ainsi été développés en premier lieu pour faire la guerre sont alors devenus une incroyable porte d'accès au monde au fur et à mesure des années. Au début du XXI<sup>ème</sup> siècle, ces avancées technologiques ont permis le développement de ces moteurs pour la génération de puissance avec des rendements autour de 40 % [3]. Des déclinaisons de différents moteurs aéronautiques ont été adaptées aux contraintes terrestres pour la production d'électricité à partir d'hydrocarbures notamment.

Aujourd'hui, le contexte environnemental et sociétal impose à l'industrie de se réinventer. Des efforts considérables ont été menés dans les dernières années pour optimiser les rendements de ces machines et diminuer la production de polluants. En guise d'exemple, les derniers moteurs civils de CFM International<sup>1</sup> permettent une réduction de 15 % de la consommation de carburant et d'émissions de CO<sub>2</sub> et jusqu'à 50 % de réduction des émissions d'oxydes d'azote par rapport à la précédente génération de réacteur. La filière investit également pour préparer l'avenir de l'aviation en contribuant activement à beaucoup de projets de recherche visant à remplacer les carburants fossiles. Les SAF<sup>2</sup>, les poudres métalliques et l'hydrogène décarboné sont par exemple des pistes prometteuses.

### 1.1.2 Architectures de base d'un turboréacteur et rendement thermique

---

La classe des turbomachines est vaste et regroupe un nombre important de technologies différentes. Nous nous intéressons ici principalement au fonctionnement général d'un turboréacteur aéronautique moderne. Un moteur de ce type est composé au minimum de trois éléments :

- Un compresseur, qui selon les technologies peut être axial ou radial,
- Une chambre de combustion,
- Une turbine reliée au compresseur par un axe de transmission.

La première étape dans un cycle thermodynamique de turboréacteur consiste à comprimer les gaz entrant dans le système grâce au compresseur. En considérant les pertes comme négligeables et les transformations comme isentropiques, la détente de ces gaz dans la turbine permettrait uniquement de fournir la puissance nécessaire à l'alimentation du compresseur. Afin de produire une force de poussée suffisante à l'aéronef, il est nécessaire d'apporter une énergie supplémentaire aux gaz comprimés par une augmentation de leur température avant expansion. Cet effet est obtenu par la combustion d'un carburant dans les gaz comprimés au sein de la chambre. Les gaz brûlés, chauds et à haute pression peuvent ensuite être détendus dans la turbine. Ils fournissent alors suffisamment d'énergie au compresseur pour alimenter le moteur et à l'aéronef pour se déplacer (voir figure 1.1).

---

<sup>1</sup>Coentreprise entre General Electric et Safran Aircraft Engines

<sup>2</sup>Sustainable Aviation Fuel

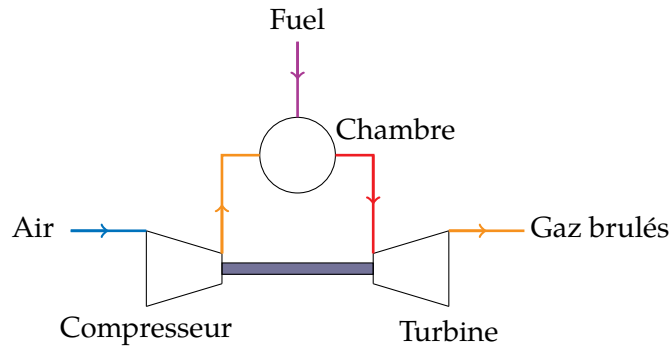


FIG. 1.1 : Schéma simplifié d'une turbomachine

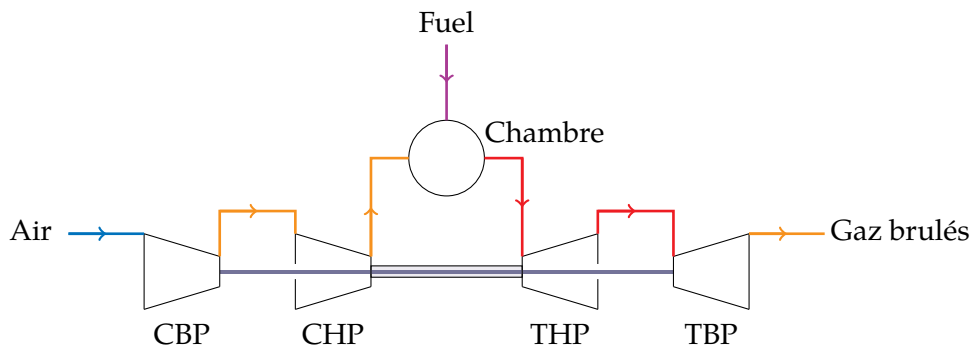


FIG. 1.2 : Schéma simplifié d'une turbomachine *double corps*

En pratique, les moteurs récents disposent de plusieurs étages de compresseurs et de turbines, généralement découplés afin d'augmenter le rendement global du moteur. Dans cette configuration de type *double corps*, on distingue généralement alors le compresseur et la turbine haute pression (respectivement *CHP* et *THP*) des éléments basse pression (*CBP* et *TBP*). Ces ensembles sont reliés par un arbre de transmission propre permettant des vitesses de rotation différentes gérées par une boîte de vitesse (voir figure 1.2).

Pour finir, les moteurs actuels, qu'ils soient commerciaux ou militaires, sont généralement à *double flux*. En entrée du compresseur *BP*, le flux d'air est divisé en deux. Une partie est dirigée vers le compresseur *HP* (flux primaire) tandis que l'autre partie contournera le reste du moteur par la veine secondaire concentrique au corps principal (flux secondaire). Ce dernier est ainsi uniquement comprimé par la *soufflante* (ou *fan*) puis éjecté. Le rapport entre la quantité d'air envoyé dans la veine secondaire et la veine principale est appelé taux de dilution ou *bypass ratio* [3]. On distingue dans cette catégorie :

- Les moteurs à fort taux de dilution qui présentent généralement des rendements élevés, une pollution sonore moindre et une consommation de carburant plus faible.
- Les moteurs à faible taux de dilution qui privilégient la compacité du groupe motopropulseur et de fortes poussées.

Globalement, les moteurs civils font partie de la première catégorie tandis que nous retrouvons la majorité des moteurs militaires dans la seconde. Certains d'entre eux comme le moteur M-88 du Rafale disposent d'un système de post-combustion consistant à injecter directement du carburant dans la tuyère d'éjection. Bien que très énergivore, ce procédé permet d'augmenter brièvement mais très fortement la poussée du moteur au décollage ou en cas de manoeuvres spécifiques.

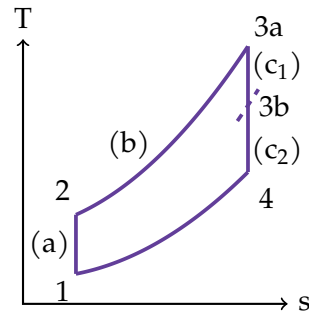


FIG. 1.3 : Cycle idéal de Brayton d'une turbomachine simplifiée

Comme démontré en détail dans l'annexe A, le rendement thermique d'une turbine à gaz idéale, en fonction des températures du cycle s'exprime selon l'équation suivante :

$$\eta_{th} = 1 - \frac{T_4}{T_3} = 1 - \frac{T_1}{T_2} \quad (1.1)$$

Les températures  $T_1$ ,  $T_2$ ,  $T_3$  et  $T_4$  désignent respectivement la température d'entrée de l'air dans le moteur, la température en sortie de compresseur, la température en entrée de turbine et la température d'éjection des gaz brûlés (voir figure 1.3). L'ensemble du développement conduisant à cette relation *rendement-températures* est construit sur des hypothèses fortes (isentropie des transformations, valeur de l'indice adiabatique  $\gamma$  constante, composition du gaz parfait inchangée). Néanmoins, il permet de mettre en exergue les leviers possibles pour maximiser le rendement thermique d'une turbomachine et par conséquent, de réduire la consommation de carburant. L'équation (1.1) montre que :

- L'augmentation de la température  $T_3$  et donc de la température de flamme dans la chambre de combustion pilote fortement le rendement de cycle de la turbomachine.
- Il est favorable de minimiser la température d'air  $T_1$  en entrée du compresseur et de maximiser celle en entrée de chambre  $T_2$  avant la combustion. De manière analogue, il est intéressant de diminuer la température des gaz brûlés en sortie de tuyère  $T_4$ .

Nous présentons dans le tableau 1.2 un exemple d'évolution des caractéristiques d'une turbomachine terrestre de type Westinghouse 501 [3]. Les tendances constatées sont en accord avec les conclusions tirées du précédent développement. Sur des systèmes terrestres de génération de puissance, il est possible d'influer sur l'ensemble des températures et sur le taux de compression. L'encombrement et le poids de la turbomachine n'étant pas un problème, les systèmes récents sont généralement équipés d'échangeurs thermiques permettant par exemple de collecter les calories des gaz en sortie de turbine (diminution de  $T_4$ ) pour les réinjecter dans le gaz en sortie de compresseur (augmentation de  $T_2$ ). La revalorisation de la chaleur fatale permet ainsi de maximiser le rendement thermique [3] du système dans son ensemble.

En pratique et dans le domaine aéronautique, il est cependant très difficile d'influer sur les températures  $T_1$ ,  $T_2$  et  $T_4$ . Notons malgré tout qu'en régime de croisière, la température d'air  $T_1$  est généralement très basse, ce qui explique en partie l'intérêt de voler à des altitudes élevées. Les contraintes physiques d'encombrement et de poids spécifiques à ce domaine rendent difficile l'ajout d'échangeurs performants. De fait, les seuls paramètres significatifs sur lesquels il est possible de jouer sont la température de combustion et le taux de compression.

Année	1968	1971	1973	1975	1981	1993
Puissance (MW)	42	60	80	95	107	160
Taux de compression	7.5	10.5	11.2	12.6	14.0	14.6
Température $T_3$ en entrée de turbine (K)	1153	1161	1266	1369	1406	1533
Rendement thermique $\eta_{th}$ (%)	27.1	29.4	30.5	31.2	33.2	35.6

Tab. 1.2 : Evolution des caractéristiques de la turbine Westinghouse 501 - Extrait de SARAVANAMUTTOO et al. [3]

*i*

Dans un système réel, le rendement d'une turbomachine dépend évidemment d'une multitude d'autres paramètres : diminution des pertes de charges, efficacité des compresseurs, architectures des veines primaire et secondaire, optimisation du *bypass* ratio, etc. Le travail de conception et les leviers d'optimisation sont donc considérables. Néanmoins, cette approche simplifiée permet de poser le contexte dans lequel s'inscrivent nos travaux.

L'optimisation du taux de compression est un problème spécifique à la conception mécanique des compresseurs et des turbines. Ici, nous nous intéressons plus en détail aux phénomènes propres à la chambre de combustion. La température visée est telle qu'elle approche aujourd'hui la température de flamme adiabatique du mélange air-kérosène. Des problématiques de plus en plus fortes se posent ainsi sur la conception des chambres :

- Résistance thermique des matériaux,
- Durée de vie de la chambre,
- Production accrue de  $NO_x$  thermiques,
- ...

Pour répondre à ces questions de conception majeures, la modélisation numérique joue un rôle fondamental. Elle permet à moindre coût de tester de nombreuses configurations et de converger vers une solution sans besoin impérieux de prototypage. Cette étape reste évidemment nécessaire mais tout le but du travail de modélisation vise à minimiser le temps et l'argent consacrés à ce procédé couteux et complexe.

L'étude et la modélisation précise des échanges thermiques au sein de la chambre deviennent de plus en plus indispensables à mesure que les températures de combustion augmentent. Plus particulièrement, les difficultés inhérentes à la modélisation du rayonnement limitent aujourd'hui sensiblement la prise en compte de l'ensemble des phénomènes physico-chimiques. Ces travaux ont ainsi pour objectif de proposer une nouvelle méthode de modélisation efficace et précise de ce mode de transfert particulier, dans le respect des coûts et des ressources numériques disponibles.

## 1.2 Grandeurs radiatives fondamentales

En guise d'avant-propos, cette première section introductive vise à présenter succinctement les grandeurs fondamentales du transfert radiatif. L'ensemble des concepts présentés ici est expliqué de manière exhaustive dans la littérature de référence [6, 5]. Ce paragraphe regroupe les notions utilisées dans la suite du manuscrit.

**Émittance, corps noir et loi de Stefan-Boltzmann** Nous définissons par émittance la densité de flux rayonné par un matériau. Cette grandeur est homogène à une puissance par élément de surface ( $W.m^{-2}$ ). Nous noterons cette grandeur  $M$ . L'émittance est fonction de la température.

Le corps noir est une notion importante en rayonnement car il représente une référence de par ses propriétés idéales. Nous définissons en effet le corps noir comme un objet dont la surface est capable d'absorber la totalité des rayons incidents, quels que soient leur longueur d'onde et leur angle d'incidence. Les rayons ne sont donc ni réfléchis ni transmis, le corps noir est un absorbeur parfait. La forme intégrée de l'émittance de ce matériau idéal est bien connue et s'exprime selon la loi de Stefan-Boltzmann :

$$M^0 = n^2 \sigma T^4 \quad (1.2)$$

L'émittance du corps noir  $M^0$  est donc proportionnelle à la température à la puissance quatrième. La grandeur  $n$  correspond à l'indice optique du milieu traversé. La constante  $\sigma$  est la constante de Stefan-Boltzmann, égale à :

$$\sigma = \frac{C_1 \pi^4}{15 C_2^4} = \frac{2 \pi^5 k_B^4}{15 c_0^2 h^3} = 5,6706 \cdot 10^8 W.m^{-2}.K^{-4} \quad (1.3)$$

Avec  $k_B = 1,380649 \cdot 10^{-23} J.K^{-1}$  la constante de Boltzmann,  $h = 6.62607015 \cdot 10^{-34} m^2.kg.s^{-1}$  la constante de Planck et  $c_0$ , la célérité de la lumière dans le vide.

**Coefficient d'absorption** Nous notons  $\kappa_{abs,\eta}$  le coefficient d'absorption au nombre d'onde  $\eta$ . Cette grandeur est homogène à l'inverse d'une longueur et s'exprime en  $m^{-1}$ . Ce coefficient caractérise la capacité d'un milieu à absorber ou non des photons. Le libre parcours moyen d'absorption, c'est à dire la longueur moyenne au terme de laquelle les photons sont absorbés dans un milieu, est égale à l'inverse de ce coefficient. Ce dernier est dépendant des conditions thermophysiques au point d'intérêt. Il dépend notamment de la concentration en espèces gazeuses rayonnantes<sup>3</sup>, de la température, de la pression et de la fraction volumique de suie. Il peut être calculé ou modélisé selon les conditions du cas d'étude de façon plus ou moins exhaustive (voir section 1.3.3 et section 1.4.4).

**Luminance spectrale et intégrée : nature directionnelle du rayonnement** Bien que l'émittance soit une grandeur fondamentale dans la science du rayonnement, elle ne prend pas en compte sa nature directionnelle. C'est une valeur macroscopique intégrée sur l'ensemble des directions. Nous pouvons alors définir une grandeur  $L_\eta$  appelée luminance monochromatique. Elle quantifie le flux rayonné par unité de surface  $\cos(\theta)dS$  par unité d'angle solide  $d\Omega$  centré sur la direction du rayon  $\mathbf{u}$  et par unité de nombre d'onde  $\eta$ . Elle s'exprime en  $W.m^{-3}.sr^{-1}.cm$  telle que :

$$L_\eta = \frac{d\dot{Q}_\eta(\Omega)}{d\eta \cos(\theta) dS d\Omega} \quad (1.4)$$

La quantité  $d\dot{Q}_\eta$  représente la puissance thermique rayonnée par la surface  $dS$  par unité de nombre d'onde  $d\eta$  dans l'angle solide  $d\Omega$ . L'angle formé entre la normale sortante  $\hat{n}$  à la surface et la direction considérée est noté  $\theta$ . Nous avons donc  $\cos(\theta) = \mathbf{u} \cdot \hat{n}$ . La luminance peut également s'exprimer de façon équivalente en fonction d'autres variables spectrales comme la fréquence ou la longueur d'onde.

<sup>3</sup>Pour les applications de type combustion, nous retenons généralement  $H_2O$ ,  $CO_2$  et  $CO$

La plupart des modèles de gaz étant exprimés en nombre d'onde, nous faisons le choix d'exprimer exclusivement la luminance en fonction de  $\eta$  dans la suite du manuscrit. On distingue alors la luminance spectrale  $L_\eta$  de la luminance totale  $L$ , cette dernière étant déterminée par l'intégration sur le spectre de ses composantes monochromatiques :

$$L(\mathbf{x}, \mathbf{u}) = \int_0^\infty L_\eta(\mathbf{x}, \mathbf{u}, \eta) d\eta \quad (1.5)$$

On note respectivement  $\mathbf{x}$  et  $\mathbf{u}$ , les coordonnées spatiales et la direction considérée. Nous utiliserons une intégrale tronquée aux bornes  $\eta_1$  et  $\eta_2$  de l'intervalle de nombres d'onde pris en compte (tel que  $\eta \in [150, 9300] \text{ cm}^{-1}$ ) :

$$L(\mathbf{x}, \mathbf{u}) \approx \int_{\eta_1}^{\eta_2} L_\eta(\mathbf{x}, \mathbf{u}, \eta) d\eta \quad (1.6)$$

**Luminance du corps noir** Comme décrit précédemment, le corps noir est un matériau de référence en physique du rayonnement. Il sert de référence pour calculer le terme d'émission dans l'équation de transfert radiatif (voir section 1.3.1). La luminance monochromatique de ce corps n'est fonction que de la température  $T$  et peut être calculée selon la loi de Planck [5] :

$$L_\eta^0(T) = \frac{2hc_0^2\eta^3}{n^2 \left[ \exp\left(\frac{\eta hc_0}{nk_B T}\right) - 1 \right]} \quad (1.7)$$

### 1.3 Rayonnement thermique en chambre de combustion : état de l'art

Comme nous l'avons vu dans la section 1.1.2 un haut rendement est atteint pour de hautes températures en entrée de turbine. Ce constat implique des contraintes de plus en plus élevées sur les matériaux constituant les organes essentiels du moteur. La prise en compte du transfert radiatif revêt alors une importance toute particulière car il influe significativement sur les températures de parois de la chambre de combustion.

Historiquement, la modélisation de ce type de transfert en milieu participatif est traitée au sein du groupe Safran par une méthode de type Ordonnées Discrètes (voir section 1.3.2). Sa formulation est adaptée au calcul des grandeurs radiatives (luminance, flux, divergence de flux, etc) sur maillage. Mathématiquement, sa formulation est très proche des méthodes traditionnellement utilisées en mécanique des fluides (Computational Fluid Dynamics CFD) de type Volumes Finis avec des composantes spectrales et angulaires supplémentaires afin de prendre en compte les spécificités des transferts par rayonnement.

Dans cette section, nous proposons de faire un point de situation sur les pratiques de modélisation courantes dans l'industrie aéronautique. Cet état de l'art débute avant tout par l'énoncé de l'équation régissant le rayonnement dans sa formulation différentielle. La méthode de résolution par Ordonnées Discrètes est ensuite présentée pour en comprendre le principe. Enfin, nous présentons les modèles de gaz actuellement utilisés pour le calcul du coefficient d'absorption.

Ce tour d'horizon permettra enfin de dresser les limites que posent les méthodologies actuelles sur la prise en compte précise et rapide du rayonnement dans ce type de système complexe. L'objectif est de démontrer la nécessité de disposer d'une nouvelle solution conceptuelle apportée par l'ensemble de ces travaux de thèse.

### 1.3.1 Formulation différentielle de l'équation de transfert radiatif

L'équation de transfert radiatif (*ETR*) dérive d'une équation de Boltzmann qui régit le transport des photons dans un milieu participant. Les hypothèses suivantes sont formulées :

- Sans interaction avec le milieu, les photons se déplacent en ligne droite et à vitesse constante. Il est donc possible de décrire la distribution des photons par leur position et leur direction,
- Nous considérons que l'équilibre thermodynamique local est atteint en chaque point du milieu,
- Les photons n'interagissent par les uns avec les autres et ne sont soumis à aucune force extérieure.

Sous ces conditions, l'expression générale d'une équation de Boltzmann est adaptée à l'étude du rayonnement et donc à la formulation de l'ETR. Une explication plus complète de cette étape est par exemple proposée dans les travaux de thèse de ROGER [11] et dans l'ouvrage de SIEGEL et HOWELL [5].

**En milieu transparent** Dans le cas du rayonnement thermique, nous utilisons la luminance spectrale comme descripteur mésoscopique de la distribution des photons sur leur espace des phases. En supposant pour le moment que le milieu est transparent, nous pouvons écrire une équation de conservation de luminance [6, 11, 7] :

$$\frac{1}{c} \partial_t L_\eta(\mathbf{x}, \mathbf{u}, t) + \mathbf{u} \cdot \nabla_{\mathbf{x}} L_\eta(\mathbf{x}, \mathbf{u}, t) = 0 \quad (1.8)$$

À partir de ce point, une nouvelle approximation peut-être faite sur la dépendance temporelle de la luminance. Dans les applications de combustion, les transferts par rayonnement sont suffisamment rapides par rapport aux autres phénomènes (réactions chimiques et transport du fluide notamment) pour considérer les transferts radiatifs comme stationnaires. L'équation (1.8) peut alors être simplifiée :

$$\mathbf{u} \cdot \nabla_{\mathbf{x}} L_\eta(\mathbf{x}, \mathbf{u}, t) = 0 \quad (1.9)$$

**En milieu participant** Les gaz de combustion sont composés d'une multitude d'espèces chimiques différentes dont certaines ne sont pas neutres vis à vis du rayonnement. Les molécules d' $H_2O$ , de  $CO_2$ , de  $CO$  et les particules de suies très présentes dans les gaz brûlés, interagissent fortement avec ce type de transfert thermique. En toute rigueur, nous pouvons faire l'inventaire de trois interactions différentes entre le milieu d'évolution et les photons : l'absorption, l'émission et la diffusion. Dans le cadre de l'étude de chambres de combustion, la prise en compte des phénomènes de diffusion n'a que peu d'influence sur les températures finales calculées aux parois. De ce fait, la prise en compte des phénomènes radiatifs s'accompagne le plus souvent d'une approximation consistant à négliger la diffusion. Ainsi, en milieu purement absorbant, il est nécessaire de prendre en compte un terme source d'émission de photons par le milieu et en un terme puits d'absorption de photons par les espèces participantes.

**Écriture du terme puits d'absorption** En milieu semi-transparent, les espèces participantes peuvent absorber spontanément les photons traversant le milieu, augmentant par conséquent son énergie interne. Il devient alors nécessaire de prendre en considération la *disparition* de photons dans le bilan de flux calculé à travers ce milieu. De manière analogue, ces espèces peuvent émettre des photons.

Ce phénomène d'absorption est uniquement dépendant du milieu, caractérisé par son coefficient d'absorption local (voir section 1.2), lui même fonction des conditions thermophysiques du champ.

L'absorption du rayonnement est également très dépendante du nombre d'onde car les espèces comme le  $CO_2$  notamment, présentent dans leur spectre d'absorption des zones transparentes distribuées de manière très hétérogène en fonction du nombre d'onde. L'écriture spectrale de ce terme puits d'absorption peut alors s'écrire simplement :

$$S_{abs} = -\kappa_{abs,\eta} L_\eta \quad (1.10)$$

**Écriture du terme source d'émission** En considérant que l'équilibre thermodynamique local est atteint. Le terme source d'émission est égal à la luminance du corps noir  $L_\eta^0$ , pondéré par un coefficient d'émission. Ce dernier est strictement égal au coefficient d'absorption selon la loi de Kirchoff. Le terme source relatif à l'émission de photons par le milieu s'écrit donc ainsi :

$$S_e = +\kappa_{abs,\eta} L_\eta^0(T) \quad (1.11)$$

**Écriture de l'ETR sous forme différentielle en milieu purement absorbant** En prenant en compte les différentes interactions possibles avec le milieu, l'équation (1.9) peut être complétée. L'équation de transfert radiatif en milieu purement absorbant, s'écrit sous forme différentielle telle que :

$$\underbrace{\mathbf{u} \cdot \nabla_{\mathbf{x}} L_\eta(\mathbf{x}, \mathbf{u}, t)}_{\text{Transport}} = \underbrace{\kappa_{abs,\eta}(\mathbf{x}) L_\eta^0(\mathbf{x}, \mathbf{u})}_{\text{Terme source}} - \underbrace{\kappa_{abs,\eta}(\mathbf{x}) L_\eta(\mathbf{x}, \mathbf{u})}_{\text{Terme puits}} \quad (1.12)$$

### 1.3.2 Résolution numérique de l'ETR sous forme différentielle par la méthode des Ordonnées Discrètes

Dans sa forme différentielle, l'équation de transfert radiatif peut être résolue par des méthodes similaires à celles utilisées pour traiter numériquement les équations de transport de Navier-Stokes par exemple. Dans le cas du rayonnement, les caractères directionnel et spectral doivent néanmoins être pris en compte. La méthode par *Ordonnées Discrètes* (ou *DOM* pour *Discrete Ordinate Method*) est ainsi comparable à la méthode des volumes finis. A ce titre, tout comme il est nécessaire de discrétiser l'espace en mailles, nous devons également le discrétiser en directions et en intervalles spectraux. Un bilan de luminance spectrale est ensuite calculée de proche en proche à travers chaque volume de contrôle pour chaque direction et chaque nombre d'onde (voir figure 1.4).

**Avantages et limites** La *DOM* est une méthode efficace et adaptée à la résolution de l'ETR sur maillage. La similarité avec les méthodes par volumes finis la rend particulièrement pertinente pour être couplée avec des solveurs CFD dans le cadre de simulations multi-physiques, que le maillage soit structuré ou non. C'est donc une méthode en champ dont la formulation en coefficient d'absorption permet l'utilisation de la plupart des modèles de gaz, la prise en compte de parois non noires, etc. Historiquement, cette méthode est donc très utilisée pour sa polyvalence et son efficacité.

Néanmoins, cette méthode trouve ses limites lorsque les maillages deviennent imposants et que la résolution spectrale recherchée est importante. En effet, en ordonnées discrètes, il est nécessaire de stocker en mémoire l'ensemble des différents champs spectraux avant de pouvoir calculer des grandeurs intégrées comme la divergence du flux radiatif.



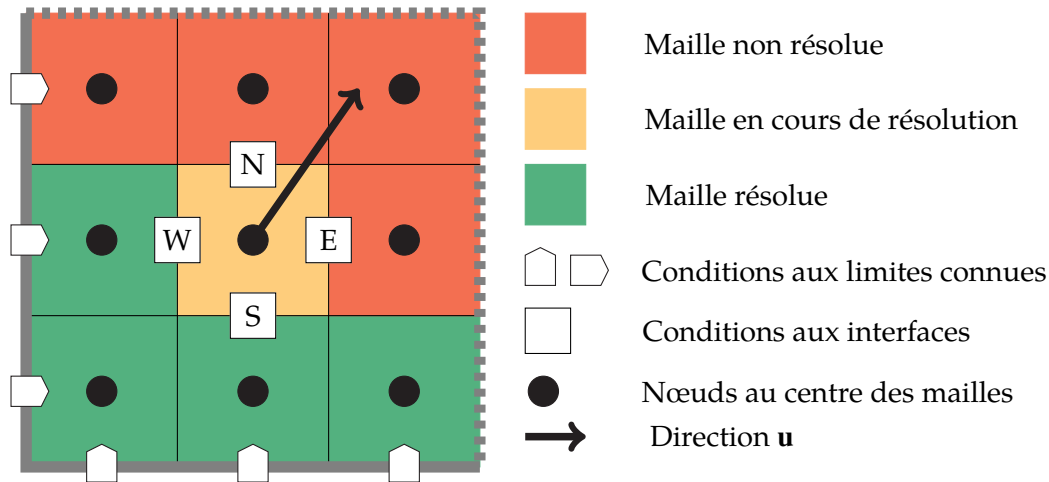


FIG. 1.4 : Résolution de proche en proche de la luminance à travers un volume de contrôle selon la direction  $u$

#### Calcul d'ordre de grandeur : taille de stockage nécessaire

L'ordre de grandeur classique du nombre de mailles pour une chambre de combustion moderne se situe autour de 250 millions. En dehors des modèles raie par raie, ceux faisant référence comportent en général 367 bandes (voir section 1.4.4). Par conséquent, nous pouvons estimer le taille de stockage nécessaire à  $250 \cdot 10^6 \times 367 \times 64 \approx 6Tbits$  soit environ 734Go rien que pour le rayonnement et pour seulement une seule direction discrète<sup>a</sup>. Autrement dit, l'utilisation de tels modèles sur des cas aussi volumineux n'est pas accessible et *a fortiori* dans le cadre de simulations couplées. Notons que cette capacité doit être disponible dans la RAM ce qui techniquement, pose également de très grosses contraintes sur les moyens de calcul utilisés. Le temps de calcul devient également problématique mais la capacité de stockage est réellement le facteur bloquant car il dépend directement des limites techniques du matériel utilisé.

<sup>a</sup>On suppose que chaque variable est encodée sur 64 bits

### 1.3.3 Les modèles globaux de rayonnement des gaz

Malgré la très vaste littérature concernant les modèles spectraux de gaz (voir section 1.4.4), leur utilisation sur cas complexes et couplés est encore assez marginale du fait du peu de ressources de calcul disponibles pour la modélisation du rayonnement thermique. En effet, et *a fortiori* dans l'industrie, les simulations menées dans les chambres de combustion modélisent beaucoup de phénomènes liés les uns aux autres, mobilisant tous beaucoup de ressources de calcul. L'impasse a ainsi souvent été faite sur l'utilisation de modèles spectraux à haute résolution en leur préférant des modèles plus simples et plus rapides. Les modèles globaux répondent à ce cahier des charges.

**Approche par somme de gaz gris : modèle WSGG** Le moyen le plus simple de modéliser le rayonnement est de considérer le gaz comme gris, ce qui implique de moyenniser l'ensemble du spectre pour déterminer un coefficient d'absorption moyen global. C'est évidemment une très forte approximation qui n'a finalement que peu d'intérêt tant les erreurs peuvent être importantes.

Une version plus avancée et encore largement utilisée est l'approche par somme pondérée de gaz gris (*Weighted Sum of Gray Gases* ou *WSGG*) introduite par HOTTEL [4]. A l'origine, ce modèle était utilisé pour calculer l'émissivité totale  $\epsilon$  d'une couche de gaz absorbante d'épaisseur  $l$  :

$$\epsilon = \sum_{j=1}^N a_j (1 - e^{-\kappa_j l}) \quad (1.13)$$

Ici,  $\kappa_j$  sont les coefficients d'absorption des gaz gris et  $a_j$  les poids qui leur sont associés. Lors de la proposition de ce modèle, les bases spectroscopiques à haute résolution n'existaient pas et ces paramètres ont été définis par mesure expérimentale. Plus tard, des corrélations ont été proposées pour déterminer des valeurs plus pertinentes pour les paramètres du modèle en fonction des conditions thermo-physiques [12, 13].

Dans un milieu non diffusant, MODEST [14] a montré que les paramètres du modèle *WSGG* pouvaient être utilisés en calculant la contribution de chaque gaz gris à la luminance totale :

$$\frac{\partial L_j}{\partial s} = \kappa_j [a_j L^0 - L_j] \quad (1.14)$$

Toutes les luminances de gaz gris  $L_j$  peuvent ensuite être simplement sommées pour obtenir une luminance totale :

$$L = \sum_{j=0}^N L_j \quad (1.15)$$

Cette version appelée *WSGG RTE* a l'avantage d'être formulée par le biais de l'ETR et donc de pouvoir être traitée par n'importe quelle méthode de résolution classique.

**Les modèles SLW** Cette famille de modèles globaux s'est rapidement imposée comme une alternative efficace à la méthode *WSGG*. Le modèle *SLW* (*Spectral Line Weighted Sum of gray Gases*) a été développé à l'origine par DENISON [15] dans ses travaux de thèse puis largement complété par une série de publications dans la littérature [16, 17, 18, 19]. Ici, contrairement au modèle précédent, on se base sur la définition physique des paramètres en s'appuyant sur des bases de données spectroscopiques à haute résolution plutôt que sur des résultats expérimentaux sur l'émissivité totale du gaz.

Dans le modèle *SLW*, on utilise une fonction de distribution appelée *ALBDF* (*Absorption Line Blackbody Distribution Function*) construite à partir de spectres haute définition qui nous permet de calculer des poids pertinents pour chaque gaz gris considéré. Cette fonction peut être calculée en amont pour chaque espèce selon des corrélations spécifiques aux conditions de pression et de température [20]. La fonction *ALBDF* est construite de la manière suivante :

- En premier lieu, le spectre d'absorption est discrétisé en intervalles de *sections efficaces* sur lesquelles la valeur du coefficient d'absorption est moyennée. Chacune de ces portions de spectre est alors considérée comme un gaz gris.
- Chacun de ces gaz gris est ensuite pondéré par un poids défini par l'épaisseur de chaque intervalle discret.
- Ces sections sont ensuite réordonnées en un histogramme croissant et approchées par une fonction de distribution appelée *ALBDF*.

Historiquement, ce modèle nécessitait la définition d'un état de référence. On supposait alors qu'il existait une relation entre les spectres d'un mélange à différents états. Cette nécessité fut longtemps une des limites de ce type de modèle malgré sa très bonne efficacité. Le choix des conditions de référence et notamment de la température, avait en effet une influence sensible sur les résultats du modèle. Cette limite a récemment été dépassée dans les travaux de ANDRÉ, SOLOVJOV et WEBB [21]. En fonction du nombre de gaz gris utilisé, les erreurs relatives à ce type de modèle peuvent aujourd'hui approcher celles de modèles spectraux à haute résolution à quelques pourcents près [22].

## 1.4 Approche stochastique par méthode de Monte Carlo et modèles spectraux de gaz

Nous proposons dorénavant de présenter l'ETR dans sa formulation intégrale et sa résolution par la méthode de Monte Carlo. Cette méthode de calcul intégral est couramment présentée comme la méthode de référence en transfert radiatif [6, 5].

Contrairement à la *DOM*, la méthode de Monte Carlo ne souffre d'aucun biais ou d'approximation dans sa formulation mathématique ou sa mise en œuvre numérique. Elle est également insensible à la complexité géométrique et indépendante du modèle de gaz utilisé. Malgré ces qualités, cette méthode est plus adaptée au calcul de point sonde : il est possible d'estimer des grandeurs de manière très précise (avec une estimation de l'erreur de mesure associée) en un point précis (ou plusieurs). Elle n'est donc pas adaptée pour des calculs de champs complets à l'inverse des méthodes par Ordonnées Discrètes et de la plupart des méthodes déterministes. Ces travaux de thèse visent à palier cette limitation en proposant d'étendre l'utilisation de la méthode de Monte Carlo au calcul de champs par couplage avec des réseaux de neurones (voir chapitre 2). Le réseau de neurones fait alors office d'interpolateur et permet d'estimer rapidement un champ d'une grandeur radiative à partir d'une base de données générée par méthode de Monte Carlo.

### 1.4.1 Formulation intégrale de l'Équation de Transfert Radiatif

Afin de résoudre l'ETR par la méthode de Monte Carlo nous l'exprimons sous sa forme intégrale. La luminance spectrale  $L_\eta$  est alors présentée comme une espérance de la luminance du corps noir  $L_\eta^0$ . Cette reformulation permet ensuite de résoudre cette équation grâce à un algorithme de Monte Carlo. Cette méthode a l'avantage de fournir des informations sur la variance des résultats et donc par extension, sur la précision et la confiance que l'on peut leur accorder. L'obtention de l'erreur statistique associée à la quantité d'intérêt permet de contrôler la qualité des résultats. La formulation théorique, mathématiquement rigoureuse, de la méthode de Monte Carlo autant que sa mise en œuvre numérique permettent de modéliser le rayonnement thermique sans aucun biais ou approximation, à la différence des méthodes *DOM* où des discrétisations de l'espace, des directions et du modèle à résoudre sont nécessaires.

**Mise en équation** Nous notons tout d'abord  $\tau_\eta(\mathbf{x}, \mathbf{u}, l)$ , la transmittivité spectrale. Elle correspond à la probabilité pour un photon à la position  $\mathbf{x}$  de parcourir un chemin  $l$  dans la direction  $\mathbf{u}$  dans le milieu d'étude sans être absorbé. Cette probabilité est indépendante de la distance déjà parcourue dans un milieu non diffusant. On peut alors définir la transmittivité selon la loi de Beer-Lambert [6, 9] comme une fonction de l'épaisseur optique telle que :

$$\tau_\eta(\mathbf{x}, \mathbf{u}, l) = \exp\left(-\underbrace{\int_0^l \kappa_{abs,\eta}(\mathbf{x} + l'\mathbf{u}) dl'}_{\text{épaisseur optique}}\right) \quad (1.16)$$

Nous pouvons alors exprimer la luminance  $L_\eta$  à une position  $\mathbf{x}_0$  dans la direction  $\mathbf{u}_0$  comme l'intégrale spatiale sur le chemin  $] -\infty, \mathbf{x}_0]$  d'un terme d'émission atténuée selon la loi de Beer-Lambert :

$$L_\eta(\mathbf{x}_0, \mathbf{u}_0) = \int_{-\infty}^{\mathbf{x}_0} \kappa_{abs,\eta}(\mathbf{x}) d\mathbf{x} \exp\left(-\int_{\mathbf{x}}^{\mathbf{x}_0} \kappa_{abs,\eta}(\mathbf{x}') d\mathbf{x}'\right) L_\eta^0(\mathbf{x}) \quad (1.17)$$

L'équation (1.17) peut être reformulée en introduisant la fonction de densité de probabilité caractéristique de la loi de Beer-Lambert afin d'obtenir une formulation statistique de la luminance :

$$L_\eta(\mathbf{x}_0, \mathbf{u}_0) = \int_{-\infty}^{\mathbf{x}_0} p_X(\mathbf{x}) d\mathbf{x} L_\eta^0(\mathbf{x}) \quad (1.18)$$

Avec  $p_X(\mathbf{x})$ , la fonction de densité de probabilité :

$$p_X(\mathbf{x}) = \kappa_{abs,\eta}(\mathbf{x}) \exp\left(-\int_{\mathbf{x}}^{\mathbf{x}_0} \kappa_{abs,\eta}(\mathbf{x}') d\mathbf{x}'\right) \quad (1.19)$$

L'équation (1.19) est une fonction de densité de probabilité qui définit la probabilité qu'un photon soit absorbé par le milieu. Cette probabilité augmente avec la distance entre  $\mathbf{x}_0$  et  $\mathbf{x}$  et la valeur du coefficient d'absorption  $\kappa_{abs,\eta}$ . Dans cette formulation, la luminance spectrale est bien écrite comme une espérance et peut être calculée par un algorithme de Monte Carlo :

$$L_\eta(\mathbf{x}_0, \mathbf{u}_0) = \mathbb{E}[L_\eta^0(\mathbf{x})] = \lim_{N_{mc} \rightarrow \infty} \left( \frac{1}{N_{mc}} \sum_{i=1}^{N_{mc}} L_\eta^0(\mathbf{x}) \right) \quad (1.20)$$

En pratique, le nombre de réalisations de Monte Carlo  $N_{mc}$  est fini et *suffisamment grand*. Comme mentionné précédemment, le contrôle de la variance et donc de l'erreur statistique faite sur la grandeur calculée est probablement une des plus grandes forces de la méthode de Monte Carlo. Si la variance est trop élevée, il est toujours possible de la réduire en augmentant la valeur de  $N_{mc}$  ou en utilisant d'autres techniques de réduction de variance [23, 24, 25, 26]. Enfin, les travaux de EL HAFI et al. [8] et IBARRART [9] par exemple démontrent l'insensibilité des algorithmes de Monte Carlo à la géométrie. Ces méthodes sont donc pertinentes pour traiter de cas réels et détaillés, notamment dans le cas de chambre de combustion aéronautique.

## 1.4.2 Résolution de l'ETR sous forme intégrale par Méthode de Monte Carlo en milieu hétérogène : introduction des collisions nulles

Dans un milieu homogène, le calcul du libre parcours moyen des photons  $l_i$  est une solution analytique de l'équation (1.21), où  $r_i$  est un nombre aléatoire compris entre 0 et 1. Ce libre parcours moyen dépend uniquement d'un coefficient d'absorption par définition constant.

$$l_i = -\frac{\ln(1 - r_i)}{\kappa_{abs,\eta}} \quad (1.21)$$

En milieu hétérogène, il n'est en revanche plus possible de définir un libre parcours moyen ni d'intégrer analytiquement l'épaisseur optique du milieu [27] comme proposé dans l'équation (1.16). En termes de résolution, cette caractéristique rend les choses plus complexes d'un point de vue numérique. Il existe des techniques pour surmonter cette difficulté, qui consistent généralement soit à discrétiser l'espace, soit à inverser l'épaisseur optique plutôt que le libre parcours. Dans les deux cas et malgré un coût de calcul très élevé, la formulation sans approximation de la méthode de Monte Carlo est perdue car un biais numérique difficile à quantifier apparaît. Afin de surmonter cette limite forte, l'algorithme basé sur la formulation de l'équation de transfert radiatif en collision nulle proposé par GALTIER [25] est utilisé. Elle est mathématiquement exacte et ne nécessite l'introduction d'aucun biais dans sa formulation ou son utilisation numérique.

Historiquement, cette formulation est apparue à la fin des années 60 dans le domaine de la physique des plasmas [28] et de la neutronique [29] et est toujours très utilisée aujourd'hui dans beaucoup de codes de calculs spécialisés [30]. Cette formulation a ensuite été étendue au rayonnement thermique dans la thèse de GALTIER [25].

Le principe de la formulation en collision nulle est de transformer un champ hétérogène en un champ homogène. Comme présenté en équation (1.17), la longueur du parcours moyen d'un photon est fonction du champ de coefficients d'absorption  $\kappa_{abs,\eta}$ . Les différentes informations concernant l'hétérogénéité du domaine sont donc incluses dans cette grandeur (voir section 1.4.4). L'idée est alors de construire artificiellement un champ théorique majorant et homogène, où le coefficient d'absorption est égal en tout point à  $\widehat{\kappa}_\eta$ , défini tel que :

$$\widehat{\kappa}_\eta = \kappa_{abs,\eta} + \kappa_{n,\eta} \quad (1.22)$$

Le coefficient de collision nulle  $\kappa_{n,\eta}$  est arbitraire et jamais exprimé autrement que comme étant la différence entre le majorant  $\widehat{\kappa}_\eta$  et le coefficient d'absorption *réel* au point considéré  $\kappa_{abs,\eta}$ . Il devient une variable d'ajustement pour assurer la constance de  $\widehat{\kappa}_\eta$ . La figure 1.5 donne une représentation du lien entre ces trois grandeurs le long d'un chemin de propagation  $\mathbf{x}$ .

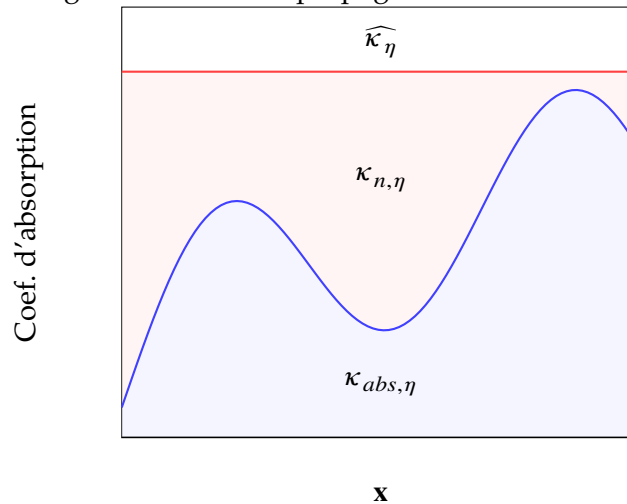


FIG. 1.5 : Homogénéisation des coefficients d'absorption par un majorant

Naturellement, l'ajout de ce nouveau coefficient ne change pas la physique du problème ni le résultat donné par l'ETR. On peut ainsi montrer que l'équation (1.18) est rigoureusement égale à [25] :

$$L_\eta(\mathbf{x}_0, \mathbf{u}_0) = \int_0^{+\infty} dl' \widehat{p}_X(\mathbf{x}') \times \left[ \frac{\kappa_{abs,\eta}(\mathbf{x}')}{\widehat{\kappa}_\eta(\mathbf{x}')} L_\eta^0(\mathbf{x}') + \frac{\kappa_{n,\eta}(\mathbf{x}')}{\widehat{\kappa}_\eta(\mathbf{x}')} \int_{4\pi} \delta(\mathbf{u}_0 - \mathbf{u}') L_\eta(\mathbf{x}', \mathbf{u}') d\mathbf{u}' \right] \quad (1.23)$$

Où  $\mathbf{x} = \mathbf{x}_0 - l$  et  $\widehat{p}_X(\mathbf{x})$  est analogue à  $p_X(\mathbf{x})$  (voir l'équation (1.19)) :

$$\widehat{p}_X(\mathbf{x}) = \widehat{\kappa}_\eta(\mathbf{x}) \exp\left(-\int_0^l \widehat{\kappa}_\eta(\mathbf{x}_0 - l' \mathbf{u}_0) dl'\right) \quad (1.24)$$

Cette formulation repose sur une image physique assez simple et intuitive. Nous pouvons notamment reconnaître dans l'équation (1.24) de fortes similitudes avec la formulation intégrale de l'ETR en milieu absorbant et diffusant [31, 32] au détail près que la fonction de phase utilisée n'est pas une fonction isotrope ou de Green [33] mais une fonction Dirac ne laissant que la possibilité d'une *pseudo-diffusion* sans changement de direction du photon. La collision nulle peut ainsi être considérée de deux manières : comme un *non-événement* ou comme une *diffusion vers l'avant*.

Le champ parcouru par le photon étant homogène et majoré par  $\widehat{\kappa}_\eta$ , le libre parcours moyen tiré aléatoirement est en moyenne toujours *plus court* par rapport au champ réel. Ainsi, à chaque tirage d'un chemin, on mesure la probabilité pour que le photon soit absorbé ou bien pour qu'il continue sa propagation vers l'avant, dans la même direction. Dans le cas où le coefficient d'absorption réel est proche du majorant, alors le rapport  $\frac{\kappa_{abs,\eta}(\mathbf{x}')}{\widehat{\kappa}_\eta(\mathbf{x}')}$  dans l'équation (1.23) est proche de 1 et la probabilité pour que l'évènement *absorption* se produise est grande. A l'inverse, si le chemin tiré se termine dans un milieu optiquement mince par rapport au reste du champ étudié, c'est à dire où  $\kappa_{abs,\eta}$  est faible devant le majorant, alors la collision nulle devient la plus probable.

*i* Bien que le choix du majorant soit mathématiquement tout à fait arbitraire, il est néanmoins important de ne pas le surestimer sans nécessité. Le nombre de collisions nulles augmentera alors sensiblement au risque d'augmenter le temps de calcul inutilement. Récemment, VILLEFRANQUE et al. [10] ont proposé une méthode pour diminuer significativement le temps de calcul en utilisant un majorant par zone pour limiter le nombre de collisions vers l'avant. Ces grilles accélératrices permettent d'adapter le majorant en fonction de la valeur des coefficients d'absorption environnant.

L'équation (1.23) peut finalement être écrite sous une forme récursive, plus proche de sa mise en oeuvre réelle dans un algorithme numérique. A chaque tirage aléatoire de longueur optique, la probabilité  $\frac{\kappa_{abs,\eta}(\mathbf{x}')}{\widehat{\kappa}_\eta(\mathbf{x}')}$  est calculée en fonction de la valeur du coefficient d'absorption au point  $\mathbf{x}'$ . Le résultat d'un nouveau tirage aléatoire uniforme entre 0 et 1 est ensuite comparé à cette probabilité telle que :

- Si la valeur obtenue est inférieure au rapport  $\frac{\kappa_{abs,\eta}(\mathbf{x}')}{\widehat{\kappa}_\eta(\mathbf{x}')}$ , alors l'évènement sera une absorption. Le chemin s'arrête et la contribution  $L_\eta^0(\mathbf{x}')$  est enregistrée pour le calcul de la luminance  $L_\eta$  (voir équation (1.20)),
- Sinon, l'évènement est une collision nulle. Une nouvelle longueur est alors tirée selon l'équation (1.21) sans changement de direction et ainsi de suite jusqu'à ce que le photon soit absorbé.

### Un exemple de réalisation de Monte Carlo

La figure 1.6 permet d'illustrer le fonctionnement d'un algorithme à collision nulle. La grille représente un maillage d'une simulation en 2 dimensions où les conditions thermo-physiques sont connues. Il est donc possible de calculer le champ de coefficients d'absorption en chaque maille et pour chaque bande spectrale considérée par un pré-traitement du champ (voir section 1.4.4). Nous souhaitons dans cet exemple calculer la luminance monochromatique au point  $\mathbf{x}_0$ . Lors de chaque réalisation de Monte Carlo, une direction et une distance  $l_0$  sont tirées aléatoirement. Nous pouvons ainsi déterminer le *point de chute*  $\mathbf{x}_1$ . Le coefficient d'absorption  $\kappa_{abs,\eta}(\mathbf{x}_1)$  est connu d'après le champ thermophysique. Nous pouvons donc calculer la probabilité  $\frac{\kappa_{abs,\eta}(\mathbf{x}')}{\overline{\kappa}_\eta(\mathbf{x}'')}$  et réaliser un test de Bernoulli pour déterminer si le photon est absorbé ou diffusé vers l'avant. Dans cet exemple, nous constatons 3 collisions nulles avant absorption du photon au point  $\mathbf{x}_4$ . Nous calculons alors la valeur de la luminance spectrale du corps noir  $L_\eta^0(\mathbf{x}_4)$  d'après la loi de Planck et ajoutons cette contribution à la somme de Monte Carlo conformément à l'équation (1.20).

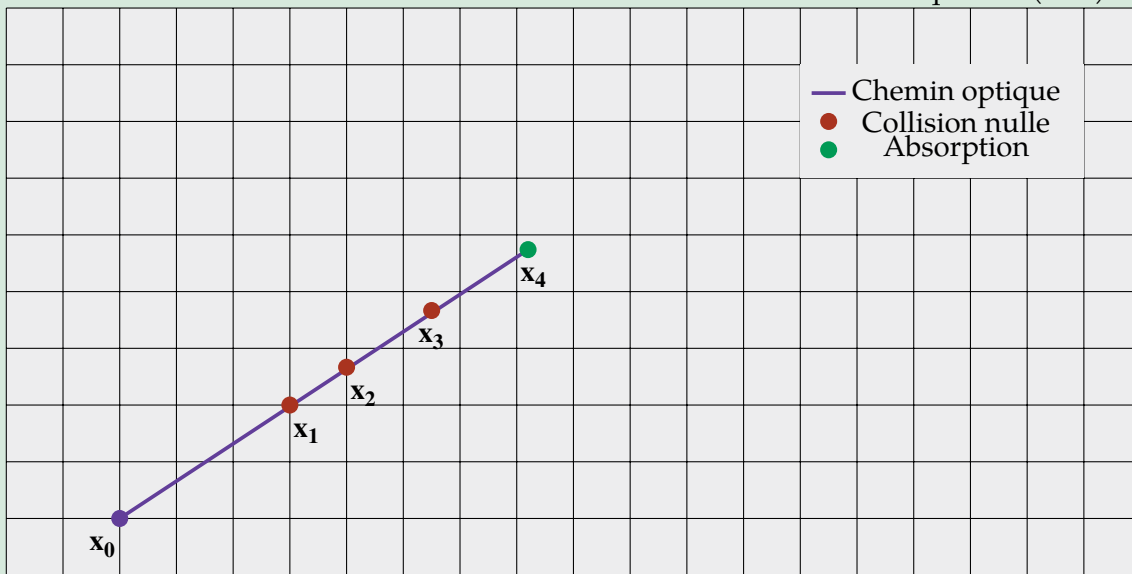


FIG. 1.6 : Exemple de chemin parcouru par un photon dans un champ majoré et uniformisé

### 1.4.3 Expression des grandeurs radiatives d'intérêt

Pour l'étude de la thermique des chambres, nous nous intéressons ici à deux observables : le flux radiatif à la paroi et la divergence du flux radiatif dans le volume. La première permet notamment de pouvoir déterminer, dans le cadre d'une analyse plus globale de la thermique, la température à la paroi. Ce flux se définit comme étant l'intégrale de l'ensemble des contributions spectrales des luminances  $L_\eta$  provenant du demi-espace *avant*  $\Omega^+$  :

$$q(\mathbf{x}) = \int_0^\infty \int_{\Omega^+} L_\eta(\mathbf{x}, \mathbf{u}) \cos(\theta) d\eta d\Omega = \int_0^\infty \int_{\Omega^+} L_\eta(\mathbf{x}, \mathbf{u}) |\mathbf{u} \cdot \hat{\mathbf{n}}| d\eta d\Omega \quad (1.25)$$

La divergence de flux  $\nabla \cdot q(\mathbf{x})$  est une grandeur volumique. Elle permet notamment de mesurer le bilan net entre émission et absorption des photons dans le milieu participant. Elle correspond exactement au terme source apparaissant dans l'équation de conservation de l'énergie d'un solveur CFD par exemple.

En considérant un domaine d'étude isotrope, nous pouvons écrire :

$$\nabla \cdot \mathbf{q}(\mathbf{x}) = \int_0^\eta d\eta \int_0^{4\pi} \kappa_{abs,\eta}(\mathbf{x}) [L_\eta^0(\mathbf{x}) - L_\eta(\mathbf{x}, \mathbf{u})] d\Omega \quad (1.26)$$

La luminance monochromatique du corps noir  $L_\eta^0$  correspond au terme d'émission au point d'intérêt. Elle est uniquement dépendante de la température et peut être calculée selon l'équation (1.7). La luminance spectrale  $L_\eta$  multipliée par  $\kappa_{abs,\eta}$  représente quant à elle le terme d'absorption au point  $\mathbf{x}$  provenant de la direction  $\mathbf{u}$  et pouvant être calculé d'après l'équation (1.17) et la connaissance des conditions thermo-physiques du champ participant. Ces dernières sont en effet nécessaires à la fois pour le calcul de la luminance mais également pour la détermination du coefficient d'absorption  $\kappa_{abs,\eta}$  (voir la section 1.4.4).

#### 1.4.4 Modèles de gaz spectraux en milieu participant

Dans la section 1.2, nous avons introduit la notion de coefficient d'absorption. Pour les applications de combustion, les principaux gaz rayonnants sont  $H_2O$ ,  $CO_2$  et  $CO$ . Bien que les spectres d'absorption de ces gaz pris individuellement soient connus et accessibles pour une large plage de pression et de température [34, 35, 36, 37, 38], le calcul d'un coefficient d'absorption unique pour un jeu de paramètres physiques, et *a fortiori* pour un mélange n'est pas trivial. Il existe qui plus est une forte corrélation entre les différentes espèces rayonnantes difficile à prendre en compte et à modéliser, en particulier à basse pression.

La nature même du rayonnement est donc une difficulté pour le modélisateur. Chaque base spectroscopique des différents gaz contient plusieurs centaines de millions de transitions devant chacune être prise en compte pour le calcul rigoureux des coefficients du spectre d'absorption. Ces spectres dits « haute résolution » regroupent entre  $10^6$  et  $10^7$  coefficients d'absorption calculés pour différents jeux de paramètres physiques de pression et de température. Généralement, chaque coefficient spectral est calculé avec un pas de  $0.01 \text{ cm}^{-1}$  regroupant potentiellement plusieurs dizaines de milliers de transitions moléculaires. Ce regroupement correspond au premier niveau de simplification pour la modélisation du rayonnement. Les coefficients d'absorption sont calculés simplement comme étant la somme de l'ensemble des contributions des différentes transitions des gaz en présence. Précisons néanmoins que ce modèle nécessite déjà un certain nombre de simplifications notamment en ce qui concerne le profil théorique des raies de la base spectroscopique. Les plus utilisés sont les profils de Lorentz, de Rayleigh ou de Voigt [39].

Ce modèle de gaz est dit *Raie par Raie* ou *Line-by-Line* [40, 39] et est une référence dans la plupart des cas. Malheureusement et bien qu'il soit déjà issu d'une simplification du réel, ce modèle est déjà largement inaccessible pour la plupart des applications en 3 dimensions ou pour les calculs couplés avec une résolution CFD par exemple. Les millions de raies contenues dans ce type de modèle sont en effet autant de résolutions nécessaires de l'ETR avec une méthode DOM et d'informations spectrales à stocker pour chaque maille du maillage. Les méthodes de Monte Carlo peuvent là encore permettre de s'affranchir de ces contraintes de par leur formulation mathématique particulière au point de rendre accessible l'utilisation de ce type de modèles sur des cas complexes voire même, d'utiliser directement les bases spectroscopiques réelles des différents gaz par échantillonnage des transitions [25, 41].

Dans la majorité des cas, il est malgré tout nécessaire d'utiliser des modèles plus approximatifs permettant à la fois de traiter les mélanges gazeux simplement et de réduire le nombre de résolutions de l'équation de transfert. Sur ce point, la littérature est riche et les modèles de gaz sont nombreux. Nous proposons ici d'en introduire quelques uns brièvement et plus particulièrement ceux utilisés dans



ces travaux. Précisons que nous restreignons notre étude à un intervalle précis de nombres d'onde, significatif vis à vis des transferts thermiques, allant de l'infrarouge proche à lointain et contenu entre  $150 \text{ cm}^{-1}$  et  $9300 \text{ cm}^{-1}$ .

### Des spectres d'absorption aux modèles à bandes : le modèle SNB de Malkmus

Le modèle de MALKMUS [42] est un modèle spectral à bandes fines exprimé en transmittivités moyennes plutôt qu'en coefficient d'absorption. Du fait de la très grande complexité des spectres d'absorption complets, ce modèle pose des hypothèses statistiques sur la répartition des raies et sur la distribution de leurs intensités au sein de bandes dites *étroites*. Ces bandes correspondent à un découpage du spectre d'absorption en intervalles réguliers de  $25 \text{ cm}^{-1}$  en général. Nous nous intéressons ici en particulier au modèle corrigé de RIVIÈRE et SOUFIANI [43] adapté aux gaz issus de la combustion à haute température. Le modèle de Malkmus initial est souvent cité sous l'acronyme SNB pour *Statistical Narrow-Band*. Dans ce manuscrit, nous confondrons cette version et celle incluant les corrections adaptées à nos conditions d'étude sous le même sigle et sans distinction.

Les hypothèses généralement énoncées pour la mise en équation de ce modèle visent principalement à assurer l'indépendance statistique de la distribution des raies d'absorption au sein de l'intervalle  $\Delta\eta$  considéré [42, 1]. Sous l'ensemble des hypothèses relatives au modèle SNB, la transmittivité moyenne d'un gaz à travers une colonne homogène et isotherme peut être écrite telle que [43] :

$$\bar{\tau}_\eta = \exp \left[ -\frac{2\bar{\gamma}}{\bar{\delta}} \left( \sqrt{1 + \frac{xp l_c \bar{k} \bar{\delta}}{\bar{\gamma}}} - 1 \right) \right] \quad (1.27)$$

Avec :

- $\bar{k}$ , l'intensité moyenne de raie ne dépendant que de la température et de la bande considérée. Ce paramètre découle directement de calculs déterministes Raie par Raie pour chaque gaz rayonnant dilué dans l'air,
- $\bar{\gamma}$ , la demi-largeur moyenne du profil de raie utilisé (ici, celui de Lorentz), également issu de calculs raie par raie,
- $\bar{\delta}$ , l'espacement moyen des raies,
- $l_c$ , la longueur de la colonne uniforme considérée,
- $p$ , la pression totale,
- $x$ , la fraction molaire de l'espèce.

Dans le cas d'un mélange de gaz rayonnants, l'approximation de Curtis-Gordson est la plus couramment appliquée [44, 45]. Elle consiste à appliquer la même formule que celle utilisée pour un gaz unique (voir équation (1.27)) mais en utilisant des coefficients moyens équivalents pour le mélange.

### Modèle en k-distribution : une introduction

La famille des modèles en k-distribution est riche et les variantes sont nombreuses, si bien qu'un inventaire complet et détaillé représente en soit un travail à part entière. Nous nous concentrerons donc ici sur la version dite *k-corrélé* ou *C-k*.

Initialement développée pour l'étude du rayonnement dans l'atmosphère [46], son utilisation a largement été étendue à d'autres domaines nécessitant l'étude de mélanges gazeux en conditions

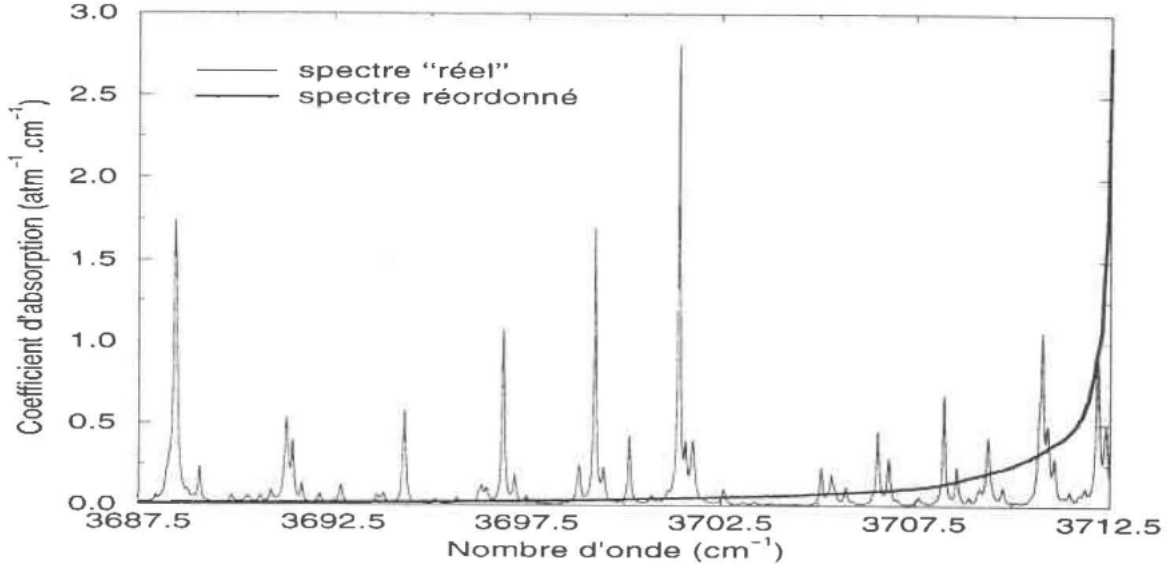


FIG. 1.7 : Exemple de réorganisation d'une portion de spectre d'absorption pour un mélange de gaz  $H_2O - N_2$  à pression atmosphérique (issu de PIERROT [1])

hétérogènes. Dans les modèles en k-distribution, nous traitons les coefficients d'absorption non plus directement par le biais de leurs spectres complets ou simplifiés mais plutôt en nous intéressant à leur distribution statistique et cumulée sur un intervalle donné. Cette dernière fonction est notée  $g(\kappa)$  et peut être explicitée telle que :

$$g(\kappa) = \frac{1}{\eta_2 - \eta_1} \int_{\eta \in [\eta_1, \eta_2], \kappa_\eta < \kappa} d\eta \quad (1.28)$$

Le principe des modèles en k-distribution consiste en fait à réorganiser le spectre d'absorption du gaz sur des portions d'intervalle  $[\eta_1, \eta_2]$  (voir figure 1.7). Nous pouvons montrer qu'en introduisant la fonction réciproque de  $g(\kappa)$  notée  $\kappa(g)$  dans l'expression de la transmittivité moyenne sur une bande spectrale, nous obtenons pour une colonne de gaz homogène :

$$\bar{\tau}_\eta(l) = \int_0^1 \exp[-xp \kappa(g)l] dg \quad (1.29)$$

L'intérêt de cette expression réside dans le fait que la fonction  $\kappa(g)$  est une fonction monotone et croissante relativement lisse. L'intégrale présentée dans l'équation (1.29) est donc facilement estimable par l'utilisation de méthodes d'intégration numériques comme des quadratures de Gauss-Legendre par exemple. Dans ce cas, l'expression de la transmittivité moyenne sur la colonne peut-être approchée pour une somme finie sur les points de quadrature :

$$\bar{\tau}_\eta(l) \approx \sum_{i=1}^I \omega_i^{GL} \exp[-xp \kappa(g_i)l] \quad (1.30)$$

Où  $I$ ,  $g_i$  et  $\omega_i^{GL}$  représentent respectivement le nombre de points de quadrature, les points et les poids qui leur sont associés. Le développement mathématique de ce modèle est largement référencé et disponible dans les ouvrages de MODEST [6] et de SIEGEL et HOWELL [5] par exemple. Néanmoins,

la version originale de ce modèle se montre inefficace si ce n'est inapplicable dans des conditions inhomogènes de température et de mélange étant donné que le coefficient d'absorption  $\kappa_\eta$  varie sur la trajectoire considérée. Il existe donc dans ce cas une fonction de distribution cumulée et une réciproque différente en chaque point d'abscisse  $s$ . Dans ce cas, une autre hypothèse peut être posée pour traiter ce type de problème et appliquer la version  $C-k$  de la méthode originale.

**Méthode  $C-k$  pour le traitement de conditions hétérogènes** Ce modèle a été développé initialement pour l'étude du rayonnement atmosphérique en astrophysique par GOODY et al. [47]. Il a par la suite été étendu pour des applications de combustion et régulièrement mis à jour [48, 43] pour sa précision [49, 50] et sa facilité de mise en œuvre sous des conditions hétérogènes. On considère tout d'abord ici que la fonction de Planck sur l'intervalle spectral de la bande considérée est constant. En toute rigueur, lorsque le réordonnement des coefficients d'absorption change avec les conditions thermo-physiques sur un intervalle spectral, le lien entre la fonction de distribution de ces coefficients et sa cumulée change également. La notion de scalabilité entre coefficients d'absorption est perdue [51].

Ce modèle se base au contraire sur le fait que cette relation est indépendante des conditions de température et de mélange et donc par conséquent, indépendante de la position du photon dans le milieu hétérogène. On considère ainsi en d'autres termes qu'il existe une *correlation* entre les fonctions de distribution de  $\kappa$  pour différentes conditions thermo-physiques.

Ce postulat est fondamentalement très fort car il ne se base sur aucune réalité physique. Néanmoins, nous pouvons remarquer que la plupart des paramètres influant sur le rayonnement sont des fonctions monotones ou tout du moins, très simples de  $\kappa$  et donc de  $g$  [47]. On se rapproche alors du comportement d'un spectre d'Elsasser ou d'un spectre à raie unique [52] où, pour la même valeur de  $g$  en différents points, nous obtiendrons une valeur de  $\kappa$  équivalente.

*i*

De manière générale, le traitement de mélanges gazeux se fait sous couvert d'une hypothèse forte qui considère les spectres d'absorption des différentes espèces rayonnantes ( $CO_2$  et  $H_2O$  par exemple en combustion) comme séparables. Factuellement, cette hypothèse n'est pas vraie mais reste une approximation acceptable dans la majorité des cas d'étude avec une erreur générée de l'ordre du pourcent [1].

Pour des modèles formulés en transmissivité, on calcule donc pour le mélange une valeur moyenne  $\overline{\tau_\eta}_{eq}$  comme étant le produit des transmittivités du mélange :

$$(\overline{\tau_\eta})_{eq} = \prod_{m=1}^M (\overline{\tau_\eta})_m \quad (1.31)$$

Où  $m$  désigne l'indice des différents gaz participant au rayonnement. D'après l'équation (1.30) et si on se limite à un mélange de deux gaz par soucis de concision ( $M = 2$ ), alors la transmittivité moyenne du mélange peut s'exprimer :

$$(\overline{\tau_\eta})_{eq} = \sum_{i=1}^I \sum_{j=1}^J (\omega_i^{GL})_1 (\omega_j^{GL})_2 \exp[-xp[(\kappa_i)_1 + (\kappa_j)_2]l] \quad (1.32)$$

On remarque ainsi que les points de quadrature du mélange sont modifiés. Les poids afférents sont égaux au produit des poids individuels de chaque gaz tandis que le coefficient d'absorption est lui égal à la somme de leurs coefficients propres.

## Un hybride pertinent pour les mélanges de gaz : le modèle SNB-CK

---

Les deux modèles précédents ont la réputation d'être particulièrement précis et sont souvent pris comme référence en l'absence de données issues de modèle RPR. Néanmoins, ils présentent malgré tout une sérieuse limitation dans leur utilisation à cause de leur expression en transmittivité en non en coefficient d'absorption. Cette particularité rend leur implémentation difficile et délicate. Pour pallier ce problème, un modèle hybride [47, 53] nommé *SNB-CK* représente une bonne alternative pour une implémentation simple et rapide en coefficient d'absorption d'un modèle approchant la précision du modèle historique de Malkmus par exemple [54].

Dans ce modèle, nous partons dans un premier temps de l'expression de la transmittivité proposée dans le modèle SNB et explicitée dans l'équation (1.27). Il est alors possible d'obtenir par l'inverse de la transformée de Laplace de cette expression, la fonction de distribution des coefficients d'absorption dans la bande spectrale d'intérêt et par extension, l'expression de sa cumulée [50, 53]. Dans la littérature, l'expression de la transmittivité est parfois écrite sous la forme suivante [50] :

$$\bar{\tau}_\eta(l) = \exp \left[ -\frac{\pi B}{2} \left( \sqrt{1 + \frac{4SL}{\pi B}} - 1 \right) \right] \quad (1.33)$$

Avec  $B = 4\frac{\bar{\nu}}{\delta}$  et  $S = \bar{k}xp$ . Avec cette notation, la fonction de distribution des coefficients d'absorption cumulée peut s'écrire analytiquement telle que [53] :

$$g(\kappa) = \frac{1}{2} \left[ 1 - \operatorname{erf} \left( \frac{a}{\sqrt{\kappa}} - b\sqrt{\kappa} \right) \right] + \frac{1}{2} \left[ 1 - \operatorname{erf} \left( \frac{a}{\sqrt{\kappa}} + b\sqrt{\kappa} \right) \right] \exp(\pi B) \quad (1.34)$$

Avec  $a = \frac{1}{2}\sqrt{\pi BS}$  et  $b = \frac{1}{2}\sqrt{\frac{\pi B}{S}}$ . Nous pouvons alors calculer le coefficient d'absorption moyen sur la bande spectrale pour chaque gaz par quadrature de Gauss-Legendre à 16 points (voir tableau 1.3) par exemple tel que :

$$\kappa_\eta \approx \sum_{i=1}^{16} \omega_i^{GL} \kappa_i(g_i) \quad (1.35)$$

La valeur de  $\kappa_i$  peut être déterminée par inversion de l'équation (1.34) pour un point de quadrature connu  $g_i$  par méthode de Newton-Raphson [50]. Une fois le coefficient d'absorption moyen sur la bande spectrale calculée pour chaque espèce rayonnante, il suffit ensuite de sommer les contributions de chaque gaz pour traiter un mélange. Nous rappelons que la fraction molaire de chaque espèce est bien prise en compte dans l'expression de la transmittivité (équation (1.27)) dans la mesure où sa contribution n'est pas explicitement mise en évidence dans ce paragraphe.

Comme nous le verrons au chapitre 3, nous avons fait le choix de travailler avec ce modèle pour plusieurs raisons :

- Il a été utilisé en complément des méthodes de Monte Carlo dans les cas d'études de Benchmark proposés par [2]. Cet article est utilisé dans le chapitre 4 pour validation de la méthode développée dans cette thèse et de l'ensemble des codes utilisés.
- Son implémentation est relativement simple et proche de celle du modèle à haute pression HPBM utilisé comme référence par Safran (voir section 1.4.4)

Indice	Point $g_i$	Poids $\omega_i^{GL}$
1	0.0052995325	0.0135762297
2	0.0277124885	0.0311267619
3	0.0671843988	0.0475792558
4	0.1222977958	0.0623144856
5	0.1910618777	0.0747979944
6	0.2709916111	0.0845782597
7	0.3591982246	0.0913017075
8	0.4524937451	0.0947253052
9	0.5475062549	0.0947253052
10	0.6408017754	0.0913017075
11	0.7290083888	0.0845782597
12	0.8089381222	0.0747979944
13	0.8777022042	0.0623144856
14	0.9328156012	0.0475792558
15	0.9722875115	0.0311267620
16	0.9947004675	0.0135762297

TAB. 1.3 : Quadrature de Gauss Legendre à 16 points

Notons que le modèle de gaz choisi n'a finalement que peu d'importance dans la construction de la méthode proposé dans ce manuscrit. Bien que nécessaire aux différentes simulations menées, il ne constitue qu'un brique interchangeable par l'utilisateur en fonction de ses contraintes.

### Passage aux hautes pressions : des modèles à bandes aux *Box models*

Les modèles présentés précédemment ont été ou sont encore très utilisés pour des gammes de pression proches de celle de l'atmosphère. Dans le cas d'une chambre de combustion aéronautique, les pressions en jeu sont bien plus élevées et sont plutôt comprises entre 10 et 30 bars. Dans ces conditions, les spectres d'absorption des espèces rayonnantes changent fondamentalement si bien qu'il n'est plus possible de modéliser les raies par un profil de Lorentz [55, 56, 57, 58]. Les raies s'élargissant avec la pression, le couplage entre chacune d'entre elle augmente et l'hypothèse de collision infiniment rapide entre les différentes molécules ne tient plus.

L'accès aux données spectroscopiques elles-mêmes fût un problème pendant longtemps pour des conditions de pression et de température aussi élevées. Ainsi et jusqu'au début des années 90, il n'existait pas de modèle à basse ou moyenne résolution capable de modéliser les transferts thermiques par rayonnement sans avoir à utiliser des modèles RPR [59, 60].

Nous nous intéressons dans ce paragraphe au modèle HPBM (*High Pressure Box Model*) développé par PIERROT [1] dans sa thèse en collaboration avec la SNECMA devenue aujourd'hui Safran Aircraft Engines. Ce modèle a donc historiquement été développé pour les applications moteurs dans le domaine aéronautique. Bien conscient des enjeux et des problématiques liés aux simulations couplées menées par l'industrie, l'auteur a proposé deux versions de son modèle. La première à 367 bandes fines est basée sur le modèle SNB historique de MALKMUS [42] et utilise les paramètres spectraux de TAINE et SOUFIANI [48]. La seconde version est une variante simplifiée à 26 bandes fines beaucoup moins gourmande en termes de ressources de calcul.

i

Notons ici que nous utilisons l'adjectif *fine* indépendamment de la largeur réelle (en terme d'intervalle) des bandes spectrales. En effet, un modèle dit à bande *large* est fondamentalement différent dans sa formulation et dans les hypothèses posées pour sa définition [61, 62, 63]

On peut montrer par un calcul raie par raie utilisant un modèle de Lorentz tronqué que pour une même quantité de gaz absorbant, le spectre d'absorption à pression atmosphérique est bien plus fluctuant qu'à 40 bars [1]. Ce comportement est également visible mathématiquement dans la formulation du modèle de MALKMUS [42]. En effet, dans l'équation (1.27), on remarque que si la valeur de  $\bar{\gamma}$  augmente du fait de sa relation de proportionnalité avec la pression, alors la racine carrée se linéarise. La transmittivité moyenne sur la bande spectrale tend alors vers l'approximation suivante :

$$\bar{\tau}_\eta = \exp(xpl\bar{\kappa}) \quad (1.36)$$

C'est sur cette relation que le modèle HPBM repose entièrement. Elle implique notamment des simplifications de mise en oeuvre considérables pour le traitement de mélanges anisothermes et inhomogènes par l'abandon de l'approximation de Curtis Gordon classiquement nécessaire avec un modèle SNB. Elle permet aussi de revenir à un modèle exprimé en coefficient d'absorption facilement exploitable pour la résolution de l'ETR. Ce modèle a été développé pour pouvoir prendre en compte les trois principales espèces rayonnantes rencontrées en combustion à savoir  $H_2O$ ,  $CO_2$  et  $CO$ .

**Modèle à 367 bandes** Ce modèle couvre la même plage spectrale *utile* dans l'infrarouge que celle sur laquelle nous travaillons depuis le début dans ce manuscrit, à savoir l'intervalle  $[150, 9300]cm^{-1}$  divisé en 367 bandes spectrales de  $25cm^{-1}$  de largeur. Les paramètres mis à jour par TAINE et SOUFIANI [48] ont été utilisés pour construire ce modèle.

La luminance totale en un point  $s$  de la trajectoire du photon considéré est égale à la somme des luminances *monochromatiques*. Notons que cette sémantique n'est pas tout à fait rigoureuse dans la mesure où la luminance sur chaque bande est moyennée sur un intervalle de nombre d'ondes. Elle n'est donc pas, dans l'absolu, purement monochromatique au sens physique du terme. Cette précision faite, nous pouvons alors écrire :

$$L(s) = \sum_{i=1}^{367} L_{\Delta\eta_i} \Delta\eta_i \quad (1.37)$$

Dans cette expression et bien que l'intervalle  $\Delta\eta_i$  soit indicé, la largeur des bandes reste constante. Nous verrons néanmoins que pour la version simplifiée du modèle HPBM à 26 bandes, cette précision de notation trouvera son intérêt. Cette équation est à compléter avec l'ETR en milieu purement absorbant, nécessaire au calcul des luminances de bande, dont nous pouvons redonner la formulation différentielle :

$$\frac{\partial L_{\Delta\eta_i}}{\partial s} = \kappa_{\Delta\eta}(i) \left[ L_{\Delta\eta_i}^0(T(s)) - L_{\Delta\eta_i}(s) \right] \quad (1.38)$$

La luminance du corps noir ne dépend que de la température et est supposée constante sur chaque bande fine comme le prévoient les hypothèses posées pour ce type de modèle. La valeur du coefficient d'absorption global  $\kappa_{\Delta\eta}$  est déterminée d'après la somme pondérée des coefficients d'absorption de chaque espèce en présence dans le mélange :

$$\kappa_{\Delta\eta} = p [x_{CO_2} \bar{\kappa}_{CO_2} + x_{H_2O} \bar{\kappa}_{H_2O} + x_{CO} \bar{\kappa}_{CO}] \quad (1.39)$$

Dans cette expression, les variables  $x$  représente les fractions molaires des espèces et  $p$ , la pression. Les valeurs des  $\kappa$  pour chaque espèce sont directement issues de TAINE et SOUFIANI [48] aux conditions thermophysiques connues au point d'abscisse curviligne  $s$ . La mise en oeuvre de ce modèle en milieu inhomogène et anisotherme est donc, en dehors de toute considération liée à la puissance de calcul nécessaire, simple et immédiate. La formulation en coefficient d'absorption de ce modèle permet l'utilisation de la plupart des méthodes de résolution de l'ETR sans problème particulier d'implémentation. Si malgré tout, l'utilisation de 367 bandes spectrales devient très vite prohibitive pour l'industrie en résolution par DOM (voir section 1.3.2), les méthodes de Monte Carlo rendent tout à fait accessible l'utilisation de modèles de ce type.

**Modèle simplifié à 26 bandes** L'objectif de cette version est en soi assez simple dans la mesure où il vise simplement à réduire la puissance de calcul nécessaire à la modélisation correcte du rayonnement dans le contexte d'une simulation plus globale, couplée avec d'autres phénomènes. L'idée est de regrouper certaines bandes spectrales sur lesquelles il est possible de postuler sans trop d'erreur, que le coefficient d'absorption moyen de ces nouvelles bandes soit proche de la moyenne des coefficients initiaux. Ce regroupement n'est possible qu'à la condition où les propriétés radiatives des bandes sont comparables. Si ce prérequis ne pose pas trop de problème pour le spectre de  $H_2O$ , les spectres du  $CO_2$  et du  $CO$  montrent des zones de transparence importantes et très localisées sur l'intervalle de nombre d'onde considéré.

Ce travail mené dans la thèse de PIERROT [1] a permis le développement d'un modèle simplifié à 26 bandes d'épaisseur variable (voir tableau 1.4) sur l'ensemble du spectre infrarouge. Notons que ces 26 bandes ont été choisies en considérant les zones de transparence du  $CO_2$  principalement. La prise en compte du rayonnement du  $CO$  est donc très approximative. Un modèle respectant les zones de transparence de ces deux espèces aurait mené à 31 bandes spectrales ; ce qui est certes, un gain substantiel par rapport aux 367 bandes d'origine mais qui reste encore insuffisant pour une mise en pratique à échelle industrielle. Dans les applications moteur, la quantité de molécules de  $CO$  reste suffisamment à la marge devant celles de  $CO_2$  et de  $H_2O$  pour se permettre ce genre d'approximation.

i

Dans cette version à 26 bandes, il n'est plus tout à fait rigoureux de considérer la fonction de Planck comme constante sur les bandes spectrales élargies. Dans le but de limiter les erreurs induites par cette hypothèse, nous prenons le résultat de cette fonction à la température locale au point  $s$  et au nombre d'onde central de la bande.

**Précision des versions du modèle HPBM** Dans sa thèse, PIERROT [1] compare la précision de ses modèles à un modèle SNB de Malkmus pris comme référence pour un intervalle de pression allant de 5 à 40 atmosphères et de température entre 900 et 2500 K. Les comparaisons ont été faites pour chaque gaz séparément dilués dans le diazote. Sans être exhaustif, il apparaît clairement que le modèle complet à 367 bandes mène à des écarts relatifs faibles par rapport à la référence de l'ordre de 0.5 % pour le  $CO_2$  jusqu'à 5 % pour le  $CO$ . En revanche, le modèle à 26 bandes engendre comme prévu une augmentation globale des erreurs en particulier pour le monoxyde de carbone pouvant aller jusque 50 %. Les résultats complets des études menées pour la validation de ces modèles sont disponibles dans le manuscrit de PIERROT [1].

Indice $i$ de la bande spectrale	Centre $\eta_i$ de la bande ( $cm^{-1}$ )	Largeur $\Delta\eta_i$ de la bande ( $cm^{-1}$ )
1	312.5	350
2	562.5	150
3	712.5	150
4	887.5	200
5	1087.5	200
6	1312.5	250
7	1562.5	250
8	1812.5	250
9	2000	125
10	2112.5	100
11	2212.5	100
12	2337.5	150
13	2612.5	400
14	3012.5	400
15	3312.5	200
16	3487.5	150
17	3662.5	200
18	3862.5	200
19	4062.5	200
20	4462.5	600
21	4912.5	300
22	5362.5	600
23	6012.5	700
24	6662.5	600
25	7262.5	600
26	8437.5	1750

TABLE. 1.4 : Définition des bandes spectrales dans le modèle HPBM26 (issu de PIERROT [1])

### Rayonnement des suies : un modèle semi-empirique

Dans les applications de combustion, les gaz participants ne sont pas les seuls contributeurs au rayonnement thermique. En effet, en présence d'une flamme de diffusion notamment, la richesse du mélange carburant/comburant n'est pas constante. Par conséquent, le mélange n'est que rarement aux conditions stoechiométriques et la production de sous-produits de combustion est inévitable. On compte notamment parmi eux du carburant imbrulé, des radicaux libres et des suies. Ces dernières sont des agrégats de carbone et participent significativement au rayonnement thermique du gaz dans son ensemble.

La présence de suies est souvent caractérisée par leur fraction volumique  $f_v$ . Leur contribution, bien que fondamentalement différente de celles des gaz, peut être prise en compte sans changement significatif de l'ETR. Elle se traduit par le simple ajout d'un coefficient d'absorption spécifique au coefficient des gaz précédemment décrit :

$$\kappa_{abs,\eta} = \kappa_{abs,\eta}^{gaz} + \kappa_{abs,\eta}^{suie} \quad (1.40)$$



En supposant que les suies sont non agglomérées et que leur propriétés optiques suivent la théorie des particules fines de Rayleigh [6], leur émission spectrale peut ainsi être approchée par un modèle semi-empirique simple, considérant que le rayonnement des suies est proportionnel à leur fraction volumique [64] :

$$\kappa_{abs,\eta}^{sui} = 5.5f_v\eta \quad (1.41)$$

La constante multiplicative fixée ici à 5.5 peut être discutée bien qu'un consensus semble exister pour les applications traitées dans ce manuscrit. Sa valeur peut néanmoins différer dans la littérature en fonction du type de flamme étudiée [65].

En comparaison avec les caractéristiques radiatives très variables des gaz, le comportement des suies est donc relativement simple et continu. Cela implique notamment un rayonnement sur l'ensemble du spectre infrarouge qui va diminuer proportionnellement avec les nombres d'onde.

### 1.4.5 Choix des modèles utilisés et solution nouvelle de modélisation du rayonnement

Ce chapitre a tout d'abord été consacré à la présentation du contexte dans lequel s'inscrivent ces travaux. Malgré des outils éprouvés et utilisés depuis longtemps, les méthodes actuelles de modélisation du rayonnement dans les conditions nouvelles d'opération des chambres de combustion ne sont plus suffisantes.

Les méthodes de Monte Carlo présentent beaucoup d'avantages très intéressants. Elles permettent notamment l'utilisation de modèles de gaz spectraux sans surcoût prohibitif. Leur fonctionnement par « points sondes » est en revanche bloquant lorsqu'il est question de travailler en champ sur des maillages très volumineux. Conceptuellement, ces méthodes ne sont pas faites pour résoudre ce genre de tâche. L'objet de ce manuscrit est de venir compléter les qualités des méthodes de Monte Carlo par l'utilisation de réseaux de neurones (voir chapitre 2). Ces outils peuvent être utilisés comme des interpolateurs universels. Nous montrons dans le chapitre 4 le potentiel qu'ils représentent pour la généralisation de solutions discrètes produites par méthode de Monte Carlo et pour la production de champs complets de grandeurs radiatives.

Comme cela a été évoqué précédemment, les modèles SNB-CK pour le mélange gazeux à pression atmosphérique et le modèle HPBM367 historique de Safran pour les cas à haute pression sont utilisés pour leurs précisions et leur facilité d'implémentation. Le chapitre 3 développe les raisons de ces choix plus en détails. Il est important de préciser encore une fois que ces modèles sont évidemment interchangeables avec d'autres en fonction des objectifs de l'utilisateur de cette méthode.

## Références

- [1] L. PIERROT. « Développement, étude critique et validation de modèles de propriétés radiatives infrarouges de CO<sub>2</sub> et H<sub>2</sub>O à haute température. Applications au calcul des transferts dans des chambres aéronautiques et à la télédétection. » Thèse de doct. 1997.
- [2] P. J. COELHO, P. PEREZ et M. EL HAFI. « Benchmark numerical solutions for radiative heat transfer in two-dimensional axisymmetric enclosures with nongray sooting media ». In : *Numerical Heat Transfer, Part B : Fundamentals* 43.5 (2003), p. 425-444.
- [3] H. I. H. SARAVANAMUTTOO et al. *Gas Turbine Theory*. Seventh Edition. Pearson education, 2017.

- [4] H. C. HOTTEL. «McAdams, Heat Transmission». In : sous la dir. de MCGRAW-HILL. 3<sup>e</sup> éd. 1954. Chap. Radiant-Heat Transmission.
- [5] R. SIEGEL et J. R. HOWELL. *Thermal Radiation Heat Transfer*. 5<sup>e</sup> éd. CRC Press, 2010. 982 p.
- [6] M. F. MODEST. *Radiative heat transfer*. Sous la dir. d'A. PRESS. 2th. 2003.
- [7] M. GALTIER et al. «Integral formulation of null-collision Monte Carlo algorithms». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 125 (2013), p. 57-68. DOI : [10.1016/j.jqsrt.2013.04.001](https://doi.org/10.1016/j.jqsrt.2013.04.001).
- [8] M. EL HAFI et al. «Three viewpoints on null-collision Monte Carlo algorithms». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 260 (2021).
- [9] 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». PhD Thesis. Ecole Nationale Supérieure des Mines d'Albi-Carmaux, Université de Toulouse, 2020.
- [10] N. VILLEFRANQUE et al. «A path tracing Monte Carlo library for 3-D radiative transfer in highly resolved cloudy atmospheres». In : *Journal of Advances in Modeling Earth systems* 11 (2019).
- [11] M. ROGER. «Modèles de sensibilité dans le cadre de la méthode de Monte-Carlo : illustrations en transfert radiatif». Thèse de doct. Institut National Polytechnique de Toulouse, 2006.
- [12] T. F. SMITH, Z. F. SHEN et J. N. FRIEDMAN. «Evaluation of Coefficients for the Weighted Sum of Gray Gases model». In : *ASME JHT* 104 (1982), p. 602-608.
- [13] A. COPPALLE et P. VERVISCH. «The total emissivities of high-temperature flames». In : *Combustion and Flame* 49.1 (1983), p. 101-108. DOI : [10.1016/0010-2180\(83\)90154-2](https://doi.org/10.1016/0010-2180(83)90154-2).
- [14] M. F. MODEST. «The Weighted-Sum-of-Gray-Gases Model for Arbitrary Solution Methods in Radiative Transfer». In : *Journal of Heat Transfer* 113.3 (1991), p. 650-656. DOI : [10.1115/1.2910614](https://doi.org/10.1115/1.2910614).
- [15] M. K. DENISON. «A spectral line-based weighted-sum-of-gray-gases model for arbitrary RTE solvers». Thèse de doct. Brigham Young University, 1994.
- [16] M. K. DENISON et B. W. WEBB. «k-distributions and Weighted-Sum-of-Gray-Gases-A hybrid model». In : *International Heat Transfer Conference Digital Library*. Begel House Inc. 1994.
- [17] «The spectral-line weighted-sum-of-gray-gases model for H<sub>2</sub>O/CO<sub>2</sub> mixtures». In : *Journal of heat transfer* 117.3 (1995), p. 788-792.
- [18] M. DENISON et B. WEBB. «An absorption-line blackbody distribution function for efficient calculation of total gas radiative transfer». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 50.5 (1993), p. 499-510. DOI : [10.1016/0022-4073\(93\)90043-H](https://doi.org/10.1016/0022-4073(93)90043-H).
- [19] M. K. DENISON et B. W. WEBB. «Development and application of an absorption- line blackbody distribution function for CO<sub>2</sub>». In : 38.10 (1995).
- [20] J. T. PEARSON. «The Development of Updated and Improved SLW Model Parameters and Its Application to Comprehensive Combustion Predictions». Thèse de doct. Brigham Young University, 2013.
- [21] F. ANDRÉ, V. P. SOLOVJOV et B. W. WEBB. «The  $\omega$ -absorption line distribution function for rank correlated SLW model prediction of radiative transfer in non-uniform gases». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 280 (2022). DOI : [10.1016/j.jqsrt.2022.108081](https://doi.org/10.1016/j.jqsrt.2022.108081).
- [22] M. GALTIER et al. «Assessment of narrow-band and full spectrum gas radiation methods in a real industrial glass furnace configuration». In : *Applied Thermal Engineering* 216 (2022), p. 119020. DOI : [10.1016/j.applthermaleng.2022.119020](https://doi.org/10.1016/j.applthermaleng.2022.119020).
- [23] J. M. HAMMERSLEY et D. C. HANDSCOMB. *Monte Carlo Methods*. London : Chapman et Hall, 1965.

- [24] J. DELATORRE et al. « Monte Carlo advances and concentrated solar applications ». In : *Solar Energy* 103 (2014), p. 653-681. DOI : [10.1016/j.solener.2013.02.035](https://doi.org/10.1016/j.solener.2013.02.035).
- [25] M. GALTIER. « Approche statistique du rayonnement dans les milieux gazeux hétérogènes : de l'échantillonnage des transitions moléculaires au calcul de grandeurs radiatives ». Thèse de doct. 2014.
- [26] W. L. DUNN et J. K. SHULTIS. *Exploring monte carlo methods*. Elsevier, 2022.
- [27] M. GALTIER et al. « Integral formulation of null-collision Monte Carlo algorithms ». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 125 (2013), p. 57-68.
- [28] H. R. SKULLERUD. « The stochastic computer simulation of ion motion in a gas subjected to a constant electric field ». In : *Journal of Physics D : Applied Physics* 1.11 (1968), p. 1567. DOI : [10.1088/0022-3727/1/11/423](https://doi.org/10.1088/0022-3727/1/11/423).
- [29] E. WOODCOCK et al. « Techniques used in the GEM code for Monte Carlo neutronics calculations in reactors and other systems of complex geometry ». In : *Proc. Conf. Applications of Computing Methods to Reactor Problems*. T. 557. 2. Argonne National Laboratory. 1965.
- [30] E. BRUN et al. « TRIPOLI-4®, CEA, EDF and AREVA reference Monte Carlo code ». In : *Annals of Nuclear Energy* 82 (2015), p. 151-160. DOI : [10.1016/j.anucene.2014.07.053](https://doi.org/10.1016/j.anucene.2014.07.053).
- [31] M. KERKER. *The scattering of light and other electromagnetic radiation*. Elsevier, 1969.
- [32] P. BOULET et al. « PROPRIETES RADIATIVES DES MILIEUX DIFFUSANTS ». École thématique CNRS « Rayonnement thermique en milieux semi-transparents », 2014.
- [33] A. LIEMERT et A. KIENLE. « Radiative transfer in two-dimensional infinitely extended scattering media ». In : *Journal of Physics A : Mathematical and Theoretical* 44.50 (2011). DOI : [10.1088/1751-8113/44/50/505206](https://doi.org/10.1088/1751-8113/44/50/505206).
- [34] N. JACQUINET-HUSSON et al. « The 2009 edition of the GEISA spectroscopic database ». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 112.15 (oct. 2011), p. 2395-2445. DOI : [10.1016/j.jqsrt.2011.06.004](https://doi.org/10.1016/j.jqsrt.2011.06.004).
- [35] L. ROTHMAN et al. « HITEMP, the high-temperature molecular spectroscopic database ». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 111.15 (oct. 2010), p. 2139-2150. DOI : [10.1016/j.jqsrt.2010.05.001](https://doi.org/10.1016/j.jqsrt.2010.05.001).
- [36] L. ROTHMAN et al. « The HITRAN2012 molecular spectroscopic database ». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 130 (nov. 2013), p. 4-50. DOI : [10.1016/j.jqsrt.2013.07.002](https://doi.org/10.1016/j.jqsrt.2013.07.002).
- [37] S. TASHKUN et V. PEREVALOV. « CDSD-4000 : High-resolution, high-temperature carbon dioxide spectroscopic databank ». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 112.9 (juin 2011), p. 1403-1410. DOI : [10.1016/j.jqsrt.2011.03.005](https://doi.org/10.1016/j.jqsrt.2011.03.005).
- [38] S. TASHKUN et al. « CDSD-296, high-resolution carbon dioxide spectroscopic databank : An update ». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 228 (mai 2019), p. 124-131. DOI : [10.1016/j.jqsrt.2019.03.001](https://doi.org/10.1016/j.jqsrt.2019.03.001).
- [39] L. SPARKS. « Efficient line-by-line calculation of absorption coefficients to high numerical accuracy ». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 57.5 (mai 1997), p. 631-650. DOI : [10.1016/S0022-4073\(96\)00154-9](https://doi.org/10.1016/S0022-4073(96)00154-9).
- [40] J. TAINE. « A line-by-line calculation of low-resolution radiative properties of CO<sub>2</sub>-CO-Transparent nonisothermal gases mixtures up to 3000K ». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 30.4 (1983), p. 371-379.
- [41] Y. NYFFENEGGER-PERE et al. « Une estimation exacte du forçage radiatif par la méthode de Monte-Carlo ». In : *SFT 2022-30ème Congrès Français de Thermique 2022*. 2022.
- [42] W. MALKMUS. « Random Lorentz Band Model with Exponential-Tailed S<sup>-1</sup> Line-Intensity Distribution Function\* ». In : *Journal of the Optical Society of America* 57.3 (1<sup>er</sup> mars 1967), p. 323. DOI : [10.1364/JOSA.57.000323](https://doi.org/10.1364/JOSA.57.000323).

- [43] P. RIVIÈRE et A. SOUFIANI. «Updated band model parameters for H<sub>2</sub>O, CO<sub>2</sub>, CH<sub>4</sub> and CO radiation at high temperature». In : *International Journal of Heat and Mass Transfer* 55.13 (2012), p. 3349-3358. DOI : [10.1016/j.ijheatmasstransfer.2012.03.019](https://doi.org/10.1016/j.ijheatmasstransfer.2012.03.019).
- [44] A. SOUFIANI, J. M. HARTMANN et J. TAINÉ. «Validity of band-model calculations for CO<sub>2</sub> and H<sub>2</sub>O applied to radiative properties and conductive-radiative transfer». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 33.3 (1985), p. 243-257.
- [45] S. J. YOUNG. «Non-isothermal band model theory». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 18 (1977), p. 1-28.
- [46] R. M. GOODY et Y. L. YUNG. *Atmospheric radiation : theoretical basis*. Oxford university press, 1995.
- [47] R. GOODY et al. «The correlated-k method for radiation calculations in nonhomogeneous atmospheres». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 42.6 (1989), p. 539-550.
- [48] J. TAINÉ et A. SOUFIANI. «High temperature gas radiative property parameters of statistical narrow-band model for H<sub>2</sub>O, CO<sub>2</sub> and CO, and correlated-K model for H<sub>2</sub>O and CO<sub>2</sub>». In : 40 (1997), p. 987-991.
- [49] L. PIERROT, A. SOUFIANI et J. TAINÉ. «Accuracy of narrow-band and global models for radiative transfer in H<sub>2</sub>O, CO<sub>2</sub>, and H<sub>2</sub>O-CO<sub>2</sub> mixtures at high temperature». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 62 (1999), p. 523-548. DOI : [10.1016/S0022-4073\(98\)00125-3](https://doi.org/10.1016/S0022-4073(98)00125-3).
- [50] F. LIU et G. J. SMALLWOOD. «Application of the statistical narrow-band correlated-k method to low-resolution spectral intensity and radiative heat transfer calculations Effects of the quadrature scheme». In : *Int. J. Heat Mass Transfer* 43 (2000).
- [51] F. ANDRE et al. *Rayonnement des gaz : des spectres de raies aux modèles approchés*. Four Solaire d'Odeillo : Ecole thématique CNRS, 2014.
- [52] C. B. LUDWIG et al. *Handbook of infrared radiation from combustion gases*. National Aeronautics and space administration (NASA). 1973. 496 p.
- [53] A. A. LACIS et V. OINAS. «A description of the correlated-k distribution method for modeling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres». In : *J. Geophysical Research* 96 (1991).
- [54] V. GOUTIERE, F. LIU et A. CHARETTE. «An assessment of real-gas modelling in 2D enclosures». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 64.3 (2000), p. 299-326. DOI : [10.1016/S0022-4073\(99\)00102-8](https://doi.org/10.1016/S0022-4073(99)00102-8).
- [55] S. A. CLOUGH, F. X. KNEIZYS et R. W. DAVIES. «Line Shape and the Water Vapor Continuum». In : *Atmospheric Research* 23 (1989), p. 229-241.
- [56] M. Y. PERRIN et J. M. HARTMANN. «Temperature-dependent measurements and modeling of absorption by CO<sub>2</sub>-N<sub>2</sub> mixtures in the far line\_wings of the 4.3 μm CO<sub>2</sub> band». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 42.4 (1989), p. 311-317.
- [57] J. M. HARTMANN et al. «The infrared continuum of pure water vapor : calculations and high-temperature measurements». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 49.6 (1993), p. 675-691.
- [58] C. BRODBECK et al. «Absorption of radiation by gases from low to high pressures. II. Measurements and calculations of CO infrared spectra». In : *Journal de Physique II* 4.12 (1994), p. 2101-2118. DOI : [10.1051/jp2:1994249](https://doi.org/10.1051/jp2:1994249).
- [59] J. M. HARTMANN et al. «Absorption of radiation by gases from low to high pressures. I. Empirical line-by-line and narrow-band statistical models». In : *Journal de Physique II* 1.7 (1991), p. 739-762. DOI : [10.1051/jp2:1991107](https://doi.org/10.1051/jp2:1991107).
- [60] M. DENISON et B. WEBB. «The absorption-line blackbody distribution function at elevated pressures». In : *1st International Symposium on Radiative Heat Transfer, Kusadasi, Turkey*. 1995.

- 
- [61] F. ANDRÉ et R. VAILLON. «The spectral-line moment-based (SLMB) modeling of the wide band and global blackbody-weighted transmission function and cumulative distribution function of the absorption coefficient in uniform gaseous media». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 109.14 (2008), p. 2401-2416. DOI: [10.1016/j.jqsrt.2008.05.002](https://doi.org/10.1016/j.jqsrt.2008.05.002).
- [62] P. S. CUMBER, M. FAIRWEATHER et H. S. LEDIN. «Application of wide band radiation models to non-homogeneous combustion systems». In : *Int. J. Heat Mass Transfer* 41.11 (1998), p. 1573-1584.
- [63] D. EDWARDS et A. BALAKRISHNAN. «Thermal radiation by combustion gases». In : *International Journal of Heat and Mass Transfer* 16.1 (1973), p. 25-40. DOI: [10.1016/0017-9310\(73\)90248-2](https://doi.org/10.1016/0017-9310(73)90248-2).
- [64] V. P. SOLOVJOV et B. W. WEBB. «An Efficient Method for Modeling Radiative Transfer in Multi-component Gas Mixtures With Soot». In : *Journal of Heat Transfer* 123.3 (2001), p. 450-457.
- [65] C. A. HOERLLE, F. H. R. FRANÇA et P. R. PAGOT. «Effects of radiation modeling on non-premixed sooting flames simulations under oxyfuel conditions». In : *Combustion and Flame* 217 (2020), p. 294-305.



# Une solution de rupture : l'intelligence artificielle pour la modélisation physique

## Sommaire du chapitre

2.1	Du cerveau aux réseaux de neurones artificiels : origines et modèle humain . . . . .	38
2.2	Les réseaux de neurones artificiels modernes . . . . .	40
2.3	Un réseau adaptée aux contraintes industrielles . . . . .	56
	Références . . . . .	67

## Notations

Symbole	Description	Unité
$a$	Sortie d'un neurone	-
$\alpha, \beta$	Hyperparamètres de régularisation	-
$\mathbf{a}^L$	Vecteur contenant les sorties du réseau	-
$b$	Biais d'un neurone	-
$b_j^l$	Biais du neurone $j$ de la couche $l$	-
$b_1, b_2$	Bornes de normalisation d'une série de données	-
$C$	Fonction de coût appliquée sur l'ensemble de la base de données	-
$\nabla C$	Gradient de la fonction de coût par rapport aux paramètres $v_i$	-
$\Delta C$	Variation de la fonction de coût $C$	-
$C_x$	Fonction de coût appliquée sur le jeu de données $x$ uniquement	-
$\tilde{\delta}^l$	Vecteur sensibilité de Marquardt sur la couche $l$	-
$\delta_j^l$	Sensibilité du neurone $j$ de la couche $l$	-
$\delta^l, \delta^L$	Vecteur contenant les sensibilités des neurones de la couche $l$ ou de la couche de sortie $L$	-
$\tilde{\delta}$	Sensibilité de Marquardt	-
$\tilde{\delta}_x$	Vecteur sensibilité de Marquardt relatif au jeu de paramètres $x$	-
$\mathbb{E}$	Espérance	-
$e_x$	Erreur entre la sortie théorique et calculée du réseau sur le jeux de données $x$	-

Symbole	Description	Unité
$\mathbf{e}$	Vecteur contenant l'ensemble des erreurs $e_x$ de la base	-
$E_d$	Fonction de coût portant uniquement sur les paramètres en cas de régularisation	-
$\mathbb{D}$	Ensemble des données d'entraînement	-
$\mathbb{M}$	Ensemble des modèles de réseau de neurones	-
$E_\omega$	Somme quadratique des poids du réseau	-
$\gamma$	Nombre effectif de paramètres	-
$\mathbf{H}_e$	Matrice Hessienne	-
$\mathbf{I}$	Matrice identité	-
$\mathbf{J}$	Matrice jacobienne	-
$L$	Exposant se rapportant à la couche dernière couche du réseau	-
$\lambda$	Offset de la fonction sigmoïde ou tangente hyperbolique	-
$MSE$	<i>Mean Squared Error</i>	-
$N$	Nombre de jeux de données dans la base	-
$\eta$	Taux d'apprentissage ou <i>learning rate</i>	-
$new$	En indice, est relatif à un paramètre mise à jour par descente de gradient	-
$norm$	Indice relatif à une donnée normalisée	-
$n_p$	Nombre de paramètres	-
$P(A)$	Probabilité d'un évènement A	-
$P(A B)$	Probabilité d'un évènement A sachant B	-
$s$	Fonction d'activation d'un neurone	-
$\sigma$	Écart type d'une série de données	-
$T$	En exposant, désigne la transposition d'une matrice ou d'un vecteur	-
$\mathbf{v}$	Vecteur contenant l'ensemble des paramètres (poids et biais) du réseau	-
$v_i$	Composante $i$ du vecteur paramètre $\mathbf{v}$ , peut être un poids ou un biais	-
$\mathbf{v}^{MP}$	Vecteur paramètre <i>le plus probable</i>	-
$\omega$	Poids d'un réseau de neurones	-
$\omega_{jk}^l$	Poids de la branche reliant le neurone $k$ de la couche $l - 1$ au neurone $j$ de la couche $l$	-
$x_i$	$i^{\text{ÈME}}$ entrée d'un neurone	-
$x$	En indice, relatif à l'identification des jeux de données de la base	-
$x_{min}, x_{max}$	Valeurs minimum et maximum d'une série de données	-
$\bar{x}$	Valeur moyenne d'une série de données	-
$\mathbf{x}_i$	Vecteur des variables d'entrées d'un réseau du $i^{\text{ÈME}}$ jeu de données	-
$\mathbf{y}_i$	Vecteur des variables de sortie d'un réseau du $i^{\text{ÈME}}$ jeu de données	-
$z$	Sortie d'un neurone avant transformation par la fonction d'activation	-

## 2.1 Du cerveau aux réseaux de neurones artificiels : origines et modèle humain

Contrairement à l'idée populaire qui voudrait que l'intelligence artificielle au sens où on l'entend soit une innovation récente ayant révolutionné nos algorithmes de calcul, les réseaux de neurones et leur architecture trouvent en réalité leur origine au milieu du  $xx^{\text{ÈME}}$  siècle. L'idée première était de construire un modèle basé sur le fonctionnement du cerveau humain, composé de neurones inter-connectés entre eux



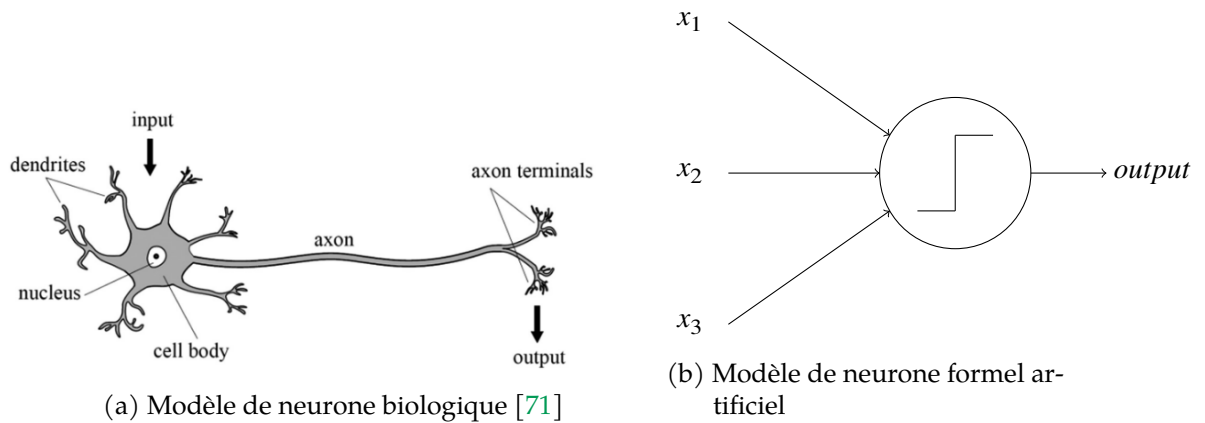


FIG. 2.1 : Analogie entre neurone biologique et artificiel

pouvant prendre deux états : excité ou non. Lorsqu'une information transite à travers un réseau neuronal humain, chaque neurone peut alors passer d'un état à l'autre en fonction d'un seuil d'activation propre, avec une temporalité différente. Les connexions entre chaque unité élémentaire du réseau peuvent également être bouclées si bien que l'information ne circule pas linéairement entre le point d'entrée du signal et sa réponse en sortie.

Un cerveau humain adulte contient entre 80 et 90 milliards de neurones connectés les uns aux autres. La réponse à une excitation initiale est donc finalement le résultat d'une logique combinatoire incroyablement complexe dictée par un apprentissage continu pendant toute la vie des individus. Des neurones et leurs liaisons peuvent être créés ou détruits continuellement dans le seul but de faire correspondre la réponse de notre cerveau à une référence connue. C'est ainsi par exemple que nous sommes capables de faire la différence entre un chien et un chat, car nous avons appris à les différencier. En réalité, notre cerveau a ordonné nos neurones de telle manière à ce que l'image que nos yeux captent active différemment notre réseau neuronal en fonction des lignes de l'animal observé.

Les premiers modèles de logique combinatoire basés sur le fonctionnement neuronal biologique ont été proposés entre 1930 et 1950. Les travaux de McCulloch et Pitts [70] sont notamment considérés comme fondateurs. Ils définissent en effet le *neurone formel*, première unité de calcul élémentaire basée sur le fonctionnement neuronal humain. Son fonctionnement est simple et similaire à celui du neurone humain. Il est composé d'un certain nombre d'entrées et d'une sortie, toutes binaires (voir figure 2.1).

L'analogie est assez directe. Nous pouvons comparer les dendrites d'un neurone biologique aux branches d'entrée d'un neurone formel. Le corps cellulaire et le nucleus sont comparables au centre de calcul du neurone artificiel schématisé par le cercle de la figure 2.1b. Enfin, la branche de sortie du neurone artificiel est comparable à l'axone. Bien que celle-ci puisse se diviser en différents terminaux, la sortie reste unique au même titre que celle du neurone biologique.

Le neurone formel renvoie 1 si la somme des signaux d'entrée dépasse son seuil d'activation. On peut donc parler ici d'une fonction d'activation de Heaviside bien que ce vocabulaire ne fût pas encore utilisé à cette époque. Cette notion sera approfondie dans la section 2.2.

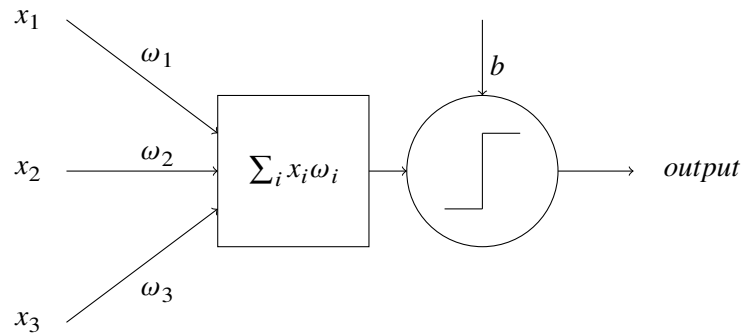


FIG. 2.2 : Le perceptron

## 2.2 Les réseaux de neurones artificiels modernes

### 2.2.1 Un premier modèle de neurone moderne «on-off» : le perceptron

Quinze ans après l'apparition du neurone formel, ROSENBLATT [72] a proposé un nouveau modèle de neurone appelé *Perceptron* (voir figure 2.2), beaucoup plus proche des outils utilisés aujourd'hui. Contrairement à son prédécesseur, les entrées de ce neurone ne sont plus binaires mais généralement comprises entre 0 et 1 (ou entre -1 et 1). La fonction d'activation du neurone est toujours un échelon, la sortie est donc quant à elle binaire. La seconde différence majeure avec le modèle original réside dans le fait que les branches reliant les entrées au neurone sont liées à des *poids* pouvant varier pendant la phase d'apprentissage (figure 2.2). Ils constitueront avec les *biais* un ensemble modifiable de grandeurs que nous appellerons *paramètres* dans la suite de ce manuscrit.

La sortie  $a$  du neurone peut ainsi se formuler en fonction des entrées  $x_i$ , des poids  $\omega_i$ , des biais  $b$  et de la fonction d'activation  $s$  du perceptron :

$$a = s(z) = s\left(\sum_i x_i \omega_i + b\right) \quad (2.1)$$

### 2.2.2 Grands principes heuristiques de construction

En se basant sur des observations sur le fonctionnement du cerveau humain, ROSENBLATT [72] a également proposé plusieurs grands principes toujours valables aujourd'hui, sur lesquels un réseau neuronal doit se fonder :

- Tout comme pour un système neuronal biologique, les connections physiques entre chaque neurone permettant l'apprentissage ou la reconnaissance ne sont pas identiques d'un réseau à l'autre. A la construction du réseau et au même titre qu'à la naissance d'un individu, ces connections sont aléatoires.
- Un réseau de neurones, qu'il soit artificiel ou non, doit être résilient et doté d'une certaine plasticité. Plus concrètement, cet axiome implique qu'un réseau de neurones doit pouvoir accepter et gérer des erreurs plus ou moins importantes dans les données utilisées.
- Deux stimuli similaires auront tendance à activer des chemins semblables à travers le réseau de neurones et donc à produire une réponse proche. C'est notamment sur ce principe fondamental

que se basent les capacités de généralisation des réseaux de neurones.

- L'application de certains *renforcements* peut faciliter ou pénaliser la formation de nouvelles connexions. Ce principe ainsi énuméré vise à conserver l'analogie avec le mécanisme biologique du cerveau. Dans la pratique, il est question ici de la phase d'apprentissage qui, grâce à des lois de mouvement des paramètres, a pour objectif de favoriser ou d'inhiber certaines branches du réseau par la modification de leurs poids.
- La notion de *similarité* des stimuli est dépendante du cas traité. Elle ne dépend pas nécessairement d'attributs particuliers qui pourraient être géométriques, fonctionnels ou existentiels mais est bien relative à l'environnement dans lequel le réseau de neurones évolue et est utilisé. Par exemple, si nous souhaitons grâce à un réseau neuronal différencier un chien et un chat, alors les images de deux chiens de races différentes vont avoir tendance à exciter les mêmes zones du réseau car la réponse souhaitée est la même : cet animal est un chien. Ces deux stimuli sont donc considérés comme semblables. En revanche, si on utilise à présent un réseau de neurones entraîné à distinguer les différentes races de chien, alors la réponse à ces deux stimuli sera très différente. Le contexte est donc primordial pour définir cette notion de similarité et dépend intrinsèquement de la manière dont le réseau neuronal perçoit le monde au sens étymologique du terme. Très concrètement, cet axiome implique une certaine rigueur à respecter lorsqu'on souhaite entraîner et utiliser un réseau de neurones. Il est en effet primordial de définir correctement son environnement. Les jeux de données utilisés pour la phase d'apprentissage doivent ainsi nécessairement être représentatifs de l'ensemble mathématique que l'outil est censé parcourir.

### 2.2.3 Les réseaux modernes : architecture et composition générale

Depuis le Perceptron, les neurones et leurs connexions se sont complexifiés de telle sorte qu'ils forment aujourd'hui une discipline propre et étendue. Cette section n'a pas vocation à être exhaustive mais permet de fixer un vocabulaire et les bases théoriques nécessaires à la compréhension globale des travaux présentés dans ce manuscrit.

**Réseau de neurones artificiel** On appelle *réseau de neurones* un outil mathématique composé de neurones reliés entre eux par des branches. Bien que les architectures possibles soient infinies, il existe plusieurs topologies de référence préférées suivant le cas à traiter. Dans ce manuscrit nous n'utiliserons que des réseaux de type *Feedforward*, abrégés par la suite par l'acronyme *FANN* (pour *Feedforward Artificial Neural Network*). Les réseaux présentés dans l'ensemble de nos travaux ne comporteront qu'une seule couche interne (*input layer* et *output layer* exclues donc) appelée couche cachée (figure 2.3).

Cette topologie est la plus simple possible mais cette famille regroupe aussi les réseaux de neurones profonds pouvant potentiellement contenir plusieurs couches cachées consécutives. On parle alors de *Deep Learning*. On dit parfois que chaque ajout de couche cachée implique un niveau d'abstraction supplémentaire. Bien que pouvant être très performants pour certaines tâches, ces réseaux à plusieurs couches peuvent devenir difficiles à entraîner par les méthodes d'apprentissage conventionnelles. Nous avons donc fait le choix ici d'opter pour une architecture plus simple à une seule couche, particulièrement adaptée au paradigme bayésien appliqué dans nos travaux (voir section 2.3.3).

D'un point de vue purement méthodologique, des preuves mathématiques ont également été apportées sur le caractère universel des réseaux de neurones à partir d'une couche cachée. En effet et sous certaines conditions sur les fonctions d'activation des neurones du réseau, ce type de FANN est capable d'approximer n'importe quelle fonction continue [66, 67]. Cette topologie est donc pertinente pour les

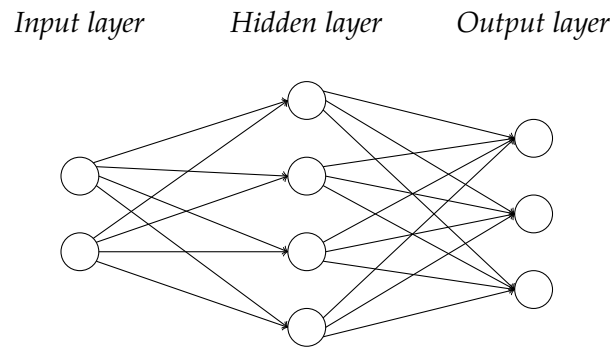


FIG. 2.3 : Un exemple de FANN à une couche cachée

travaux de régression. Notons néanmoins que théoriquement, cette preuve n'est valable que si aucune restriction n'est imposée sur le nombre de neurones dans la couche cachée.

Pour des tâches plus spécifiques comme la reconnaissance d'image par exemple, on utilise plutôt aujourd'hui des réseaux profonds convolutifs (*CNN*), alternant des couches particulières aux fonctions différentes (couche de *convolution*, de *pooling*, de *correction* et *fully-connected*). Les nombreuses couches de ce type d'architecture permettent d'identifier les tendances ou les *similarities* entre chaque image pour qu'elle soit ensuite interprétée et classifiée. Une description complète de ce type de réseau a été récemment proposée par LI et al. [73].

Enfin, nous pouvons également citer les réseaux bouclés ou récurrents qui se distinguent fondamentalement des *FANN* par le fait que l'information ne transite plus dans un seul sens mais peut potentiellement remonter les couches grâce aux boucles créées par les branches de ce type de réseaux. Cette architecture est notamment utilisée pour l'analyse de séries temporelles, de reconnaissance vocale ou de traduction manuscrite par exemple. Ces réseaux présentent un effet de mémoire intéressant permettant de prédire une sortie qui ne serait plus uniquement fonction d'un jeu de données à un instant spécifique mais également d'entrées antérieures. Cette fonction complètement hors de portée d'un *FANN* classique est donc particulièrement intéressante pour la traduction linguistique par exemple, dont le sens dépend bien souvent d'un contexte plus général donné par les phrases ou les mots précédents. Une description plus exhaustive de ce type de réseau est disponible dans les travaux de LIPTON, BERKOWITZ et ELKAN [74].

**Les fonctions d'activation** Cette notion a déjà été évoquée précédemment. Elle désigne la fonction mathématique associée à chaque neurone. La somme pondérée  $z$  des entrées  $x_i$  par leurs poids respectifs  $\omega_i$  subit une transformation par cette fonction. Les neurones formels et les perceptrons ont tous les deux une fonction d'activation de Heaviside qui implique donc une sortie binaire. Les réseaux de neurones modernes utilisent aujourd'hui des fonctions d'activation continues qui permettent d'obtenir une sortie quelconque comprise entre les bornes de la fonction d'activation.

Dans le cas des *FANN*, il a été démontré que le caractère universel des approximateurs de type réseaux de neurones était assuré si les fonctions d'activation utilisées étaient continues, bornées et non-constantes [67]. Il est également nécessaire que les fonctions choisies soient dérivables pour pouvoir appliquer l'algorithme d'apprentissage par descente de gradient (voir section 2.2.4). Pour les travaux de régression, les fonctions les plus utilisées sont de la famille des sigmoïdes  $s_{sig}(x)$  ou des variantes proches comme la tangente hyperbolique  $s_{th}$  (figure 2.4) :

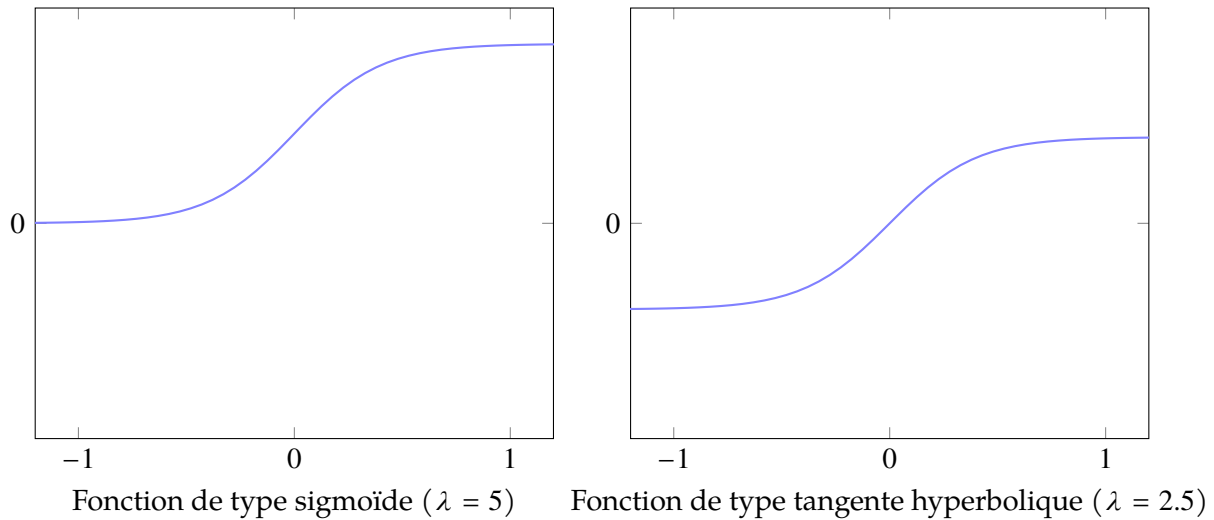


FIG. 2.4 : Exemples de fonctions d'activation utilisées pour la regression

$$s_{sig}(x) = \frac{1}{1 + e^{\lambda x}} \quad (2.2)$$

$$s_{th}(x) = \frac{e^{\lambda x} - e^{-\lambda x}}{e^{\lambda x} + e^{-\lambda x}} \quad (2.3)$$

Lorsque l'on souhaite faire de la classification ou de la reconnaissance d'image par exemple, les fonctions d'activation sont plutôt de type *ReLU* ou *ELU* pour les couches cachées et de type *SoftMax* pour la couche de sortie. Bien qu'elles ne satisfassent pas aux critères pour garantir l'universalité, leur efficacité n'est plus à démontrer. Les travaux décrits dans la suite de ce manuscrit ne traitant pas de ce type de tâches, nous n'explicitons pas plus ces fonctions. Notons également que cet inventaire n'est pas exhaustif. La liste détaillée des fonctions utilisables pour l'entraînement des réseaux de neurones est disponible dans les travaux de APICELLA et al. [75].

**Base de données et importance de la normalisation** Afin de pouvoir entraîner un réseau de neurones, il est nécessaire de pouvoir s'appuyer sur une base de données. Elle est indispensable à tout apprentissage supervisé et est constituée de jeux de données d'entrée avec leurs sorties correspondantes. Cet ensemble constitue alors la base théorique sur laquelle s'appuie le réseau pour apprendre le comportement qu'il doit copier. Ces données serviront notamment dans l'algorithme par descente de gradient à calculer les erreurs effectuées par le réseau entre la sortie calculée et la sortie théorique directement issue de la base.

Dans la suite de ce manuscrit, les entrées et les sorties théoriques, issues de la base de donnée seront respectivement notées  $x_i$  et  $y_i$  où  $i$  correspond au  $i^{\text{ème}}$  jeu de données composant la base. Elles correspondent toutes les deux potentiellement à des vecteurs dans la mesure où les entrées et les sorties peuvent être multiples.

i

Par convention et dans l'ensemble de ce manuscrit, les vecteurs seront notés en lettres minuscules et gras comme  $x_i$  et  $y_i$ . Nous les différencierons ainsi des matrices (en lettres majuscules et en gras) et des variables (en lettres minuscules simples). Ces notations seront à nouveau précisées plus loin dans ce chapitre.

Il est difficile de donner un ordre de grandeur de la taille nécessaire minimale ou idéale pour une base de données car elle est très dépendante du problème à traiter. Le plus important est qu'elle soit représentative de ce que l'on souhaite modéliser, notamment pour des tâches de régression. Cela dit, les réseaux de neurones peuvent également être des outils plus exploratoires pour mettre en exergue des tendances ou des liens entre données qu'on ne soupçonnerait pas. Dans l'industrie aéronautique par exemple, ce type d'outils est très utilisé pour déterminer les plans de maintenance préventive des systèmes propulsifs. Les bases sont constituées d'un très grand nombre de données de vol ou d'essais sans qu'aucun lien *a priori* évident n'existe. L'objectif de ces réseaux est de déterminer les meilleures dates ou actions de maintenance avant incident. Cet exemple nous permet de préciser que le terme générique d'Intelligence Artificielle est sans doute incorrect en toute rigueur. Les réseaux de neurones tiennent plus des statistiques avancées. Ils se basent sur des liens existants et implicites entre un grand nombre de jeux de données et en recherchent des tendances ou des corrélations complexes.

Ajoutons également que de manière générale, plus une base de données est grande, plus l'apprentissage sera complet. En revanche, le temps nécessaire à l'entraînement sera plus élevé. En fonction de la taille de cette base de données, nous adapterons également l'algorithme d'entraînement choisi. Certains sont plus adaptés aux très grandes bases de données tandis que d'autres seront bien plus performants sur des petits jeux bien choisis. Nous n'explicitons pas en détail dans ce manuscrit les différents algorithmes existants. Nous nous focaliserons sur les bases et principes des méthodes d'entraînement par descente de gradient (voir section 2.2.4) et sa variante dans le paradigme bayésien (voir les sections 2.3.2 et 2.3.3).

Enfin, nous insistons sur l'importance de normaliser les données utilisées par le réseau de neurones. En effet, une différence d'échelle entre les données d'entrée peut impliquer des poids très importants sur certaines branches. L'algorithme par descente de gradient pourra être alors beaucoup plus long pour aboutir voire ne jamais réussir à sortir d'un extremum local. Enfin, des valeurs absolues trop grandes en entrée risquent de saturer la majorité des neurones aux bornes des fonctions d'activation. Ainsi, des grandes différences de valeurs en entrée pourront n'impliquer qu'une très faible variation de la sortie. Le réseau ne pourra alors rien apprendre correctement. De la même façon, il est important de normaliser la sortie ou bien d'ajouter une couche dédiée à cette opération utilisant une fonction d'activation non-bornée (type *ReLU*) ou utilisant des bornes contenant les extrema des valeurs possibles en sortie.

Bien que cette étape soit parfois négligée, nous recommandons fortement de systématiser ce protocole. Plusieurs normalisations peuvent être utilisées telle que :

- La normalisation *Min/Max* :

$$x_{norm} = \frac{(x - x_{min})(b_2 - b_1)}{x_{max} - x_{min}} \quad (2.4)$$

Avec  $b_1$  et  $b_2$ , les valeurs des bornes de normalisation.

- La normalisation centrée réduite :

$$x_{norm} = \frac{x - \bar{x}}{\sigma} \quad (2.5)$$

Avec  $\bar{x}$  et  $\sigma$ , la moyenne et l'écart type de la série de valeurs. Avec cette technique, les valeurs minimale et maximale de la série normalisée peuvent être inférieures (resp. supérieures) à -1 (resp.

1). Il convient donc d'utiliser des fonctions d'activation avec des bornes contenant les nouvelles valeurs normalisées. Il est par exemple possible de modifier la valeur de  $\lambda$  (voir équations (2.2) et (2.3)) dans le cas où la fonction d'activation est une sigmoïde ou une tangente hyperbolique.

La liste n'est, encore une fois, pas exhaustive. Toute autre méthode permettant de respecter les principes précédemment cités permettra globalement d'améliorer l'efficacité de l'apprentissage. La normalisation fait ainsi assurément partie des bonnes pratiques à adopter dans la mise en œuvre de ce type d'outils.

**Paramètres et notations** Au-delà du vocabulaire indispensable, ce paragraphe va nous permettre de définir une notation particulière pour l'ensemble de ce manuscrit. Tout d'abord, on appelle *paramètres* l'ensemble des grandeurs modifiables et utilisées pour la production d'une réponse par le réseau de neurones. On regroupe ainsi sous ce terme les poids  $\omega$  et les biais  $b$ . L'ensemble de ces paramètres se note sous la forme d'un vecteur  $\mathbf{v}$ .

#### Identification des poids et des biais

Les poids  $\omega$  et les biais  $b$  sont identifiés par leur position dans le réseau. On note par exemple  $\omega_{jk}^l$  le poids de la branche reliant le neurone  $k$  de la couche  $l - 1$  au neurone  $j$  de la couche  $l$ . Les biais étant eux assignés à des neurones, nous noterons  $b_j^l$  le biais du neurone  $j$  de la couche  $l$  (voir figure 2.5).

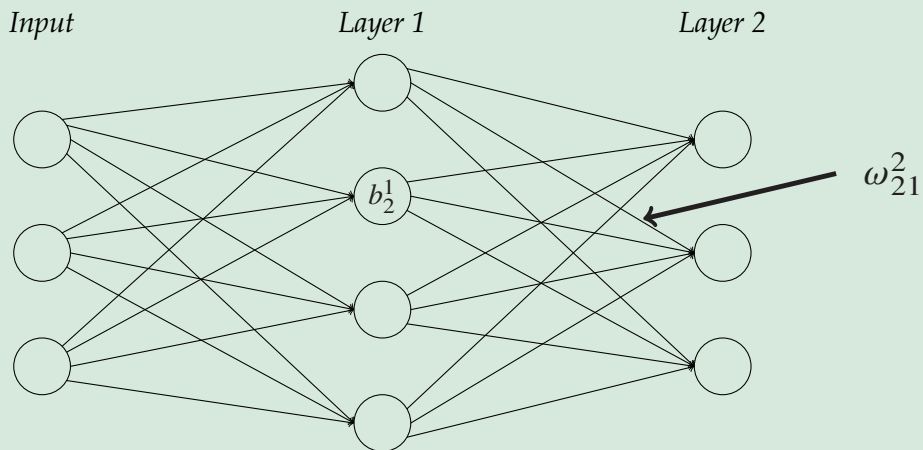


FIG. 2.5 : Notations des poids et des biais

Si nous reprenons l'exemple de la figure 2.5, nous pouvons alors écrire :

$$\mathbf{v} = \{\omega_{11}^1, \omega_{12}^1, \dots, b_1^1, \omega_{21}^1, \omega_{22}^1, \dots, b_2^1, \omega_{31}^1, \dots, \omega_{3,3}^2, \omega_{3,4}^2, b_3^2\} \quad (2.6)$$

Cette notation nous permettra alors d'exprimer les paramètres du réseau sous différentes formes de manière indifférenciée. Nous pourrions en effet utiliser leur rang dans le vecteur  $\mathbf{v}$  autant que la notation explicite décrite dans le cadre ci-dessus. Nous avons donc par exemple si  $i = 5$  :

$$\mathbf{v}(i) = v_i = \omega_{21}^1 \quad (2.7)$$

**Les Hyperparamètres** On appelle *hyperparamètres* toutes les autres grandeurs du réseau de neurones relatives à la mise en données ou au déroulement de l'algorithme d'entraînement. Nous pouvons par exemple citer :

- Le nombre de couches et le nombre de neurones par couche
- Le taux d'apprentissage (ou *learning rate*)  $\eta$  (voir section 2.2.4)
- Les coefficients de régularisation  $\alpha$  et  $\beta$  (voir sections 2.2.6 et 2.3.3)
- Le nombre d'époques
- Le nombre de jeux de données d'entraînement, de test et de validation
- ...

Nous reviendrons plus en détail sur chacun d'entre eux dans la suite de ce chapitre. Notons néanmoins que ces hyperparamètres revêtent une grande importance dans le processus d'entraînement car ils peuvent grandement favoriser l'apprentissage s'ils sont correctement paramétrés tout comme ils peuvent le rendre totalement inefficace. Malheureusement, ils sont dans la majorité des cas difficiles à définir *a priori* car ils dépendent très fortement du cas d'application. Bien que des ordres de grandeur puissent être donnés pour certains d'entre eux, un paramétrage efficace est souvent issu d'une série d'essais infructueux. Cette caractéristique limite fortement l'utilisation de ce type d'outils pour l'interpolation de fonctions complexes et le développement de modèles réduits à l'échelle industrielle.

**Fonction de coût** On entend par apprentissage la phase où l'on modifie les paramètres du réseau afin d'obtenir des réponses les plus proches possibles de celles attendues. Il est donc nécessaire de mesurer cet écart par une grandeur observable. Cette évaluation se fait par une fonction de coût notée  $C$  pouvant potentiellement prendre plusieurs formes mais devant être dérivable.

Au-delà de la performance du réseau, cette fonction joue également un rôle fondamental dans le processus d'apprentissage. Un parallèle avec la notation d'une copie d'un élève peut par exemple être fait. La correction d'un devoir permet en principe à la fois au professeur d'évaluer l'élève autant que de lui montrer où se trouvent ses erreurs pour qu'il puisse les corriger. Ici, les erreurs calculées par la fonction de coût seront ensuite utilisées par l'algorithme d'entraînement afin de modifier de manière efficace et ciblée, les paramètres  $\mathbf{v}$  du réseau.

Pour les travaux de régression, la fonction de coût la plus utilisée est la MSE pour *Mean Square Error*. Elle correspond à la moyenne quadratique des écarts entre les sorties théoriques  $\mathbf{y}_x$  et celles produites par le réseau  $\mathbf{a}_x^L$ . On distingue  $C_x$  correspondant à la MSE calculée sur un jeu de données  $x$  et  $C$  calculée sur l'ensemble de la base :

$$C_x = \frac{1}{2} \|\mathbf{y}_x - \mathbf{a}_x^L\|^2 = \frac{1}{2} \sum_j (y_{x,j} - a_{x,j}^L)^2 \quad (2.8)$$

$$C = \frac{1}{2N} \sum_x \|\mathbf{y}_x - \mathbf{a}_x^L\|^2 = \frac{1}{2N} \sum_x \sum_j (y_{x,j} - a_{x,j}^L)^2 \quad (2.9)$$

Avec  $j$ , le  $j^{\text{ème}}$  neurone de la couche de sortie  $L$  et  $N$  le nombre total de jeux de données d'entraînement. On remarque d'après les équations (2.8) et (2.9) que la fonction de coût globale, calculée sur la totalité de la base de données est finalement égale à la moyenne des contributions individuelles de chaque jeu de données, d'où :

$$C = \frac{1}{N} \sum_x C_x \quad (2.10)$$



i

Pour les tâches de régression et en particulier pour celles présentées dans le reste de ce manuscrit, la sortie de l'ANN est unique. Par conséquent, les vecteurs  $\mathbf{y}_i$  et  $\mathbf{a}_i^L$  sont assimilables à des scalaires. L'indice  $j$  utilisé ci-dessus ne sera donc plus nécessaire, au même titre que la notation vectorielle. Nous avons donc :

$$\begin{aligned} \mathbf{y}_x &\sim y_x \\ \mathbf{a}_x^L &\sim a_x^L \end{aligned}$$

Les équations (2.8) et (2.9) peuvent également se simplifier :

$$\begin{aligned} C_x &= \frac{1}{2} (y_x - a_x^L)^2 \\ C &= \frac{1}{2N} \|\mathbf{y} - \mathbf{a}\|^2 = \frac{1}{2N} \sum_x (y_x - a_x^L)^2 \end{aligned}$$

Enfin, nous notons  $\mathbf{e} = \{e_x\}_{x=1:N}$  le vecteur contenant l'ensemble des erreurs pour chaque jeu de données d'entraînement. Encore une fois, du fait de la sortie unique du réseau de neurones,  $e_x$  est ici un scalaire et est le résultat de la différence  $y_x - a_x^L$  pour le jeu de données  $x$ .

## 2.2.4 Algorithme par descente de gradient : description, intuïfs et explication d'une méthode d'entraînement incontournable

Il n'est pas envisageable de parler de méthode d'entraînement d'un réseau de neurones sans évoquer la descente de gradient. Cet algorithme est en effet un incontournable tant il est utilisé et efficace. Notons également que cette méthode ne se limite pas à l'apprentissage supervisé. C'est en effet globalement une technique d'optimisation de paramètres permettant de trouver le minimum de n'importe quelle fonction convexe de manière itérative.

Si sa formulation de base est aujourd'hui obsolète, la quasi-totalité des algorithmes d'entraînement modernes n'en sont que des variantes. Ce paragraphe a pour but de comprendre le principe de fonctionnement finalement assez intuitif de cette méthode. Nous présenterons également les quatre équations fondamentales sur lesquels la descente de gradient se base lorsqu'elle est appliquée à l'apprentissage supervisé.

i

Dans ce paragraphe, nous supposons que le jeu de données d'entraînement  $x$  est fixé. Pour alléger la notation, nous pourrions ainsi écrire  $C = C_x$ .

Souvent également appelé par son nom anglophone, la *Backpropagation* a été introduite par RUMELHART, HINTON et WILLIAMS [76] dans l'un des plus célèbres papiers de la communauté traitant des techniques de Machine Learning. Dans ce paragraphe, nous nous basons principalement sur les travaux de vulgarisation et une partie des analogies proposées par NIELSEN [77].

*i*

Afin de faciliter la compréhension générale de la méthode, l'ensemble de ce paragraphe est basé sur l'étude d'un réseau de neurones de type *Feedforward* (voir section 2.2.3) à une ou plusieurs couches cachées. La fonction de coût considérée est la *MSE* et les fonctions d'activation de chaque neurone sont des tangentes hyperboliques.

**Principe de l'algorithme par descente de gradient appliquée à l'apprentissage** L'apprentissage est une phase qui consiste à modifier les paramètres du réseau de telle manière à minimiser la fonction de coût. En d'autres termes, nous cherchons à optimiser le vecteur  $\mathbf{v}$  à chaque itération de l'algorithme, appelée *époque*. Nous supposons dans ce paragraphe que  $\mathbf{v}$  est composé de deux scalaires telles que  $\mathbf{v} = \{v_1, v_2\}$  et que la sortie du réseau de neurones est unique, c'est à dire que  $\mathbf{a}^L = \{a^L\}$ . Chercher le minimum de  $C$ , fonction des deux variables  $\{v_1, v_2\}$  revient donc finalement à parcourir la surface construite à partir de l'ensemble des couples possibles pour atteindre *le point le plus bas*. Pour des fonctions dépendant de plus de variables, le principe reste le même au détail près que l'on parcourt alors une hypersurface comportant autant de dimensions que de paramètres à optimiser.

Au début de l'algorithme, lors de l'initialisation aléatoire des paramètres, nous pouvons imaginer que nous lançons aléatoirement une balle sur une surface vallonnée. A cet endroit précis, nous pouvons alors connaître la forme de la surface et donc, vers quelle direction la balle aura tendance à rouler. L'ingénieur, afin de déterminer la loi exacte régissant le mouvement de la balle, écrirait les équations de Newton en prenant en compte la pente de la surface, correspondant à la dérivée de  $C$  dans l'ensemble de ses *directions*  $v_1$  et  $v_2$ , le coefficient de frottement de la balle avec le sol, etc. jusqu'à ce que la balle atteigne le point le plus bas de la surface.

Cette loi de mouvement de la balle, basée sur la physique, est précisément le comportement que nous recherchons. L'idée de l'algorithme par descente de gradient est donc bien de construire une loi de mouvement afin que la balle, correspondant dans cette analogie à la valeur de la fonction de coût, descende sans cesse.

Nous notons  $\Delta C$ , la variation de la fonction  $C(v_1, v_2)$  telle que :

$$\Delta C \approx \frac{\partial C}{\partial v_1} \Delta v_1 + \frac{\partial C}{\partial v_2} \Delta v_2 \quad (2.11)$$

Faire en sorte que la balle *descende* le long de la surface à chaque itération revient à dire que la variation  $\Delta C$  est toujours négative. Le vecteur variation  $\Delta \mathbf{v} = (\Delta v_1, \Delta v_2)^T$  des paramètres doit donc être choisi de telle manière à respecter ce principe. On définit également le gradient de la fonction de coût telle que :

$$\overrightarrow{\text{grad}} C = \nabla C = \left( \frac{\partial C}{\partial v_1}, \frac{\partial C}{\partial v_2} \right)^T \quad (2.12)$$

D'après équations (2.11) et (2.12), nous pouvons donc écrire :

$$\Delta C \approx \nabla C \cdot \Delta \mathbf{v} \quad (2.13)$$

Afin de garantir le signe négatif que doit porter  $\Delta C$ , il nous suffit alors de définir une loi de variation des paramètres telle que :

$$\Delta \mathbf{v} = -\eta \nabla C \quad (2.14)$$

On a ainsi pour chaque paramètre :

$$v_i \rightarrow v'_i = v_i - \eta \frac{\partial C}{\partial v_i} \quad (2.15)$$

De cette façon, l'équation (2.13) devient alors :

$$\Delta C \approx -\eta \nabla C \cdot \nabla C = -\eta \|\nabla C\|^2 \quad (2.16)$$

L'hyperparamètre  $\eta$  est appelé *taux d'apprentissage* (ou *learning rate*) et est strictement positif. Avec cette définition, la variation de la fonction de coût au fur et à mesure de l'entraînement ne peut être que négative, autrement dit, les erreurs produites par le réseau de neurones ne peuvent que diminuer en moyenne. Cela signifie que le réseau de neurones apprend à chaque époque pour se rapprocher d'un minimum global vers lequel il tend asymptotiquement. Pour reprendre l'analogie de la balle parcourant une surface, le taux d'apprentissage pourrait être assimilé à la force avec laquelle la balle est poussée à chaque itération.

i

Le taux d'apprentissage  $\eta$  est un hyperparamètre très important dans tout algorithme par descente de gradient. Il agit de la même manière qu'une boîte de vitesse pour l'algorithme. En effet, il guide la *vitesse* d'apprentissage en ajoutant un facteur multiplicatif à la norme du gradient. Lorsque  $\eta$  est grand (de l'ordre de l'unité), la variation des paramètres du réseau est grande et le réseau apprend donc rapidement (figure 2.6, gauche). C'est le cas de figure qui se pose en début d'apprentissage, lorsque la réponse du réseau est très éloignée de la solution finale. Au cours de l'apprentissage, sa valeur diminue afin que les paramètres puissent s'ajuster plus finement à mesure que les erreurs faites par le réseau diminuent (figure 2.6, droite).

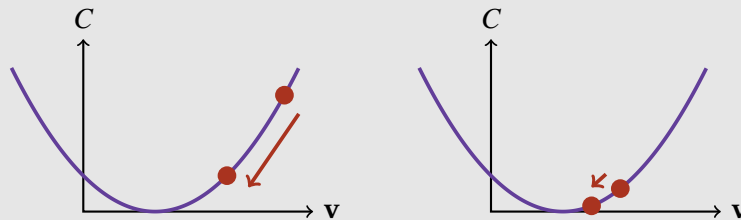


FIG. 2.6 : Évolution recherchée de  $\eta$

Dans la pratique, il n'est pas toujours simple de définir une loi de mouvement de ce paramètre  $\eta$  car sa valeur dépend beaucoup du cas d'étude. En général, il est nécessaire de tester plusieurs configurations par *essais/erreurs* pour déterminer l'évolution de cet hyperparamètre. En effet, si le taux d'apprentissage est trop faible au début de l'algorithme (figure 2.7, gauche), l'apprentissage peut être très lent voire ne pas avancer du tout. À l'inverse, s'il est trop grand lorsque le réseau commence à converger (figure 2.7, droite), alors la solution peut potentiellement osciller autour de l'optimum sans jamais l'atteindre.

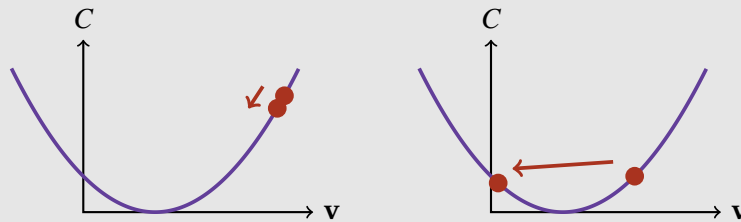


FIG. 2.7 : Cas problématiques en début (gauche) ou fin d'algorithme (droite)

Enfin, il est parfois possible que le processus d'apprentissage soit bloqué dans un minimum local (figure 2.8, gauche). Cette situation peut devenir très problématique et est probablement la difficulté majeure à surmonter avec ce type d'algorithme. Dans ce cas, une valeur de  $\eta$  ponctuellement élevée peut potentiellement permettre de sortir de ce minimum local pour continuer ensuite l'entraînement (figure 2.8, droite). Toute la difficulté réside dans le fait d'identifier et de différencier une solution optimale locale et globale.

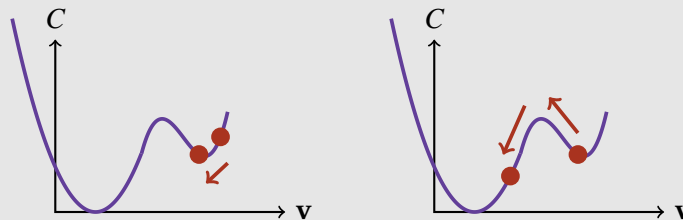


FIG. 2.8 : Présence d'un minimum local

Afin de pouvoir modifier chaque paramètre  $v_i$  dans le réseau, il est nécessaire de connaître la valeur de chaque dérivée partielle  $\frac{\partial C}{\partial v_i}$  dans toutes les couches du réseau. Pour ce faire, l'algorithme par descente de gradient se base sur quatre équations fondamentales permettant de *rétro-propager* les erreurs calculées en sortie de neurone vers les couches supérieures. Nous définissons pour cela  $\delta_j^l$ , la sensibilité au neurone  $j$  de la couche  $l$  telle que :

$$\delta_j^l = \frac{\partial C}{\partial z_j^l} \quad (2.17)$$

Avec  $z_j^l$ , la sortie du neurone  $j$  avant transformation par la fonction d'activation (voir équation (2.1)).

**Déroulement de l'algorithme** La descente de gradient fonctionne en deux temps. On calcule tout d'abord la sortie du réseau de neurones  $\mathbf{a}^L$  au début de l'époque considérée. On appelle cette étape la *Feedforward Pass*. Elle permet notamment de calculer toutes les grandeurs intermédiaires comme les sorties de chaque neurone avant transformation  $\mathbf{z}_j^l$ , après transformation  $\mathbf{a}_j^l$  et le vecteur erreur  $\mathbf{e}$  (voir figure 2.9).

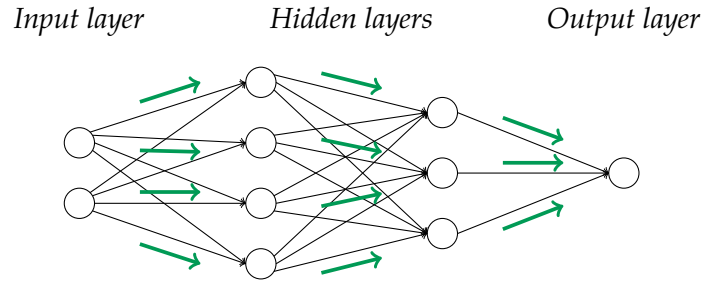


FIG. 2.9 : *Feedforward Pass* : calcul des sorties de chaque neurone et des erreurs

Nous pouvons alors écrire la première équation fondamentale de l'algorithme, permettant de calculer les sensibilités de chaque neurone sur la dernière couche :

$$\delta_j^L = \frac{\partial C}{\partial a_j^L} s'(z_j^L) \quad (2.18)$$

Dans cette équation, tout est connu ou facilement calculable grâce à la *Feedforward pass*. La dérivée partielle  $\frac{\partial C}{\partial a_j^L}$  dépend évidemment de la fonction de coût choisie mais dans notre cas, nous pouvons simplement écrire :

$$\frac{\partial C}{\partial a_j^L} = \frac{\partial \left[ \frac{1}{2} \sum_j (a_j^L - y_j)^2 \right]}{\partial a_j^L} = (a_j^L - y_j) \quad (2.19)$$

D'après les équations (2.18) et (2.19), nous pouvons également réécrire la première équation de l'algorithme sous forme matricielle telle que :

$$\boldsymbol{\delta}^L = (\mathbf{a}^L - \mathbf{y}) \odot s'(\mathbf{z}^L) \quad (2.20)$$

**Produit de Hadamard**

L'opérateur  $\odot$  correspond à un produit de Hadamard prenant deux matrices de même dimension et produisant une nouvelle matrice dont les coefficients sont les produits termes à termes des deux matrices en argument. Nous avons par exemple :

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \odot \begin{bmatrix} c_1 & d_1 \\ c_2 & c_2 \end{bmatrix} = \begin{bmatrix} a_1c_1 & b_1d_1 \\ a_2c_2 & b_2d_2 \end{bmatrix} \quad (2.21)$$

La seconde équation fondamentale de l'algorithme permet de *rétro-propager* les sensibilités calculées par l'équation (2.20) dans les autres couches du réseau de neurone. C'est ce qu'on appelle la *Backward Pass*.

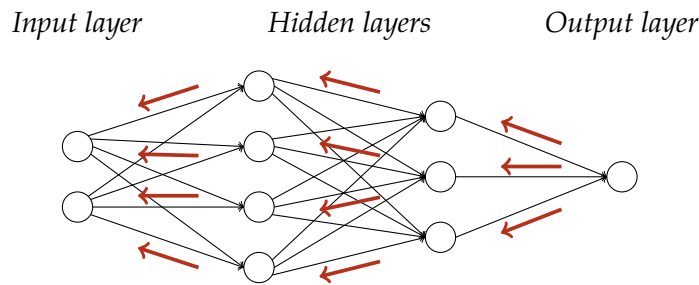


FIG. 2.10 : *Backward Pass* : Rétro-propagation des sensibilités

C'est en fait une relation de récurrence que l'on peut écrire sous forme matricielle telle que :

$$\delta^l = [(\omega^{l+1})^T \delta^{l+1}] \odot s'(z^l) \quad (2.22)$$

Là encore et après application de l'équation (2.20), chaque terme est connu. Le calcul des sensibilités de chaque neurone est donc facilement réalisable de proche en proche.

La troisième équation permet de calculer la dérivée partielle de la fonction de coût par rapport aux biais  $b_j^l$  :

$$\frac{\partial C}{\partial b_j^l} = \delta_j^l \quad (2.23)$$

En intégrant l'équation (2.23) à l'équation (2.15), nous pouvons ainsi calculer simplement les nouvelles valeurs des biais  $b_j^l$  à partir de la loi de mouvement suivante :

$$b_{j,new}^l = b_j^l - \eta \frac{\partial C}{\partial b_j^l} = b_j^l - \eta \delta_j^l \quad (2.24)$$

Enfin, la quatrième et dernière équation permet de calculer la dérivée partielle de la fonction de coût par rapport aux poids dans le réseau :

$$\frac{\partial C}{\partial \omega_{jk}^l} = a_k^{l-1} \delta_j^l \quad (2.25)$$

De la même manière que précédemment, en intégrant l'équation (2.25) à l'équation (2.15), nous obtenons la loi de mouvement de chaque poids du réseau de neurones :

$$\omega_{jk,new}^l = \omega_{jk}^l - \eta \frac{\partial C}{\partial \omega_{jk}^l} = \omega_{jk}^l - \eta a_k^{l-1} \delta_j^l \quad (2.26)$$

Finalement, ces quatre équations fondamentales suffisent à entraîner un réseau de neurones par *BackPropagation*, tout du moins en principe. Dans la pratique, ces équations de bases ne sont pas utilisées en l'état. On utilise plutôt des variantes plus efficaces et avec une vitesse de convergence plus importante mais l'idée et le déroulement de l'algorithme restent globalement toujours les mêmes. Ces équations et ce cas simplifié permettent ainsi de bien comprendre la procédure mise en place.

i

Pour appliquer les équations précédentes, nous devons connaître la valeur des paramètres à l'époque  $n - 1$ . Si cette condition est évidemment toujours remplie au cours de l'algorithme, il est nécessaire de les initialiser lors du lancement du calcul. Théoriquement, cette étape n'est pas censée influencer sur le résultat final car l'algorithme doit converger vers un optimum commun quel que soit le point de départ. Néanmoins, la manière d'initialiser les paramètres du réseau peut influencer, parfois significativement, sur la vitesse de convergence. Pour reprendre l'analogie utilisée dans ce paragraphe, il est assez aisé de comprendre que plus le point de départ de la balle dans la vallée est éloigné du point le plus bas, plus la balle mettra de temps à atteindre son objectif. L'ordre de grandeur des paramètres d'un réseau de neurones est généralement compris entre 0.01 et 10 en valeur absolue. Ces valeurs sont données à titre indicatif mais ne constituent pas une règle incontournable. Une manière simple et efficace consiste à tirer la valeur de chaque paramètre selon une loi normale centrée en 0 et d'écart type égal à 1 [69]. D'autres techniques plus poussées existent mais sont généralement destinées à des topologies plus complexes ou à des fonctions d'activation particulières comme la *ReLU* par exemple [78, 79].

## 2.2.5 Le sur-apprentissage : le mal de l'algorithme par descente de gradient

Bien que généralement très efficace, la *BackPropagation* peut être sujette à plusieurs difficultés. Nous avons déjà évoqué précédemment le risque de rester bloqué dans un minimum local pendant l'entraînement mais il existe également une autre limite importante à ce type d'algorithme. Le sur-apprentissage ou *overfitting* survient généralement lorsqu'on ne l'attend pas dans le sens où il peut être difficile à identifier.

En effet, les critères de qualité classiquement utilisés comme la valeur de la fonction de coût ou la *RMSE* par exemple sont toujours bons et continuent à décroître en cas d'*overfitting*. Ce phénomène semble donc parfois paradoxal. Qualitativement, nous sommes en situation de sur-apprentissage lorsque le réseau ne parvient plus à proposer une solution générale au problème traité. Très concrètement et bien que le réseau copie de mieux en mieux ses données d'entraînement (figure 2.12, points rouges), son comportement sur des jeux inconnus devient aberrant (figure 2.12, points verts).

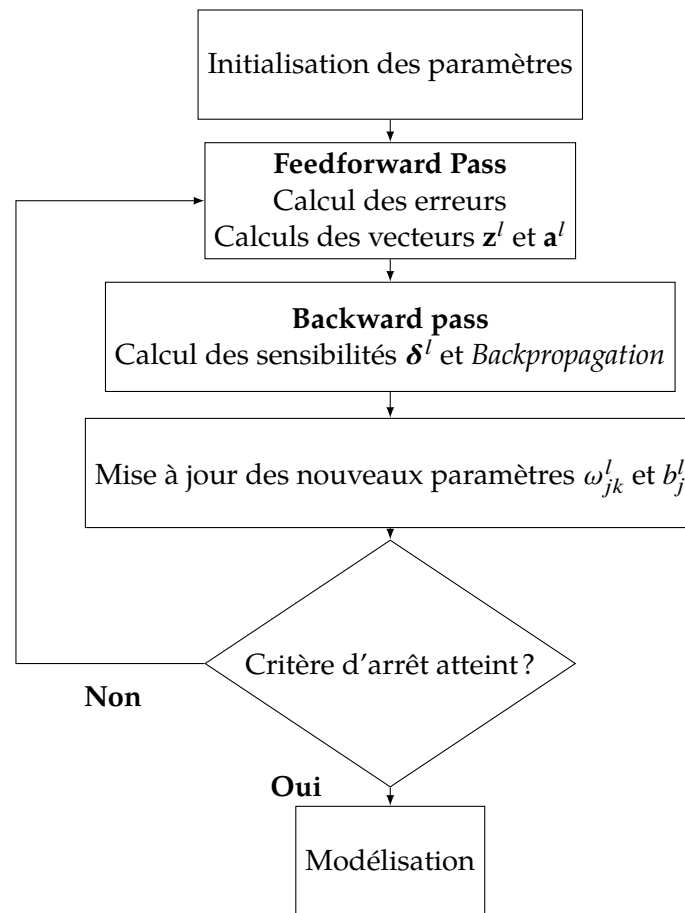


FIG. 2.11 : Logigramme de l'algorithme d'entraînement par descente de gradient

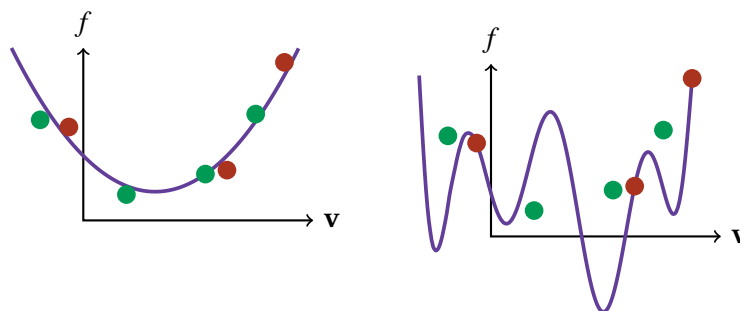


FIG. 2.12 : Exemple de sortie d'un réseau de neurone correctement entraîné (gauche) et en sur-apprentissage (droite). Les points rouges représentent la réponse du réseau à des données d'entraînement et les points verts, celle à des données inconnues

Quantitativement, ce problème intervient lorsque certains poids du réseau sont trop grands. Dans ce cas de figure, une faible variation des entrées peut impliquer une grande variation des sorties. Bien qu'exceptionnellement, ce genre de comportement puisse être désiré, nous recherchons en général deux réponses similaires pour deux jeux d'entrée similaires. On dit également que le réseau *apprend le bruit* dans le cas où par exemple, les données d'entraînement proviennent d'expériences réelles (enregistrement vocaux, photographies, etc.) ou de simulations numériques.



Il existe plusieurs stratégies pour limiter les risques de sur-apprentissage [80, 81]. Nous n'en ferons pas une liste exhaustive ici mais nous pouvons malgré tout citer quelques principes pertinents. Premièrement, la méthode qui semble la plus simple mais qui n'est pas toujours possible est d'ajouter des points d'entraînement à la base de données. En effet, plus celle-ci sera grande et représentative du problème à modéliser, moins les risques de résultats aberrants sur des cas inconnus seront grands. Néanmoins, une base de données trop importante peut mener à des temps d'apprentissage assez élevés. Il est également possible et conseillé de ne pas trop *pousser* l'entraînement, c'est à dire de limiter le nombre d'époques au strict nécessaire [82]. Dans le cas des réseaux de neurones, l'adage « Le mieux est l'ennemi du bien » est en ce sens particulièrement pertinent. Afin de contrôler ce phénomène, une partie de la base de données d'entraînement est extraite avant l'apprentissage et utilisée pour valider les prédictions du réseau. Si l'erreur sur les données d'entraînement diminue mais que celle sur cette base de validation augmente, alors il est nécessaire de stopper l'apprentissage. Cette technique se nomme *Early stopping* (voir figure 2.13).

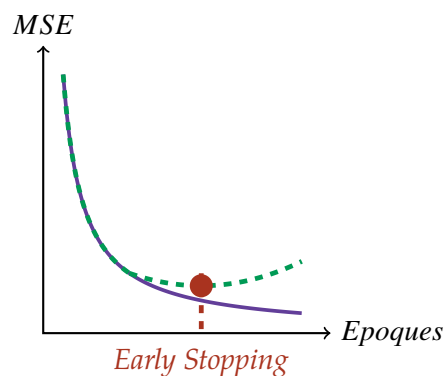


FIG. 2.13 : Représentation de la MSE sur les points d'entraînement (bleu) et sur les points de validation (vert) en cas d'overfitting

La topologie d'un réseau peut également avoir une grande importance dans sa tendance au sur-apprentissage. Le nombre de neurones est représentatif du nombre de degrés de liberté de l'outil. En ce sens, il serait donc assez naturel en tant qu'utilisateur de vouloir surestimer le nombre de neurones nécessaires. L'idée derrière cette tendance est de se dire que plus le réseau sera libre, mieux il pourra s'adapter à la fonction à modéliser. Cependant, au delà de l'augmentation du temps d'entraînement, un réseau sur-dimensionné aura une forte tendance au sur-apprentissage justement car il lui sera *plus facile* d'aller parfaitement copier les points d'entraînement potentiellement bruités au détriment de sa capacité de généralisation. Le choix de la topologie est donc une vraie question et nécessite souvent plusieurs essais avant de trouver une combinaison adaptée. Il existe des méthodes d'optimisation permettant de déterminer *la meilleure topologie* basées par exemple sur les algorithmes génétiques [83]. Elles nécessitent néanmoins souvent de très importantes capacités de calcul pour un gain relativement faible, tout du moins pour la plupart des applications. Retenons en tous les cas qu'il est la plupart du temps préférable de sous estimer le nombre de neurones nécessaires que l'inverse.

## 2.2.6 Régularisation des poids : principes et techniques

Pour compléter le panel de techniques présentées dans le paragraphe précédent, le phénomène de sur-apprentissage peut également être prévenu assez simplement par l'utilisation d'une méthode de régularisation des poids. Nous avons en effet évoqué le fait que *l'overfitting* était finalement lié au fait que certains poids du réseau devenaient particulièrement grands avec pour conséquence les problématiques

soulevées dans la section 2.2.5. Les méthodes de régularisation de poids consistent ainsi à ajouter à la fonction de coût un terme de pénalisation liée à la grandeur des poids tel que :

$$C = \beta E_d + \alpha E_\omega \quad (2.27)$$

Avec  $E_d$ , la fonction de coût choisie et définie dans la section 2.2.3 et  $E_\omega$ , le terme de pénalisation. Bien que nombreuses et variées, une des méthodes les plus connues et les plus simples est probablement la régularisation  $L2$  [77]. Le terme de pénalité est alors égal à la somme quadratique des poids du réseau :

$$E_\omega = \sum_i \omega_i^2 \quad (2.28)$$

Les coefficients  $\alpha$  et  $\beta$  symbolisent la compétition entre la minimisation de  $E_d$  et la valeur des poids du réseau. Si par exemple,  $\beta$  est élevé par rapport à  $\alpha$ , alors la minimisation de la MSE est favorisée. A l'inverse, si nous savons que les données d'entraînement à disposition sont par exemple très bruitées, nous chercherons plutôt à favoriser la minimisation des poids. En d'autres termes, une prédominance de  $\beta$  sur  $\alpha$  favorisera une sortie du réseau plus sensible aux variations des entrées alors que si  $\alpha \gg \beta$ , nous imposons un comportement beaucoup plus lisse au réseau. Nous comprenons alors bien que les valeurs de ces deux hyperparamètres vont être particulièrement dépendantes du cas d'étude. Au même titre que pour la topologie ou le taux d'apprentissage, il peut donc être difficile et fastidieux de déterminer un couple de valeurs optimales. Néanmoins, il existe des méthodes avancées permettant de les ajuster efficacement et automatiquement au cours de l'algorithme par descente de gradient (voir section 2.3.3).

En injectant l'équation (2.27) dans l'équation (2.25), on obtient alors simplement :

$$\frac{\partial C}{\partial \omega_{jk}^l} = \frac{\partial (\beta E_d + \alpha E_\omega)}{\partial \omega_{jk}^l} = \beta a_k^{l-1} \delta_j^l + \alpha 2\omega_{jk}^l \quad (2.29)$$

D'après l'équation (2.26), il vient :

$$\omega_{jk, new}^l = \omega_{jk}^l - \eta (\beta a_k^{l-1} \delta_j^l + \alpha 2\omega_{jk}^l) = (1 - 2\alpha) \omega_{jk}^l - \eta \beta a_k^{l-1} \delta_j^l \quad (2.30)$$

De la même façon, l'équation (2.24) devient pour les biais :

$$b_{j, new}^l = b_j^l - \eta \beta \frac{\partial C}{\partial b_j^l} = b_j^l - \eta \beta \delta_j^l \quad (2.31)$$

## 2.3 Un réseau adaptée aux contraintes industrielles

---

### 2.3.1 Les contraintes d'un outil industriel

---

La section 2.2 dans son ensemble met en exergue plusieurs points importants. Tout d'abord, l'algorithme par descente de gradient décrit précédemment est mathématiquement relativement simple à mettre en œuvre et à développer, tout du moins dans sa version de base. Néanmoins, il existe dans sa mise en œuvre pratique un nombre relativement important d'hyperparamètres à fixer, tous dépendants du cas d'étude étudié. En d'autres termes, dans le cadre d'un déploiement global de l'outil à l'échelle de l'entreprise, la maîtrise de la théorie relative aux réseaux de neurones devient un pré-requis. De plus

et bien que certaines recommandations peuvent être formulées, il est très difficile de définir en avance les bons jeux de paramètres sans plusieurs itérations entre entraînement et conception du réseau. Ces difficultés sont très clairement un point bloquant et un frein au développement de ce type de méthodes. Il est donc nécessaire de pouvoir proposer un outil clé en main capable de s'adapter au problème à résoudre.

Bien que la *BackPropagation* soit un algorithme d'entraînement puissant et polyvalent, elle reste dans la version présentée ci-dessus très sujette au sur-apprentissage et sensible aux minimum locaux potentiels se trouvant sur l'hypersurface de la fonction de coût. Il est donc nécessaire que la méthode proposée soit suffisamment robuste et fiable pour limiter ces phénomènes pernicious et parfois difficiles à détecter. Dans le cadre de ces travaux, la minimisation à l'extrême de la fonction de coût n'est pas recherchée ou tout du moins, n'est pas une fin en soit. L'outil que nous proposons est un modèle réduit du rayonnement thermique dans les chambres de combustion. A ce titre et étant donné les erreurs générées par les modèles actuellement en production, nous privilégions avant tout les capacités de généralisation du réseau de neurones, c'est à dire sa capacité à interpoler des résultats de qualité en présence de données inconnues, plutôt que sa capacité à reproduire exactement des données d'entraînement. Pour nos applications, il sera notamment particulièrement important de saisir les maxima d'une grandeur radiative spécifique qui influenceront significativement sur les températures de paroi des chambres et la structure de flamme.

### 2.3.2 Une optimisation de la descente de gradient : algorithme de Levenberg-Marquardt

Depuis sa proposition par RUMELHART, HINTON et WILLIAMS [76], l'algorithme par descente de gradient a connu un nombre important d'améliorations ayant pour but à la fois d'améliorer sa convergence et de diminuer sa propension aux différents risques cités précédemment. Nous pouvons notamment citer parmi les algorithmes les plus utilisés, les méthodes SGD [84], RMSProp [85] et Adam [86], ainsi que leurs variantes. Ces trois méthodes sont aujourd'hui des incontournables pour les tâches de régression et sont notamment disponibles dans des bibliothèques très connues comme TensorFlow<sup>1</sup> et Keras<sup>2</sup>.

Nous nous intéressons ici à l'algorithme de Levenberg-Marquardt ou *LM* [87] reconnu pour sa rapidité de convergence et sa grande stabilité. C'est une méthode d'optimisation développée pour la minimisation numérique d'une fonction dépendant de plusieurs variables. Bien que non exclusivement destinée à l'entraînement d'un réseau de neurones, sa formulation est particulièrement adaptée à ce type de tâche. Globalement plus gourmande numériquement que les méthodes précédemment citées, elle est en revanche parfaitement adaptée au couplage avec la méthode de régularisation bayésienne (voir section 2.3.3).

*i*

L'algorithme *LM* s'inscrit exactement dans le même esprit que la descente de gradient. La procédure explicitée dans ce paragraphe est également proposée de manière détaillée dans le chapitre 12 de HAGAN et al. [69]. Au delà des différences de notation, il est important de préciser que l'algorithme présenté ci-dessous est simplifié pour un réseau de neurones avec une sortie unique comparativement au cas plus général traité dans l'ouvrage mentionné.

<sup>1</sup><https://www.tensorflow.org/>

<sup>2</sup><https://keras.io/>

**Introduction du Jacobien et notations** Dans ce paragraphe, nous privilégierons l'écriture des différentes équations sous forme matricielle. L'algorithme *LM* est particulièrement adapté à l'entraînement des réseaux de neurones lorsque la fonction de coût utilisée est une somme quadratique.

Nous rappelons tout d'abord que dans l'ensemble de ce chapitre, nous avons utilisé la MSE comme fonction de coût. L'équation (2.10) peut notamment s'écrire sous la forme d'une espérance telle que :

$$C = \mathbb{E} [(\mathbf{t} - \mathbf{a}^L)^T(\mathbf{t} - \mathbf{a}^L)] = \mathbb{E} [\mathbf{e}^T \mathbf{e}] \quad (2.32)$$

Dans cet algorithme, nous travaillons désormais non plus avec l'espérance de l'erreur quadratique mais directement avec la somme des erreurs quadratiques. Ainsi,  $C$  devient simplement :

$$C = \mathbf{e}^T \mathbf{e} = \sum_i e_i^2 \quad (2.33)$$

Contrairement à la descente de gradient classique, l'algorithme de Levenberg-Marquardt nous impose de raisonner non plus *cas d'entraînement par cas d'entraînement* mais sur l'ensemble de la base de données contenant  $N$  jeux. L'expression du gradient devient alors dans ce cas [69] :

$$\nabla C(\mathbf{v}) = 2\mathbf{J}^T \mathbf{e} \quad (2.34)$$

La matrice  $\mathbf{J}$  est ici le jacobien de la base de données. Elle se définit par la concaténation de ses composantes  $\mathbf{J}_x$  telle que :

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \\ \vdots \\ \mathbf{J}_N \end{bmatrix} = \begin{bmatrix} \frac{\partial e_1}{\partial v_1} & \frac{\partial e_1}{\partial v_2} & \cdots & \frac{\partial e_1}{\partial v_{n_p}} \\ \frac{\partial e_2}{\partial v_1} & \frac{\partial e_2}{\partial v_2} & \cdots & \frac{\partial e_2}{\partial v_{n_p}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial e_N}{\partial v_1} & \frac{\partial e_N}{\partial v_2} & \cdots & \frac{\partial e_N}{\partial v_{n_p}} \end{bmatrix} \quad (2.35)$$

Ici,  $n_p$  et  $N$  sont définis respectivement comme le nombre de paramètres du réseau et le nombre de jeux d'entraînement de la base de données. Nous pouvons également remarquer que la matrice jacobienne  $\mathbf{J}$  est de taille  $N \times n_p$  et que  $\mathbf{E}$  est de taille  $N \times 1$  dans le cas où la sortie du réseau est unique. Dans cette situation, la matrice gradient est donc une matrice colonne de taille  $n_p \times 1$ . Les composantes de cette matrice correspondent donc à un gradient global calculé pour l'ensemble des jeux d'entraînement et ce, pour l'ensemble des paramètres du réseau. Nous retrouvons alors exactement la même forme mathématique que définie plus tôt dans l'équation (2.12).

**Calcul du Jacobien et sensibilités de Marquardt** Le coeur de l'algorithme *LM* se trouve dans la procédure de calcul de la matrice jacobienne qui se base donc sur une variation de la descente de gradient standard. On remarque notamment que d'après l'équation (2.35), il est nécessaire ici de calculer les dérivées partielles des erreurs plutôt que des erreurs quadratiques.

Nous définissons ici la sensibilité de Marquardt, notée  $\tilde{\delta}$  telle que :

$$\tilde{\delta}_{j,x}^l = \frac{\partial e_x}{\partial z_{j,x}^l} \quad (2.36)$$

Chaque composante de la matrice jacobienne se décompose alors comme suit :

$$[\mathbf{J}]_{i,x} = \frac{\partial e_x}{\partial v_i} = \frac{\partial e_x}{\partial z_{j,x}^l} \times \frac{\partial z_{j,x}^l}{\partial v_i} \quad (2.37)$$

Si  $v_i$  est le poids  $\omega_{jk}^l$  (voir les notations utilisées dans la figure 2.5 et la section 2.2.4), alors :

$$[\mathbf{J}]_{i,x} = \tilde{\delta}_{j,x}^l \times a_{k,x}^{l-1} \quad (2.38)$$

Si  $v_i$  est le biais  $b_j^l$ , alors :

$$[\mathbf{J}]_{i,x} = \tilde{\delta}_{j,x}^l \quad (2.39)$$

Sur la base de ces nouvelles définitions, la relation de récurrence entre les différentes couches définie dans la section 2.2.4 reste la même. Le calcul de la sensibilité de Marquardt sur la dernière couche  $L$  diffère quant à lui par rapport à l'algorithme classique. Nous avons en effet :

$$\tilde{\delta}_x^L = \frac{\partial e_x}{\partial z_x^L} = \frac{\partial (t_x - a_x^L)}{\partial z_x^L} = -\frac{\partial a_x^L}{\partial z_x^L} = -s'(z_x^L) \quad (2.40)$$

L'équation (2.22) nous permet finalement d'écrire la rétropropagation du vecteur sensibilité dans le cas de l'algorithme  $LM$  :

$$\tilde{\delta}_x^l = [(\boldsymbol{\omega}^{l+1})^T \tilde{\delta}_x^{l+1}] \odot s'(z_x^l) \quad (2.41)$$

Enfin, au même titre que la matrice Jacobienne, la matrice totale des sensibilités de Marquardt se construit par concaténation des vecteurs lignes  $\tilde{\delta}_x$ . Il suffit alors d'appliquer les équations (2.38) et (2.39) pour construire la Jacobienne de l'ensemble de la base de données.

**Approximation de la matrice Hessienne et loi de mouvement** Le calcul de la matrice jacobienne nous permet ensuite de construire une approximation de la matrice Hessienne relative à l'ensemble de la base de données et nécessaire à la définition de la loi de mouvement des paramètres du réseau. En effet, la méthode de Levenberg-Marquardt est certes une variation de la *BackPropagation* lorsqu'elle appliquée aux réseaux de neurones mais est également et avant tout une variation de la méthode d'optimisation de Newton [88]. A ce titre et dans le cadre de l'entraînement du réseau, les paramètres à l'époque  $p + 1$  sont déduits des paramètres de l'époque  $p$  à partir de :

$$\mathbf{v}_{p+1} = \mathbf{v}_p - [\nabla^2 \mathbf{C}(\mathbf{v}_p)]^{-1} \cdot \nabla \mathbf{C}(\mathbf{v}_p) \quad (2.42)$$

Une approximation correcte de la matrice Hessienne [69] peut être exprimée uniquement en fonction de la Jacobienne :

$$\mathbf{H}_e = \nabla^2 \mathbf{C} \approx \mathbf{J}^T \mathbf{J} \quad (2.43)$$

En injectant les équations (2.34) et (2.43) dans l'équation (2.42), nous obtenons alors la méthode d'optimisation de Gauss-Newton [89] :

$$\mathbf{v}_{p+1} = \mathbf{v}_p - [\mathbf{J}^T(\mathbf{v}_p)\mathbf{J}(\mathbf{v}_p)]^{-1} \mathbf{J}^T(\mathbf{v}_p)\mathbf{e} \quad (2.44)$$

La méthode de Gauss-Newton, du fait de l'approximation de la matrice Hessienne, a l'avantage d'un point de vue calculatoire de ne pas rendre nécessaire le calcul de dérivées secondes. Néanmoins, il est possible que  $\mathbf{H}_e$  telle qu'elle est exprimée dans l'équation (2.43) ne soit pas inversible. L'algorithme de Levenberg-Marquardt utilise ainsi une autre approximation de  $H_e$  permettant d'assurer que la matrice soit définie positive :

$$\mathbf{H}_e = \mathbf{J}^T\mathbf{J} + \eta_p\mathbf{I} \quad (2.45)$$

Avec  $\mathbf{I}$ , la matrice identité. D'où finalement :

$$\mathbf{v}_{p+1} = \mathbf{v}_p - [\mathbf{J}^T(\mathbf{v}_p)\mathbf{J}(\mathbf{v}_p) + \eta_p\mathbf{I}]^{-1} \mathbf{J}^T(\mathbf{v}_p)\mathbf{e} \quad (2.46)$$

*i*

Nous pouvons constater que la matrice Hessienne est construite par le produit matriciel de la transposée de la Jacobienne par elle-même. Les dimensions de  $H_e$  sont donc égales à  $(n_p \times N) \times (N \times n_p)$  soit une matrice carrée de dimension  $(n_p \times n_p)$ . On comprend ainsi assez aisément pourquoi cette méthode est considérée comme plus gourmande en termes de ressources de calculs car cette matrice en particulier peut rapidement devenir très importante jusqu'à potentiellement saturer la RAM. Notons également qu'il est nécessaire dans l'algorithme d'inverser cette matrice. Cette opération est sans aucun doute la plus coûteuse numériquement dans cette méthode d'entraînement. La taille de la Jacobienne (voir l'équation (2.35)) est également directement liée au nombre de paramètres mais aussi au nombre de jeux d'entraînement. En d'autres termes, la maîtrise de la topologie du réseau et de la base de données est essentielle lors de l'utilisation d'un tel algorithme.

**Adaptation automatique du taux d'apprentissage** On retrouve dans l'équation (2.46) un équivalent du taux d'apprentissage déjà rencontré dans la section 2.2.4. Cet algorithme a notamment le gros avantage de proposer une loi de mouvement également pour le taux d'apprentissage  $\eta_p$  et donc, d'automatiser son évolution quel que soit le cas étudié. Le taux d'apprentissage est généralement initialisé à 0.01 au début de l'algorithme. Puis, quand on calcule le vecteur paramètre  $\mathbf{v}_{p+1}$  mis à jour au cours de l'entraînement, il est nécessaire de relancer une *FeedForward Pass* afin de calculer une nouvelle valeur de  $C$ . Si cette valeur est inférieure à la valeur précédente, alors le réseau s'améliore et dans ce cas, le taux d'apprentissage  $\eta$  est divisé par un facteur 10 et les nouvelles valeurs des paramètres sont validées pour cette époque. A l'inverse, si  $C$  augmente, alors la valeur de  $\eta$  est multipliée par 10. On calcule alors de nouveau les valeurs actualisées des paramètres et ainsi de suite. Ainsi, il existe donc en quelque sorte, des époques au sein de chaque époque afin de déterminer la bonne valeur de  $\eta_p$  à chaque itération de l'algorithme.

### 2.3.3 Paradigme et régularisation bayésienne

**Introduction** Dans la section 2.2.6, nous avons introduit la notion de régularisation comme technique permettant de limiter le sur-apprentissage du réseau. Néanmoins, certaines limites pratiques et de mises en œuvre sont susceptibles de réduire son efficacité et son déploiement à grande échelle. Conformément

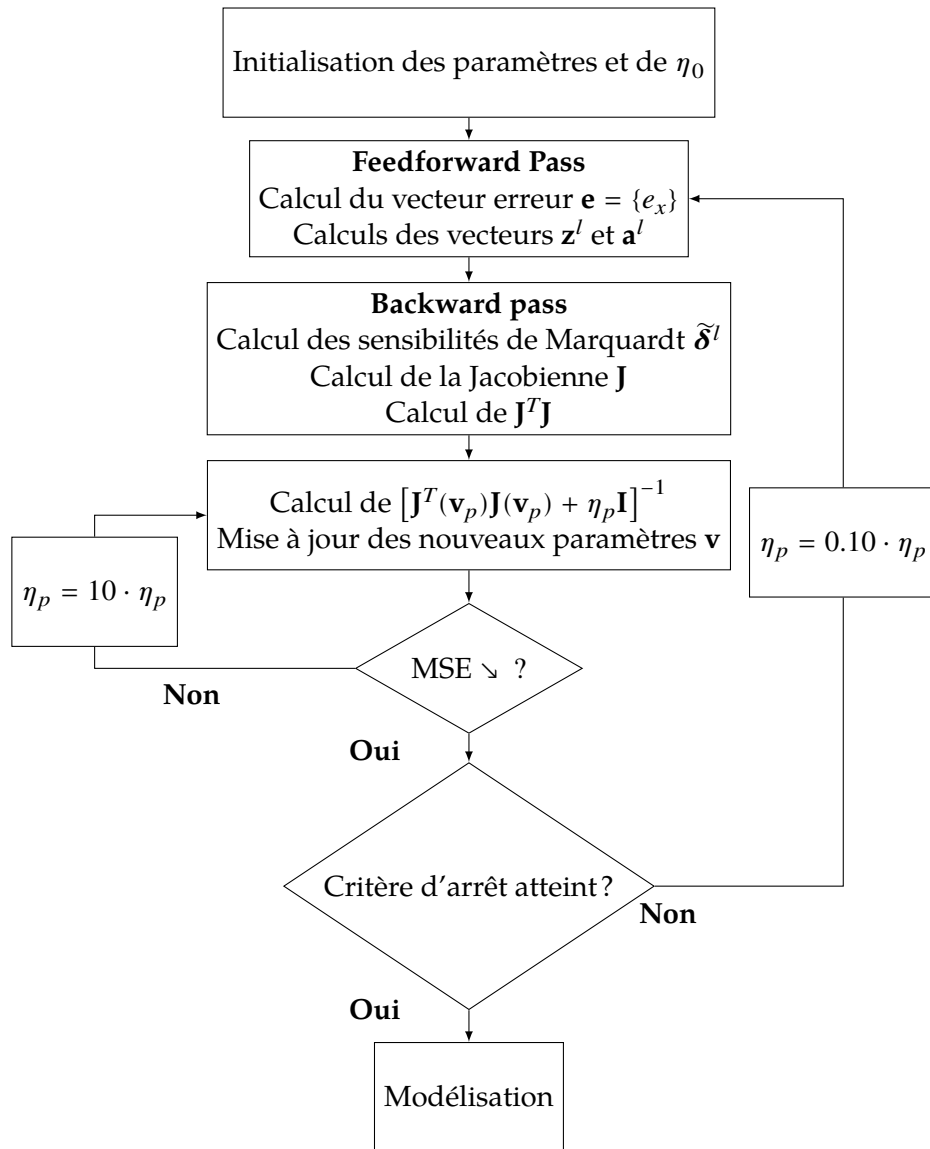


FIG. 2.14 : Logigramme de l'algorithme de Levenberg-Marquardt

au besoin industriel, nous proposons ici de se placer dans le cadre du paradigme Bayésien, appliqué aux techniques de régularisation afin notamment d'automatiser la définition des hyperparamètres  $\alpha$  et  $\beta$  (voir équation (2.27)).

Le théorème de Bayes est un des fondamentaux de la théorie des probabilités conditionnelles. Dans sa forme très générale, il permet d'exprimer la probabilité d'un évènement A sachant B  $P(A|B)$  en fonction de  $P(A)$ ,  $P(B)$  et  $P(B|A)$  :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (2.47)$$

Comme la plupart des grands théorèmes, cette formulation pourtant très simple permet de poser un paradigme plus complet dont la portée va bien au delà de celui qu'on pourrait lui donner de prime abord. Du précédent théorème découle notamment les méthodes d'inférence et d'optimisation bayésienne.

Cette écriture permet de mesurer la confiance que l'on peut accorder à une hypothèse en connaissant sa conséquence, à savoir la réalisation d'un évènement. En quelque sorte et de manière vulgarisée, le point de vue bayésien s'intéresse à la probabilité des théories sachant les évènements tandis que les statistiques plus standards traitent des évènements sachant les théories.

En un sens, on peut dire que le paradigme bayésien n'est finalement que la formulation mathématique du processus de construction de l'opinion chez l'Homme. En effet, pour expliquer et comprendre le monde qui nous entoure, nous commençons tous par poser un modèle dit *a priori*, issu de choses entendues çà et là, de notre éducation ou de nos croyances personnelles. Il advient bien souvent que ce modèle ne corresponde pas à la réalité du monde ou plutôt que certaines observations nouvelles influent sur le postulat que nous avons fait au départ. Dans un monde idéal, nous changeons d'avis, nous évoluons et adaptons notre modèle du monde aux constats que nous pouvons faire dans la vie de tous les jours. De ce point de vue, nous adoptons alors un fonctionnement purement bayésien où chaque observation d'évènement, de conséquence, *infère* sur la théorie, sur la cause, que nous avons posée jusqu'alors.

Nous pouvons alors commencer à voir poindre un lien avec le sujet qui nous intéresse, à savoir l'apprentissage d'un réseau de neurones. Dans ce cas, le modèle *a priori* serait le vecteur paramètre  $\mathbf{v}$  initialisé aléatoirement. Les évènements observés correspondraient à chaque jeu de données contenus dans la base d'entraînement. La régularisation bayésienne va ainsi nous permettre d'optimiser, d'inférer les valeurs des hyperparamètres de notre modèle sur la base des *faits constatés dans le réel*.

**Hypothèses et formulations probabilistes de l'apprentissage dans le paradigme bayésien** Ce paragraphe vise à présenter le développement mathématique et les hypothèses posées afin de déterminer les lois de mouvement des hyperparamètres de régularisation  $\alpha$  et  $\beta$ . Le paradigme bayésien nous permet également de connaître le *besoin* du réseau en termes de paramètres libres, et donc d'avoir des informations sur sa topologie. Nous nous basons sur les travaux de MACKAY [68], de FORESEE et HAGAN [90] ainsi que sur le chapitre 13 de HAGAN et al. [69] dans la suite de ce paragraphe.

Nous notons  $\mathbb{D}$ , l'ensemble des données d'entraînement et  $\mathbb{M}$ , le modèle de réseau de neurones (nombres de neurones par couche et nombre de couches). Le théorème de Bayes (équation (2.47)) nous permet d'écrire sous forme probabiliste [68] :

$$P(\mathbf{v}|\mathbb{D}, \alpha, \beta, \mathbb{M}) = \frac{P(\mathbb{D}|\mathbf{v}, \beta, \mathbb{M})P(\mathbf{v}|\alpha, \mathbb{M})}{P(\mathbb{D}|\alpha, \beta, \mathbb{M})} \quad (2.48)$$

Cette fonction est dite *a posteriori*. Dans la continuité de l'analogie donnée en début de section, cette fonction représenterait une opinion après inférence de nos expériences sur notre postulat de départ. Afin de pouvoir poursuivre le développement, il est nécessaire de préciser que nous supposons que les sorties théoriques de la base de données sont bruitées et que ce bruit répond à une gaussienne centrée en zéro avec un écart type  $\sigma_t$ . Dans ces conditions, nous pouvons alors écrire :

$$P(\mathbb{D}|\mathbf{v}, \beta, \mathbb{M}) = \frac{1}{Z_d(\beta)} \exp(-\beta E_d) \quad (2.49)$$

Où  $\beta = \frac{1}{2\sigma_t^2}$  et  $Z_d(\beta) = (\pi/\beta)^{N/2}$ . La probabilité définie dans l'équation (2.49) est appelée la fonction de vraisemblance ou *likelihood function* en anglais. Elle traduit la probabilité qu'un jeu de données d'entrées corresponde à sa sortie théorique pour un ensemble de paramètres  $\mathbf{v}$ , une valeur de  $\beta$  et un modèle  $\mathbb{M}$  connus. On s'aperçoit alors que minimiser la valeurs de  $E_d$  revient à maximiser cette probabilité et inversement.



La probabilité  $P(\mathbf{v}|\alpha, \mathbb{M})$  est appelée la fonction *a priori* (ou *Prior function*). Elle représente les connaissances que nous avons avant toute inférence de données sur les poids du réseau. En pratique, nous pouvons tout simplement définir la forme de cette probabilité en imposant de tirer les paramètres du réseau à l'initialisation de telle manière à ce qu'ils correspondent également à une loi normale centrée en zéro :

$$P(\mathbf{v}|\alpha, \mathbb{M}) = \frac{1}{Z_\omega(\alpha)} \exp(-\alpha E_\omega) \quad (2.50)$$

Avec  $\alpha = \frac{1}{\sigma_\omega^2}$ ,  $\sigma_\omega$  l'écart type sur les poids du réseau et  $Z_\omega(\alpha) = (\pi/\alpha)^{n_p/2}$ .

Enfin, le dernier terme dans l'équation (2.48) correspond à la fonction *Evidence* et ne dépend pas des poids du réseau mais seulement des hyperparamètres de régularisation et de topologie. En l'état, on peut la considérer comme une fonction de normalisation. Finalement et compte-tenu des équations (2.49) et (2.50) et en injectant l'équation (2.27) dans l'équation (2.47) nous pouvons réécrire la fonction *a posteriori* telle que :

$$P(\mathbf{v}|\mathbb{D}, \alpha, \beta, \mathbb{M}) = \frac{\frac{1}{Z_d(\beta)} \frac{1}{Z_\omega(\alpha)} \exp(-(\beta E_d + \alpha E_\omega))}{P(\mathbb{D}|\alpha, \beta, \mathbb{M})} = \frac{1}{Z_c(\alpha, \beta)} \exp(-C(\mathbf{v})) \quad (2.51)$$

Au même titre que précédemment pour la fonction de vraisemblance, nous pouvons remarquer que trouver le vecteur paramètre  $\mathbf{v}$  le plus probable, c'est à dire celui permettant de maximiser la correspondance entre les entrées et les sorties théoriques de la base de données, revient finalement à minimiser cette fois-ci la fonction de coût régularisée  $C$ .

*i*

Dans cette formulation probabiliste, basée sur les distributions normales des différentes grandeurs, nous pouvons déjà réussir à reconnaître une certaine élégance dans les formulations précédentes. En effet, l'hyperparamètre de régularisation  $\beta$  est inversement proportionnel à la variance sur les données d'entraînement. En d'autres termes, plus les données utilisées seront bruitées, plus la valeur de  $\beta$  sera faible. Conformément à la section 2.2.6, l'algorithme ne favorisera donc pas une correspondance à tout prix entre les données d'entraînement et les prédictions du réseau de neurone. La réponse du réseau sera donc plus lisse et le bruit ne sera pas *appris*. Cette formulation prévient donc naturellement l'*overfitting* en fonction des données disponibles.

**Détermination des Hyperparamètres de régularisation** Le cadre bayésien et ses implications sur la formulation probabiliste du problème étant posé, nous souhaitons à présent déterminer l'expression des hyperparamètres  $\alpha$  et  $\beta$  en fonction des données d'entraînement disponibles. L'idée ici est d'appliquer la même procédure que précédemment en cherchant cette fois la probabilité *a priori*  $P(\alpha, \beta|\mathbb{D}, \mathbb{M})$  en lui appliquant le théorème de Bayes [68]. Nous avons donc :

$$P(\alpha, \beta|\mathbb{D}, \mathbb{M}) = \frac{P(\mathbb{D}|\alpha, \beta, \mathbb{M})P(\alpha, \beta|\mathbb{M})}{P(\mathbb{D}|\mathbb{M})} \quad (2.52)$$

Dans ce cas, nous supposons que la fonction *a priori*  $P(\alpha, \beta|\mathbb{M})$  est uniforme. La maximisation de la fonction  $P(\alpha, \beta|\mathbb{D}, \mathbb{M})$  revient alors à la maximisation de la fonction de vraisemblance  $P(\mathbb{D}|\alpha, \beta, \mathbb{M})$ . On remarque par ailleurs que cette probabilité correspond à la fonction évidence dans l'équation (2.48). Au

vu de ce qui précède et en nous basant sur l'ensemble des équations (2.48) à (2.51), nous pouvons donc écrire :

$$P(\mathbb{D}|\alpha, \beta, \mathbb{M}) = \frac{P(\mathbb{D}|\mathbf{v}, \beta, \mathbb{M})P(\mathbf{v}|\alpha, \mathbb{M})}{P(\mathbf{v}|\mathbb{D}, \alpha, \beta, \mathbb{M})} = \frac{\left[ \frac{1}{Z_d(\beta)} \exp(-\beta E_d) \right] \cdot \left[ \frac{1}{Z_\omega(\alpha)} \exp(-\alpha E_\omega) \right]}{\frac{1}{Z_c(\alpha, \beta)} \exp(-C(\mathbf{v}))} \quad (2.53)$$

En simplifiant l'équation (2.53), on obtient finalement une expression pour la fonction évidence de l'équation (2.48) et donc de la fonction de vraisemblance de l'équation (2.52) :

$$P(\mathbb{D}|\alpha, \beta, \mathbb{M}) = \frac{Z_c(\alpha, \beta)}{Z_d(\beta)Z_\omega(\alpha)} \quad (2.54)$$

Dans cette expression, nous connaissons les constantes  $Z_d(\beta)$  et  $Z_\omega(\alpha)$  dans la mesure où nous supposons que  $\alpha$  et  $\beta$  sont connus. En revanche, le calcul de  $Z_f(\alpha, \beta)$  demande un développement supplémentaire. En décomposant la fonction de coût  $C(\mathbf{v})$  selon une série de Taylor à son point minimum [69], c'est à dire lorsque le vecteur paramètre *le plus probable*  $\mathbf{v}^{MP}$  maximise la fonction *a posteriori* (voir équation (2.51)), on obtient finalement :

$$P(\mathbf{v}|\mathbb{D}, \alpha, \beta, \mathbb{M}) \approx \left[ \frac{1}{Z_c} \exp(-C(\mathbf{v}^{MP})) \right] \exp \left[ -\frac{1}{2}(\mathbf{v} - \mathbf{v}^{MP})^T \mathbf{H}\mathbf{e}^{MP} (\mathbf{v} - \mathbf{v}^{MP}) \right] \quad (2.55)$$

Par identification avec la forme standard d'une densité de probabilité répondant à une loi normale, nous pouvons donc en déduire :

$$Z_c(\alpha, \beta) \approx (n_p \pi)^{\frac{1}{2}} (\det((\mathbf{H}\mathbf{e}^{MP})^{-1}))^{\frac{1}{2}} \exp(-C(\mathbf{v}^{MP})) \quad (2.56)$$

En injectant cette expression dans l'équation (2.54), il devient alors possible de déterminer les valeurs optimales de  $\alpha$  et  $\beta$  au point *le plus probable*, c'est à dire permettant de maximiser la probabilité *a posteriori* et donc de minimiser la fonction de coût  $C(\mathbf{v})$ . Mathématiquement, le développement n'est pas trivial mais les travaux de МАСКАУ [68] mènent aux expressions suivantes pour les hyperparamètres de régularisation :

$$\alpha = \frac{\gamma}{2E_\omega(\mathbf{v})} \quad (2.57)$$

$$\beta = \frac{N - \gamma}{2E_d(\mathbf{v})} \quad (2.58)$$

On note alors l'introduction d'une nouvelle grandeur  $\gamma$ . Elle correspond au *nombre effectif de paramètres* et est définie comme suit :

$$\gamma = n_p - 2\alpha \operatorname{tr}(\mathbf{H}\mathbf{e}^{-1}) \quad (2.59)$$

Les équations (2.57) à (2.59) sont liées les unes aux autres mais l'algorithme d'entraînement étant itératif, cette dépendance n'est pas un problème. Il suffira d'initialiser un des hyperparamètres pour connaître les autres. Leurs valeurs convergeront ensuite vers les valeurs *les plus probables* au sens du théorème de Bayes, c'est à dire vers celles permettant de minimiser au mieux la fonction de coût.

i

Nous pouvons relever dans les équations (2.55), (2.56) et (2.59) l'apparition de la matrice Hessienne  $\mathbf{H}_e$ . La méthode de régularisation bayésienne est donc particulièrement adaptée au couplage avec un entraînement par algorithme de Levenberg-Marquardt et peut être appliquée sans aucun surcoût computationnel.

Dans la section 2.3.2, nous avons pu définir la loi de mouvement des paramètres du réseau. Si la fonction de coût est régularisée, alors l'équation (2.46) devient :

$$\mathbf{v}_{p+1} = \mathbf{v}_p - \frac{[2\beta\mathbf{J}^T(\mathbf{v}_p)\mathbf{J}(\mathbf{v}_p) + 2\alpha\eta_p\mathbf{I}]}{\mathbf{H}_e}^{-1}\mathbf{J}^T(\mathbf{v}_p)\mathbf{e} \quad (2.60)$$

**Application du rasoir d'Ockham : la parcimonie comme principe** Comme discuté plus haut dans ce chapitre, nous savons que les différents algorithmes par descente de gradient sont globalement sujets au sur-apprentissage. Il est donc nécessaire de prendre des contre-mesures pour minimiser au maximum cette tendance. La régularisation fait partie de ces techniques mais a ses défauts. Nous avons montré précédemment que le paradigme bayésien permettait néanmoins de dépasser les limitations sur la mise en place de telles méthodes.

Néanmoins, pour des raisons à la fois de performance de calcul et de rigueur de l'algorithme, la régularisation à elle seule ne suffit pas. Nous avons déjà cité dans la section 2.2.5 un adage particulièrement approprié. Nous souhaitons mettre en évidence dans ce paragraphe l'intérêt du théorème de Bayes en accord avec cette maxime. Dans ses travaux, MACKAY [68] formalise la portée du paradigme bayésien par l'application du *Rasoir d'Ockham*. C'est un principe bien connu qui indique que la multiplication des hypothèses est toujours dommageable. Sa formulation latine originelle peut être textuellement traduite par *Les multiples ne doivent pas être utilisés sans nécessité*. Par contraposition, si nous nous replaçons dans le contexte général de nos travaux, nous pouvons directement tirer de ce principe que la multiplication des degrés de liberté d'un réseau par l'ajout de neurones est non seulement inutile mais problématique. Le rasoir d'Ockham est donc globalement un principe qui invite le scientifique à la parcimonie.

Précédemment, dans l'équation (2.59), nous avons défini le nombre effectif de paramètres  $\gamma$ , borné entre 0 et le nombre de paramètre réel  $n_p$ . Cette grandeur, calculée pendant l'algorithme, permet de donner des informations sur le nombre de paramètres réellement nécessaires à la modélisation de la fonction visée. Implicitement,  $\gamma$  nous indique donc si le réseau utilisé est sur-dimensionné ou non. Il est alors très simple d'implémenter dans un code d'apprentissage une méthode permettant d'ajouter ou de supprimer pendant le calcul des neurones de manière automatique et relative à la base de données (voir figure 3.2).

Si  $\gamma$  est égal ou très proche au nombre de paramètres réels du réseau, alors nous pouvons dire qu'il manque de souplesse et ajouter des neurones. A l'inverse, si  $\gamma \ll n_p$ , l'outil est sur-dimensionné et il devient nécessaire d'adapter la topologie en conséquence. Dans la pratique, nous recherchons une valeur de  $\gamma$  comprise entre 90 à 85 % de la valeur de  $n_p$ .

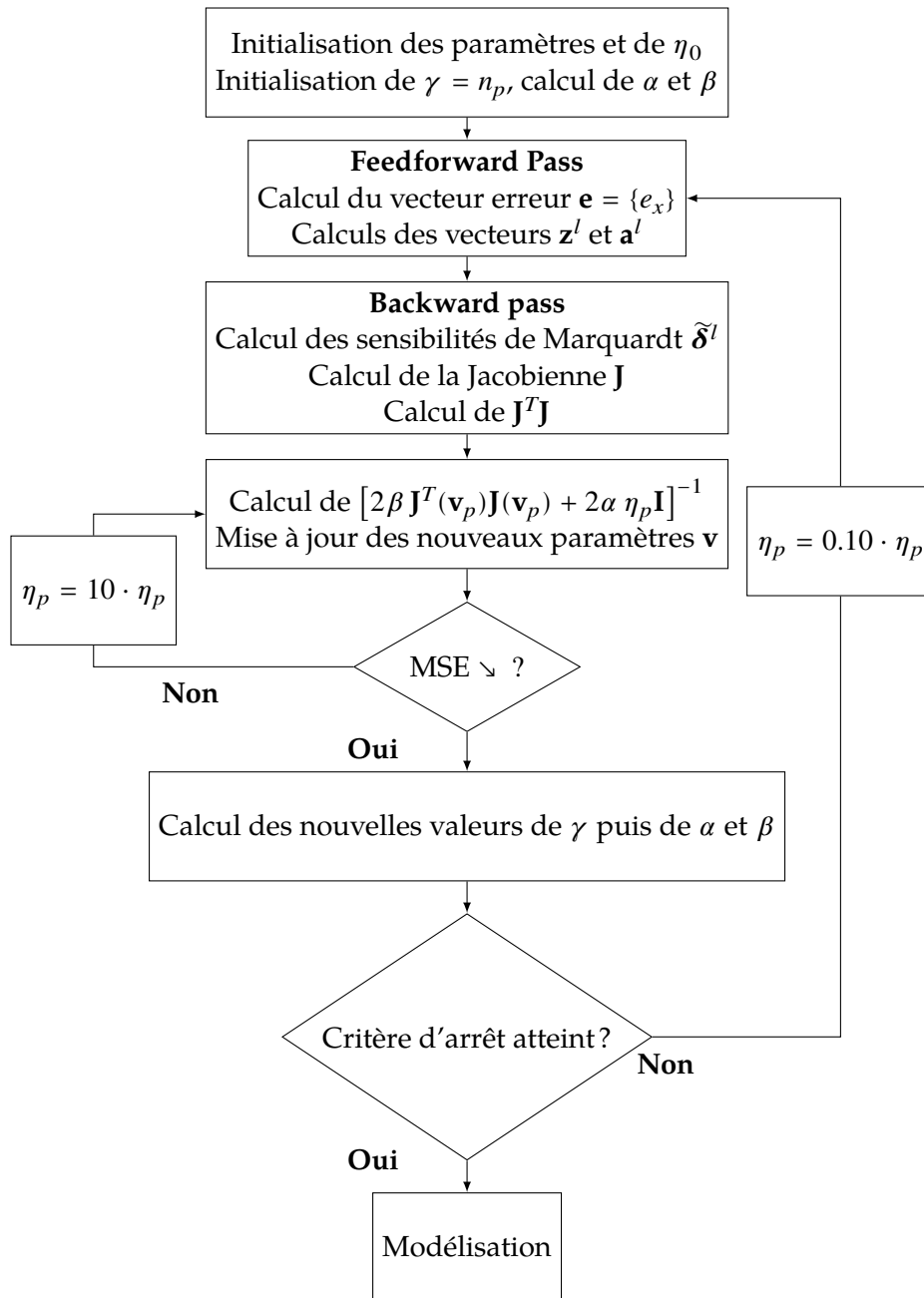


FIG. 2.15 : Logigramme de l'algorithme *LM* couplé à la régularisation bayésienne

**Conclusion** Pour résumer l'ensemble de cette section, le paradigme bayésien nous permet de surmonter la plupart des difficultés soulignées précédemment. Par conséquent, lorsque la régularisation est couplée avec l'entraînement *LM*, il n'existe plus aucun hyperparamètre *ad hoc* à fixer à l'exception, dans l'absolu, du nombre de couches cachées du réseau. L'intérêt de n'en utiliser qu'une seule permet ainsi de ne pas avoir à se demander sur quelle couche il est nécessaire d'ajouter ou de supprimer les nouveaux neurones, toujours dans l'esprit de proposer un outil clé en main. L'algorithme s'adapte et s'optimise automatiquement en fonction des données qui lui sont proposées pour devenir un outil performant et optimisé pour chaque cas d'étude. Il existe en revanche une limitation sur la quantité de données

utilisables dans la base d'entraînement et sur le nombre de neurones, imposée par les ressources de calcul disponibles.

## Références

- [66] K. HORNIK, M. STINCHCOMBE et H. WHITE. «Multilayer feedforward networks are universal approximators». en. In : *Neural Networks* 2.5 (1989), p. 359-366. DOI : [10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- [67] M. LESHNO et al. «Multilayer feedforward networks with a nonpolynomial activation function can approximate any function». en. In : *Neural Networks* 6.6 (jan. 1993), p. 861-867. DOI : [10.1016/S0893-6080\(05\)80131-5](https://doi.org/10.1016/S0893-6080(05)80131-5).
- [68] D. J. C. MACKEY. «Bayesian Interpolation». In : *Neural Computation* 4.3 (1992), p. 415-447.
- [69] M. T. HAGAN et al. *Neural Network Design*. 2<sup>e</sup> éd. Boston : PWS Publishing Co., 1996.
- [70] W. S. McCULLOCH et W. PITTS. «A logical calculus of the ideas immanent in nervous activity». en. In : *Bulletin of Mathematical Biophysics* 5 (1943), p. 19.
- [71] A. C. NEVES et al. «A New Approach to Damage Detection in Bridges Using Machine Learning». en. In : *Experimental Vibration Analysis for Civil Structures*. Sous la dir. de J. P. CONTE et al. T. 5. Series Title : Lecture Notes in Civil Engineering. San Diego, USA : Springer International Publishing, 2018, p. 73-84. DOI : [10.1007/978-3-319-67443-8\\_5](https://doi.org/10.1007/978-3-319-67443-8_5).
- [72] F. ROSENBLATT. «The perceptron : A probabilistic model for information storage and organization in the brain.» en. In : *Psychological Review* 65.6 (1958), p. 386-408. DOI : [10.1037/h0042519](https://doi.org/10.1037/h0042519).
- [73] Z. LI et al. «A Survey of Convolutional Neural Networks : Analysis, Applications, and Prospects». en. In : *IEEE Transactions on Neural Networks and Learning Systems* 33.12 (déc. 2022), p. 6999-7019. DOI : [10.1109/TNNLS.2021.3084827](https://doi.org/10.1109/TNNLS.2021.3084827).
- [74] Z. C. LIPTON, J. BERKOWITZ et C. ELKAN. «A critical review of recurrent neural networks for sequence learning». arXiv preprint arXiv :1506.00019. 2015.
- [75] A. APICELLA et al. «A survey on modern trainable activation functions». In : *Neural Networks* 138 (2021), p. 14-32. DOI : [10.1016/j.neunet.2021.01.026](https://doi.org/10.1016/j.neunet.2021.01.026). arXiv : [2005.00817\[cs, stat\]](https://arxiv.org/abs/2005.00817).
- [76] D. E. RUMELHART, G. E. HINTON et R. J. WILLIAMS. «Learning representation by back-propagating errors». In : *Nature* 323.9 (1986), p. 533-536.
- [77] M. NIELSEN. *Neural Networks and Deep Learning*. 2015.
- [78] S. K. KUMAR. «On weight initialization in deep neural networks». 2017.
- [79] L. LU et al. «Dying relu and initialization : Theory and numerical examples». 2019.
- [80] S. LAWRENCE, C. L. GILES et A. C. TSOI. «Lessons in neural network training : Overfitting may be harder than expected». In : *AAAI/IAAI*. 1997, p. 540-545.
- [81] Y. TIAN et Y. ZHANG. «A comprehensive survey on regularization strategies in machine learning». In : *Information Fusion* 80 (2022), p. 146-166. DOI : <https://doi.org/10.1016/j.inffus.2021.11.005>.
- [82] L. PRECHELT. «Early Stopping | but when?» In : *Neural Networks : Tricks of the Trade*. Second Edition. 2012, p. 53-67.
- [83] D. WHITLEY et al. «Genetic algorithms and neural networks». In : *Genetic algorithms in engineering and computer science* 3 (1995), p. 191-201.
- [84] L. BOTTOU. «Stochastic Gradient Descent Tricks». In : *Neural Networks : Tricks of the Trade*. Sous la dir. de G. MONTAVON, G. B. ORR et K.-R. MÜLLER. T. 7700. Series Title : Lecture Notes in Computer Science. Berlin, Heidelberg : Springer Berlin Heidelberg, 2012, p. 421-436. DOI : [10.1007/978-3-642-35289-8\\_25](https://doi.org/10.1007/978-3-642-35289-8_25).

- [85] T. TIELEMAN, G. HINTON et al. «Lecture 6.5-rmsprop : Divide the gradient by a running average of its recent magnitude». In : *COURSERA : Neural networks for machine learning* 4.2 (2012), p. 26-31.
- [86] D. P. KINGMA et J. BA. «Adam : A method for stochastic optimization». In : *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*. 2014.
- [87] D. W. MARQUARDT. «An algorithm for least-squares estimation of nonlinear parameters». In : *Journal of the society for Industrial and Applied Mathematics* 11.2 (1963), p. 431-441.
- [88] J. J. MORÉ. «The Levenberg-Marquardt algorithm : Implementation and theory». In : *Numerical Analysis*. Sous la dir. de G. A. WATSON. T. 630. Series Title : Lecture Notes in Mathematics. Berlin, Heidelberg : Springer Berlin Heidelberg, 1978, p. 105-116. DOI : [10.1007/BFb0067700](https://doi.org/10.1007/BFb0067700).
- [89] S. GRATTON, A. S. LAWLESS et N. K. NICHOLS. «Approximate Gauss-Newton Methods for Nonlinear Least Squares Problems». In : *SIAM Journal on Optimization* 18.1 (2007), p. 106-132. DOI : [10.1137/050624935](https://doi.org/10.1137/050624935).
- [90] D. F. FORESEE et M. HAGAN. «Gauss-Newton approximation to Bayesian learning». In : *Proceedings of International Conference on Neural Networks (ICNN'97)*. International Conference on Neural Networks (ICNN'97). T. 3. Houston, TX, USA : IEEE, 1997, p. 1930-1935. DOI : [10.1109/ICNN.1997.614194](https://doi.org/10.1109/ICNN.1997.614194).

# Modélisation du rayonnement par réseau de neurones bayésien et algorithme de Monte Carlo

## Sommaire du chapitre

3.1	Rappel du constat et proposition	70
3.2	Interfaçage des méthodes : Monte Carlo pour la base de données et IA pour interpolation	71
3.3	Synthèse de la méthode	76
	Références	78

## Notations

Symbole	Description	Unité
$a_{j,i}$	Coefficient de Monte Carlo symbolique	-
$\alpha, \beta$	Hyperparamètres de régularisation	-
$\mathbf{z}^l, \mathbf{a}^l$	Vecteur de sortie d'un réseau avant et après transformation par la fonction d'activation	-
$b$	Base de Halton	-
$C_b$	Suite de Van der Corput en base $b$	-
$c_i$	Coordonnée du $i^{me}$ point d'une suite de Van der Corput	-
$\mathbf{c}_k$	Vecteur de Halton	-
$e_x$	Erreur entre la sortie théorique et calculée du réseau sur le jeux de données $x$	-
$\mathbf{e}$	Vecteur contenant l'ensemble des erreurs $e_x$ de la base	-
$f_v$	Fraction volumique de suie	$\text{m}^3 \text{m}^{-3}$
$\gamma$	Nombre effectif de paramètres	-
$\mathbf{h}_i$	Vecteur coordonnées du $i^{me}$ point d'une suite de Halton	-
$\mathbb{I}$	Espace des phases de la grandeur d'intérêt	-
$\mathbf{J}$	Matrice jacobienne	-
$L_\eta$	Luminance monochromatique	$\text{W m}^{-2} \text{sr}^{-1} \text{cm}$

Symbole	Description	Unité
$n_p$	Nombre de paramètres du réseau	-
$\Omega$	Ensemble des points de référence	-
$P$	Pression	bar
$\mathbf{p}_x$	Vecteur caractéristique d'un point de référence	-
$\mathcal{S}$	Espace des solutions	-
$T$	Température	K
$\mu$	Taux d'apprentissage par époque	-
$\mathbf{v}$	Vecteur contenant l'ensemble des paramètres (poids et biais) du réseau	-
$\mathbf{x}$	Cordonnées spatiales	-
$x_{H_2O}, x_{CO_2}$	Fraction molaire des espèces participatives	-

### 3.1 Rappel du constat et proposition

Après la description détaillée du chapitre précédent, nous nous intéressons ici à la construction de la base de données d'entraînement du réseau. Au delà de la performance de la méthode d'apprentissage choisie, cette étape est fondamentale pour l'obtention d'un réseau entraîné efficace et dans lequel nous pouvons avoir confiance. Dans la section 1.3.2, nous avons montré que les méthodes par ordonnées discrètes présentaient des limites pour le calcul en champ du rayonnement thermique notamment vis à vis des ressources de calcul nécessaires.

Dans le chapitre 2, nous présentons les réseaux de neurones comme un outil d'interpolation complexe, qui se base sur des points théoriques connus, sur un savoir pré-établi en quelque sorte, afin de pouvoir interpoler des comportements ou des grandeurs à des situations non connues. Contrairement aux méthodes de résolutions globales, nous proposons ici de construire notre base de données grâce à une résolution par algorithme de Monte Carlo. Bénéficiant des avancées méthodologiques et informatiques, la méthode est bien maîtrisée dans les applications traitant du rayonnement thermique en milieux participants [91, 92] et est considérée comme la méthode de référence dans la plupart des publications.

L'extension de cette méthode aux milieux participants hétérogènes par la formulation en collision nulle [7, 93, 8] (voir section 1.4.2) permet d'étendre son utilisation à des problématiques réelles et complexes sans approximation dans sa formulation mathématique. Néanmoins, cette méthode n'est pas adaptée aux calculs en champ mais bien à celui de grandeurs radiatives intégrées à des points précis du domaine d'étude. Son utilisation exclusive sur des maillages de plusieurs centaines de millions de mailles est donc souvent inenvisageable au vu du temps et des ressources de calcul nécessaires.

L'utilisation des réseaux de neurones en support de cette méthode stochastique permet de dépasser cette limitation et de produire d'excellents résultats proches des calculs de références menés uniquement par Monte Carlo [94]. Nous présentons dans ce chapitre et en détails, la procédure mise en place et les différents modèles utilisées pour la modélisation du rayonnement thermique.



## 3.2 Interfaçage des méthodes : Monte Carlo pour la base de données et IA pour interpolation

### 3.2.1 Un modèle réduit statistique : quelques analogies

Notre proposition se base sur les capacités de généralisation puissantes des réseaux de neurones. Contrairement à la plupart des autres méthodes d'interpolation, les ANN sont en effet capables de prendre en compte une quantité illimitée de paramètres non homogènes, sous réserve de la puissance de calcul suffisante. L'apprentissage permet ensuite de construire des corrélations statistiques avancées entre les différentes données à disposition afin de pouvoir interpoler des résultats sur des cas inconnus. En ce sens, la méthode que nous proposons ici est fondamentalement ce qu'on peut appeler un modèle réduit car chaque réseau de neurones est construit et entraîné spécifiquement pour le cas d'étude à traiter. Ces outils n'ont donc en aucun cas la prétention de pouvoir résoudre le rayonnement par simple combinaison linéaire de plusieurs unités de calcul mais bien de proposer des corrélations complexes adaptées et optimisées à chaque problème.

De ce point de vue, la méthode proposée ici peut être rapprochée des méthodes de Monte Carlo symbolique. Différents travaux ont été menés sur ces thématiques dont l'idée globale revient à définir une fonction de transfert paramétrique, construite à partir de réalisations de Monte Carlo. L'objectif est de pouvoir calculer des observables physiques directement grâce à cette fonction de transfert explicite de par ses dépendances à différents paramètres de la simulation. Dans les travaux de PENAZZI et al. [95], la fonction de transfert construite par Monte Carlo permet de calculer la température en un point dans le domaine et pour un point de fonctionnement, c'est à dire pour une valeur du symbole choisi. Nous supposons que l'ensemble des autres conditions thermophysiques ou géométriques sont inchangées. Cette fonction fonctionne alors comme un extrapolateur qui permet de donner une estimation de la température pour des valeurs du symbole s'éloignant du point de fonctionnement. La confiance dans la valeur obtenue est maîtrisée grâce au contrôle de la barre d'erreur, en fonction de la valeur du symbole.

Chez GALTIER et al. [96], l'idée diffère quelque peu dans le sens où la grandeur modélisée n'est plus la température mais la luminance. L'ETR est reformulée et approximée sous forme polynomiale avec le coefficient d'absorption  $\kappa_\eta$  comme symbole (pour un cas purement absorbant) et les coefficients  $a_{j,i}$  déterminés par réalisations de Monte Carlo :

$$L_\eta(\mathbf{x}_0, \mathbf{u}_0) \approx \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{M_i} a_{j,i} \kappa_\eta^j \quad (3.1)$$

Dans les deux cas, l'idée sous-jacente derrière ces méthodes est la même : être capable sur la base de calculs stochastiques, de proposer un modèle réduit de l'observable recherchée qui serait paramétré par les données thermophysiques du problème. Ce que nous proposons de faire s'inscrit dans cette démarche mais avec une approche plus systémique. La fonction de transfert implicite, modélisée par le réseau de neurones, permet dans notre cas de prédire une grandeur radiative sur la base de simulations connues *dans l'ensemble* du domaine étudié. Ainsi, dans notre cas, les paramètres ne portent plus seulement sur les données physico-chimique mais s'étendent aussi sur l'espace géométrique. Cette démarche permet donc à terme de proposer non plus un modèle réduit de la grandeur d'intérêt en un point mais bien d'une solution approchée en champ, spécifique au cas sur lequel le réseau aura été entraîné. Ce changement d'échelle se paye dans notre cas par un cadre théorique moins rigoureux dans la mesure où les travaux précédemment cités sont issus d'une reformulation, parfois approchée, du problème physique. Dans les deux cas par exemple, il reste possible de donner une estimation de l'écart type sur les résultats. Notre

modèle se rapproche plutôt d'une méthode d'interpolation se basant sur des méthodes d'approximations mathématiques que sur la physique du problème.

### 3.2.2 Procédure : du calcul stochastique par point à l'interpolation en champ

Cette section est dédiée à la description complète de la méthode développée dans le cadre de cette thèse. Elle s'articule autour de trois axes :

- La construction de la base de données par méthode de Monte Carlo
- L'entraînement du réseau de neurones
- L'interpolation de l'observable d'intérêt

Il est important de préciser en préambule que nous considérons le champ thermophysique sur lequel nous travaillons comme entièrement connu. Dans la pratique, il peut être paramétrique pour une étude de cas académique (voir chapitre 4) ou bien issu d'une simulation CFD. Le couplage avec un éventuel solveur ne sera pas détaillé dans cette partie.

**1<sup>ère</sup> étape : Construction de la base de données** Avant toute utilisation d'un réseau de neurones, il est au préalable nécessaire de construire une base de données représentative du problème à modéliser. Elle sera constituée d'un ensemble de points de référence  $\Omega$ , dépendant des données thermophysiques dont est fonction l'observable d'intérêt :

$$\Omega = \{\mathbf{p}_x(\mathbf{i}, t) \mid \mathbf{i}, t \in \mathbb{I}, \mathbb{S}\} \quad (3.2)$$

Avec  $\mathbb{I}$  et  $\mathbb{S}$ , respectivement l'espace des phases de l'observable d'intérêt et l'espace des solutions. Dans ce cas présent, le vecteur  $\mathbf{i}$  est composé des coordonnées spatiales du point de référence  $\mathbf{p}_x$  et des données thermophysiques en ce point ayant une influence sur la quantité radiative d'étude :

$$\mathbf{i} = (\mathbf{x}, T(\mathbf{x}), P(\mathbf{x}), x_{CO_2}, x_{H_2O}, x_{CO}, f_v) \quad (3.3)$$

*i* Comme évoqué dans le chapitre 2 et en cohérence avec les notations déjà utilisées,  $t$  est ici un scalaire dans la mesure où nous ne nous intéressons qu'à une quantité radiative par point. Dans les résultats présentés dans le chapitre suivant, il s'agira soit du flux radiatif à la paroi, soit de la divergence du flux dans le domaine.

Les points de référence sont tout d'abord échantillonnés dans le domaine d'étude (voir section 3.2.3). La taille de la base de données dépend du cas d'étude, de ses caractéristiques géométriques ainsi que de la complexité du champs thermophysique. Ce point sera discuté au cas par cas dans le chapitre 4. Une fois la liste initiale des points établie, un calcul de Monte Carlo en chacun de ces points est réalisé afin de calculer la grandeur d'intérêt selon les équations (1.25) et (1.26). Notons que les modèles de gaz ou de suies permettant de calculer le coefficient d'absorption n'influent pas sur la procédure. Afin de traiter au mieux les hétérogénéités, la formulation en collisions nulles présentée dans la section 1.4.2 est utilisée.

A ce stade, nous disposons alors d'une liste de points de référence, calculés par méthode de Monte Carlo sans approximation à l'exception du modèle de coefficient d'absorption choisi. En considérant que la convergence des calculs soit suffisante, c'est à dire que la variance sur les points de référence soit

bonne, nous disposons d'une base de données solide, représentative du problème. Nous montrerons dans le chapitre 4 que les réseaux de neurones bayésiens utilisés parviennent par ailleurs, à gérer particulièrement bien le bruit numérique sur les données d'entraînement. Par conséquent, il nous sera possible par soucis d'efficacité numérique de réduire grandement le nombre de réalisations de Monte Carlo tout en conservant de très bons résultats interpolés.

**2<sup>ème</sup> étape : entraînement du réseau** Une fois la base de données construite, la deuxième phase peut commencer. Nous suivons la procédure décrite en détail dans le chapitre 2, reprise et complétée dans la figure 3.2. Avant tout, l'ensemble des données est normalisé par une méthode *Min-Max* de telle sorte que les entrées et sorties du réseau soient comprises entre -1 et 1. Nous utilisons en effet une fonction d'activation de type *tanh* dont les droites d'équation  $y = -1$  et  $y = 1$  sont des asymptotes. Nous nous assurons de la sorte à ne pas saturer les réseaux par des valeurs trop élevées en valeurs absolues.

Un réseau de neurones bayésien est ensuite entraîné. Nous utilisons un algorithme de Levenberg-Marquardt dans sa formulation dans le paradigme bayésien (voir section 2.3.3). Le nombre de neurones initiaux dépend là encore de la situation. Néanmoins, au vu des résultats obtenus dans le chapitre 4, un total de 10 unités dans la couche cachée est une estimation qui semble globalement pertinente pour la prédiction de la divergence du flux radiatif en condition hétérogène par exemple. Quoiqu'il en soit, l'algorithme adaptera la topologie pendant l'entraînement en fonction du nombre effectif de paramètres  $\gamma$ . Une bonne initialisation du réseau peut néanmoins permettre une convergence plus rapide.

En cas d'ajout ou de suppression de neurones, on constate qu'environ 500 époques suffisent généralement pour que  $\gamma$  converge vers une valeur représentative du nombre de paramètres nécessaires (voir chapitre 4). Une valeur de  $\gamma$  comprise entre 80 et 90 % du nombre total de paramètres est généralement recherchée. Cet ordre de grandeur n'est évidemment pas absolu. Si la valeur de  $\gamma$  est constante avec l'augmentation du nombre de neurones, cela signifie que les unités ajoutées sont superflues. Dans la pratique, ce principe se traduit ainsi en terme de critère par une valeur de  $\gamma$  comprise dans cet intervalle. En dehors de celui-ci, il est soit nécessaire d'ajouter des neurones (si  $\gamma \approx n_p$ ), soit bienvenu d'en supprimer (si  $\gamma \ll n_p$ ).

**3<sup>ème</sup> étape : interpolation multivariante sur le maillage** Une fois entraîné, le réseau de neurone peut finalement être considéré d'un certain point de vue comme la représentation implicite d'une fonction de transfert complexe permettant d'estimer la grandeur d'intérêt dans le domaine. A partir de quelques points de référence, par nature ponctuels, nous avons donc construit un modèle continu dans l'espace des phases de l'observable. Plus particulièrement et en supposant avoir accès aux données en tout point du domaine, nous sommes capables d'interpoler la valeur de l'observable en tout point de l'espace géométrique.

Dans la pratique, cette situation n'est jamais rencontrée dans la mesure où ce modèle vise à être couplé à un solveur CFD et donc, à être appliqué sur un maillage surfacique ou volumique. Très concrètement, il s'agit de faire ce que nous avons appelé dans le chapitre 2, une *FeedForward Pass* sur le réseau en chaque nœud du maillage. Néanmoins, l'interpolation étant quasiment instantanée une fois le réseau entraîné, le modèle réduit est quasiment insensible à la qualité du maillage (voir chapitre 4). Dans le même ordre d'idée, il a été montré que le calcul des points sondes par méthode de Monte Carlo est également insensible au niveau de discrétisation spatiale de l'espace. Par conséquent et contrairement à une méthode DOM ou Monte Carlo complète (c'est à dire calculant stochastiquement l'observable en chaque point du maillage), la procédure présentée ici ne consomme pas plus de ressources quelle que soit la complexité géométrique du problème.

### 3.2.3 Tirage des points de référence

Dans le processus de construction de la base de données, nous évoquons le fait que les points de référence sont tirés dans le domaine. Néanmoins, aucun commentaire n'a été fait sur la méthode utilisée. Nous précisons ce point dans cette section.

Nous proposons dans ces travaux deux méthodes de sélection des points d'entraînement. La première est certainement la plus simple et la plus naturelle. Elle consiste simplement à tirer aléatoirement ces points dans le domaine. La seconde est plus efficace mais plus difficilement adaptable aux géométries complexes. Elle correspond à un tirage déterministe selon une suite à discrédance faible.

**Tirage aléatoire** Il s'agit ici simplement de tirer aléatoirement dans le domaine les points de référence. Cela dit, précisons tout de même que si cette méthode semble *a priori* immédiate, sa mise en œuvre n'est pas forcément triviale.

En effet, un tirage aléatoire *fini* n'implique pas nécessairement un tirage uniforme ou tout du moins, ne garantit pas de minimiser la discrédance entre chacun des points [97]. Beaucoup de questions peuvent alors se poser sur le tirage le plus efficace, impliquant par extension le moins de points de référence possible pour une interpolation de qualité. Cette problématique reste une question ouverte qui n'a pas pu être traitée en profondeur dans le cadre de ces travaux. Quelques pistes néanmoins peuvent être données.

Tout d'abord et bien que ceci puisse paraître contre-intuitif, il n'est pas nécessairement évident qu'un point de référence tiré dans un endroit où l'observable est relativement constante, soit toujours moins intéressant qu'un point dans une zone très hétérogène. La présence de zones blanches est en effet toujours une prise de risque dans la mesure où rien n'indique au réseau de neurones les tendances à suivre en l'absence d'information, quand bien même celles-ci seraient simples et lisses. Le risque d'interpolation aberrante grandit alors sans qu'il ne soit possible de les détecter facilement par des indicateurs d'entraînement comme la *MSE*. Notons que la régularisation permet malgré tout de limiter assez fortement ce phénomène.

La pertinence des tirages est donc un vrai sujet et mérite une étude de fond. Nous pouvons par exemple imaginer qu'un tirage préférentiel sur les zones à forts gradients de température ou de concentration en espèces rayonnantes puisse améliorer la convergence et l'efficacité de l'outil. L'ajout des points extrêmes à la base de données est également une idée intéressante au risque sinon, de saturer les neurones de calcul. En effet, si les bornes de normalisation des paramètres n'encadrent pas l'ensemble des valeurs possibles, ce phénomène peut alors se produire. Ce dernier point a été étudié et les résultats mettent en exergue les conséquences qu'une mauvaise normalisation peut engendrer sur la qualité des interpolations (voir section 4.4.1).

**Tirage selon une suite à discrédance faible de Halton** La deuxième méthode investiguée dans ces travaux consiste à tirer les points de référence non plus de manière aléatoire mais selon une suite à discrédance faible de Halton [97]. Ces suites mathématiques optimisent la distribution de points au sein d'un hypercube unité de dimension quelconque dans le but de réduire leur discrédance. Ce terme mesure en quelques sortes, la disproportion de la distribution de points dans un espace. Dit autrement, la discrédance d'une suite de points sera haute s'ils ne sont pas *correctement* répartis au sein de l'espace considéré (voir figure 3.1).

Bien que la théorie mathématique soit particulièrement complexe pour mesurer cette grandeur [98], il existe des suites simples à programmer qui par définition minorent la discrédance comme celle de

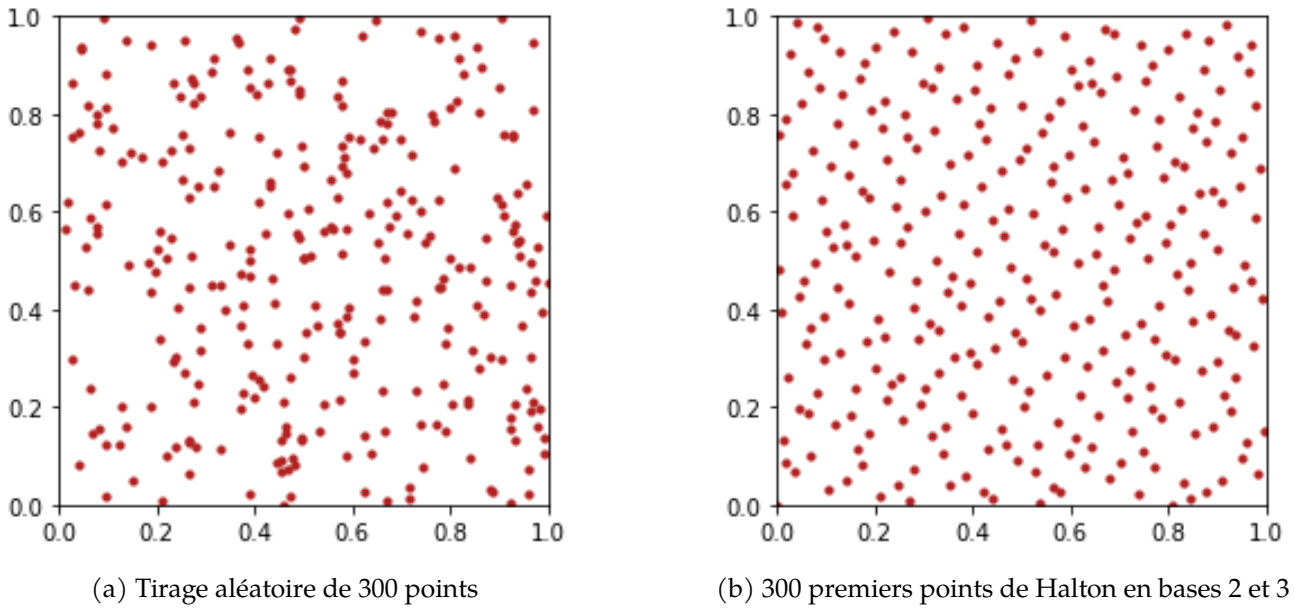


FIG. 3.1 : Comparaison de la disparité des points dans un carré unité tirés aléatoirement ou suivant une séquence de Halton

Halton. Ce type de suites est d'ailleurs central pour les algorithmes de *Quasi Monte Carlo* [99]. Les suites de Halton en particulier, sont une version étendue des suite de Van der Corput en dimension  $N$ .

Ce type de distribution permet de supprimer les potentielles zones blanches pouvant apparaître avec un tirage aléatoire. L'ensemble du domaine est ainsi couvert au mieux par des points de référence dans l'ensemble des dimensions spatiales. On note ainsi une réelle amélioration de la qualité de l'apprentissage grâce à ce type de tirage (voir chapitre 4) par rapport à un tirage aléatoire uniforme par exemple. Néanmoins, cette méthode est difficilement applicable à des géométries complexes et en particulier à une chambre de combustion du fait de son domaine de définition. Cette méthode a été appliquée pour un cylindre axisymétrique dans le chapitre 4 avec succès en mettant à l'échelle les différentes coordonnées mais n'a pas pu être transposé à un cas industriel jusqu'à présent. Cet exemple a néanmoins le mérite de démontrer l'importance d'une base de données bien construite pour un apprentissage complet et efficace.

### Construction d'une suite de Halton

Pour construire une suite de Van der Corput, il est tout d'abord nécessaire de choisir une base  $b$  telle que  $b \in \mathbb{Z}$ . On définit  $i$  l'indice du point de la suite. Le vecteur  $\mathbf{c}_k$  correspond à l'indice  $i$  exprimé dans la base  $b$ . Pour simplifier la compréhension et l'écriture, nous supposons que  $i \in [0, 26]$ . Ainsi, en base 3, l'écriture de  $i$  peut s'écrire comme un vecteur de dimension 3 :

$$\mathbf{c}_k = (i_0, i_1, i_2)^{(3)} \text{ avec } i = \sum_{j=0}^{b-1} \mathbf{c}_k(j) \cdot 3^j \quad (3.4)$$

On a ainsi dans cet exemple :

$$i = 0 \text{ s'écrit } \mathbf{c}_k = (0, 0, 0) \Leftrightarrow i = 0 \cdot 3^0 + 0 \cdot 3^1 + 0 \cdot 3^2 \quad (3.5)$$

$$i = 1 \text{ s'écrit } \mathbf{c}_k = (1, 0, 0) \Leftrightarrow i = 1 \cdot 3^0 + 0 \cdot 3^1 + 0 \cdot 3^2 \quad (3.6)$$

$$i = 5 \text{ s'écrit } \mathbf{c}_k = (2, 1, 0) \Leftrightarrow i = 2 \cdot 3^0 + 1 \cdot 3^1 + 0 \cdot 3^2 \quad (3.7)$$

$$i = 22 \text{ s'écrit } \mathbf{c}_k = (1, 1, 2) \Leftrightarrow i = 1 \cdot 3^0 + 1 \cdot 3^1 + 2 \cdot 3^2 \quad (3.8)$$

Les points  $c_i$  d'une suite de Van der Corput  $C_b = \{c_0, c_1, \dots\}$  en base  $b$ , repérés par leur indice  $i$ , se calculent ensuite par la formule explicite suivante :

$$c_i = \sum_{j=0}^{\infty} \mathbf{c}_k(j) \cdot b^{-(j+1)} \quad (3.9)$$

Dans cette expression, nous faisons apparaître une somme infinie. En pratique, cette somme s'arrête en fonction du nombre de *digit* nécessaire à la description de  $i$  dans la base choisie. Au delà, les termes de la somme seront toujours égaux à zéro. En nous basant sur cette définition, les premiers termes d'une somme de Van der Corput en base 3 sont donc  $C_3 = \{0, \frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \dots\}$ . L'extension des suites de Van der Corput à des dimensions supérieures est simple en pratique. En deux dimensions par exemple, les coordonnées d'un point  $\mathbf{h}_i$  suivant une séquence de ce type seront toutes les deux issues d'une suite de Van der Corput dans deux bases différentes tel que :

$$\mathbf{h}_i = (c_i^{(b_1)}, c_i^{(b_2)}) \mid b_1 \neq b_2, (b_1, b_2) \in \mathbb{Z}^{*2} \quad (3.10)$$

C'est la définition d'une séquence de Halton en deux dimensions dont il est prouvé qu'il s'agit d'une suite à discrétion faible [100]

## 3.3 Synthèse de la méthode

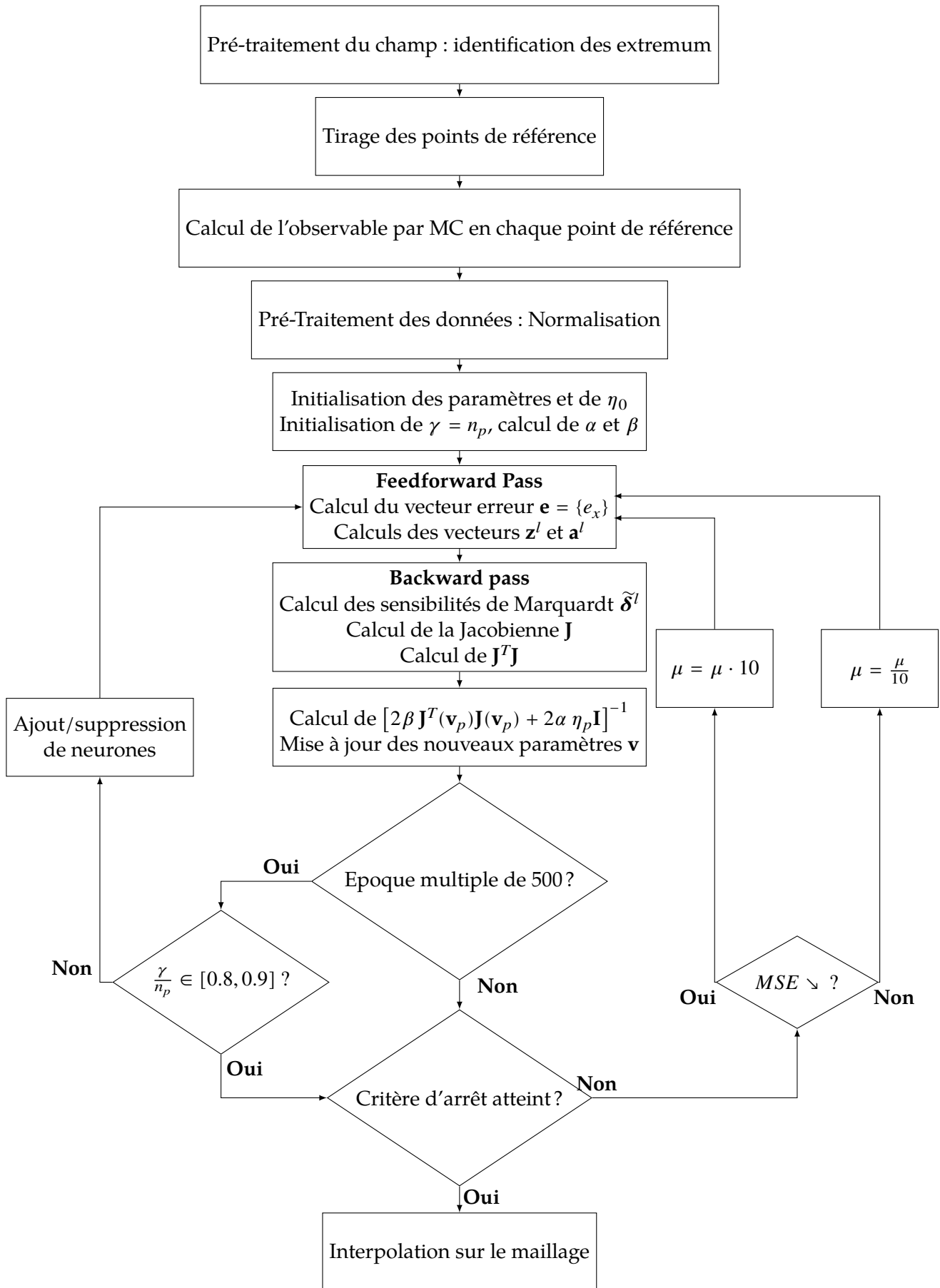


FIG. 3.2 : Logigramme complet de l'algorithme développé

Le logigramme présenté dans la figure 3.2 reprend la procédure complète de calcul développée pendant ces travaux de thèse. Nous proposons ici quelques commentaires :

- Dans un premier temps, il est nécessaire d'identifier dans le champ thermophysique les extremum en température et en espèces rayonnantes pouvant mener à des valeurs majorantes de la grandeur radiative à modéliser. Cette étape est nécessaire pour une normalisation efficace de la base de données et une interpolation de qualité.
- L'étape suivante concerne le tirage des points de référence qui constitueront la base de données d'entraînement. Les extremums précédemment identifiés devront être intégrés à cette base. Une fois les points d'entraînement échantillonnés et les données thermophysiques en ces points relevés, les valeurs correspondantes de la grandeur d'intérêt sont calculées par méthode de Monte Carlo.
- La base de données une fois contruite, les entrées et les sorties sont normalisées selon leurs extremum propres.
- A partir de cette étape et après initialisation des différents hyperparamètres, la base de données est prête et l'algorithme d'entraînement peut commencer. Les *Feedforward Pass* et *Backward Pass* ainsi que le calcul des nouveaux paramètres du réseaux suivent la procédure décrite dans la section 2.3.3.
- Toutes les 500 itérations, nous contrôlons la valeur du nombre effectif de paramètres  $\gamma$ . Ce nombre d'époque suffit généralement pour atteindre une convergence de ce paramètre. Si la valeur de  $\gamma$  sort des critères proposées dans la section 2.3.3, le nombre de neurones est ajusté à la hausse ou à la baisse puis l'algorithme d'entraînement reprend.
- Le dernier test consiste à contrôler si les critères de convergence sont atteints ou non. Ils peuvent concerner la valeur de la MSE, de la RMSE, du taux d'apprentissage  $\mu$  ou encore le nombre d'itérations. Les besoins peuvent être différents selon les usages. Si aucun des critères n'est atteint, la valeur du taux d'apprentissage est ajusté en fonction de la comparaison entre la valeur de la MSE calculée avec les nouveaux paramètres du réseau et les anciens (voir la procédure dans la section 2.3.3 et le chapitre 13 de HAGAN et al. [69]).

Une fois l'algorithme terminé, le réseau de neurones est alors opérationnel pour interpoler sur le maillage une valeur de la grandeur radiative d'intérêt en chaque cellule, d'après les données thermophysiques hétérogènes en chaque noeud. L'ordre de grandeur en terme de temps de calcul d'un tel algorithme se situe autour de quelques minutes. La durée de l'entraînement dépend évidemment très fortement du nombre d'époques maximum choisi et du nombre de points d'entraînement dans la base de données. Un ensemble de valeurs représentatives est présenté dans le chapitre 4 pour différents cas d'application.

## Références

---

- [7] M. GALTIER et al. «Integral formulation of null-collision Monte Carlo algorithms». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 125 (2013), p. 57-68. DOI : [10.1016/j.jqsrt.2013.04.001](https://doi.org/10.1016/j.jqsrt.2013.04.001).
- [8] M. EL HAFI et al. «Three viewpoints on null-collision Monte Carlo algorithms». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 260 (2021).
- [69] M. T. HAGAN et al. *Neural Network Design*. 2<sup>e</sup> éd. Boston : PWS Publishing Co., 1996.



- 
- [91] M. F. MODEST. «Narrow-band and full-spectrum k-distributions for radiative heat transfer—correlated-k vs. scaling approximation». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 76.1 (2003), p. 69-83.
- [92] J. R. HOWELL et K. J. DAUN. «The Past and Future of the Monte Carlo Method in Thermal Radiation Transfer». In : *Journal of Heat Transfer* 143.10 (2021).
- [93] V. EYMET et al. «Null-collision meshless Monte-Carlo - Application to the validation of fast radiative transfer solvers embedded in combustion simulators». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 129 (2013), p. 145-157.
- [94] A. ROYER et al. «A new method for modeling radiative heat transfer based on Bayesian artificial neural networks and Monte Carlo method in participating media». In : *International Journal of Heat and Mass Transfer* 201 (2023), p. 123610.
- [95] L. PENAZZI et al. «Toward the use of Symbolic Monte Carlo for Conduction-Radiation Coupling in Complex Geometries». In : *RAD-19 - 9th International Symposium on Radiative Transfer*. Athens, Greece : Begellhouse, 2019, 8 p. DOI : [10.1615/RAD-19.380](https://doi.org/10.1615/RAD-19.380).
- [96] M. GALTIER et al. «A symbolic approach for the identification of radiative properties». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 196 (2017), p. 130-141. DOI : [10.1016/j.jqsrt.2017.03.026](https://doi.org/10.1016/j.jqsrt.2017.03.026).
- [97] L. KOCIS et W. J. WHITEN. «Computational investigations of low-discrepancy sequences». In : *ACM Transactions on Mathematical Software* 23.2 (1997), p. 266-294. DOI : [10.1145/264029.264064](https://doi.org/10.1145/264029.264064).
- [98] E. THIÉMARD. «Sur le calcul et la majoration de la discrèpance à l’origine». Thèse de doct. Ecole Polytechnique Fédérale de Lausanne, 2000.
- [99] L. PALLUOTTO et al. «Assessment of randomized Quasi-Monte Carlo method efficiency in radiative heat transfer simulations». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 236 (2019). DOI : [10.1016/j.jqsrt.2019.07.013](https://doi.org/10.1016/j.jqsrt.2019.07.013).
- [100] J. H. HALTON. «On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals». In : *Numerische Mathematik* 2 (1960), p. 84-90. DOI : [10.1007/BF01386213](https://doi.org/10.1007/BF01386213).



# Evaluation de la méthode sur cas académiques

## Sommaire du chapitre

4.1	Choix des modèles physiques et du réseau de neurones . . . . .	82
4.2	Etude d'un cas axisymétrique homogène de Benchmark . . . . .	84
4.3	Cas hétérogène : étude d'une flamme de diffusion simplifiée . . . . .	86
4.4	Prise en compte de perturbations . . . . .	98
4.5	Ouverture et perspective . . . . .	106
	Références . . . . .	109

## Notations

Symbole	Description	Unité
$\alpha, \beta$	Hyperparamètres de régularisation	-
$f_v$	Fraction volumique de suie	$\text{m}^3 \text{m}^{-3}$
$\gamma$	Nombre effectif de paramètres	-
$\kappa_{abs,\eta}$	Coefficient d'absorption monochromatique	$\text{m}^{-1}$
$l, r$	Coordonnées axiale et radiale	m
$L, R$	Longueur et rayon du cylindre d'étude	m
$l_{PC}, r_{PC}$	Coordonnées axiale et radiale du centre du point chaud	m
$\mathcal{N}(\mu_n, \sigma)$	Loi normale centrée en $\mu_n$ d'écart type $\sigma$	-
$q$	Flux radiatif	$\text{W m}^{-2}$
$\nabla \cdot q$	Divergence du flux radiatif	$\text{kW m}^{-3}$
$T$	Température	K
$T_{PC}$	Température au point chaud	K
$\mu$	Taux d'apprentissage par époque	-
$x_E$	Fraction molaire de l'espèce $E$	-
$x_{E,PC}$	Fraction molaire au Point chaud de l'espèce $E$	-

## 4.1 Choix des modèles physiques et du réseau de neurones

Ce chapitre est consacré à la présentation de résultats donnés par l'application de l'outil développé sur différents cas d'étude. Dans les chapitres 2 et 3, il n'a jusqu'ici pas été question de choisir un modèle de gaz en particulier dans la mesure où la méthode présentée dans ces travaux est indépendante de la méthode de calcul du coefficient d'absorption  $\kappa_{abs,\eta}$ . Cette section introductive vise à présenter les modèles utilisés dans le cadre des applications étudiées dans la suite de ce chapitre et d'en justifier le choix.

### 4.1.1 Modélisation du coefficient d'absorption

**A pression atmosphérique** Bien que la méthodologie développée ici ait pour objectif d'être appliquée à des chambres de combustion, dans des conditions de pression élevée (autour de 20 bars), elle a été dans un premier temps validée sur cas académiques à pression atmosphérique par comparaison avec des résultats de benchmark issus de la littérature. Les configurations étudiées sont tirées des travaux de COELHO, PEREZ et EL HAFI [2] pour la validation de la méthode. L'objectif est de pouvoir comparer les résultats obtenus par notre méthode d'interpolation basée sur les réseaux de neurones avec des données éprouvées indépendamment du modèle de gaz. Dans ce papier, les résultats de référence calculés à partir d'un algorithme de Monte Carlo utilisant une formulation de l'ETR en coefficient d'absorption ont été construits à partir d'un modèle SNB-CK [47, 53, 54]. Ce modèle de gaz a donc été choisi pour la construction de la base de données d'entraînement.

Au delà de sa précision proche d'un modèle SNB classique [42], ce modèle présente d'autres avantages intéressants par rapport à sa facilité d'implémentation et sa gestion simplifiée des mélanges gazeux [93, 101] (voir section 1.4.4). Il est donc adapté à la modélisation du rayonnement dans des environnements proches des champs de combustion réels, tant en termes de niveaux de température rencontrés que de mélange de différentes espèces rayonnantes. Nous avons utilisé une quadrature de Gauss-Legendre à 7 points pour déterminer le coefficient d'absorption de chaque gaz sur les bandes spectrales échantillonnées. La valeur des différents  $\kappa_{abs,\eta}$  peut ensuite être déterminée par l'inversion de la fonction cumulée  $g(k)$  grâce à une méthode de Newton-Raphson [50]. La version du modèle SNB-CK utilisée comporte 367 bandes spectrales d'une largeur de  $25 \text{ cm}^{-1}$ . Les paramètres sont tirés des travaux de RIVIÈRE et SOUFIANI [43].

**A haute pression** Pour les cas d'étude à haute pression, nous utiliserons le modèle HPBM développé par PIERROT [1] à 367 bandes de  $25 \text{ cm}^{-1}$ . Cette version fait aujourd'hui office de référence pour les applications Safran. Au delà de sa précision et de la prise en compte des trois espèces rayonnantes principales ( $\text{CO}_2$ ,  $\text{H}_2\text{O}$  et  $\text{CO}$ ) en combustion, ce modèle est également facilement implémentable et peut être utilisé sur une très large plage de pression allant de 5 à 40 bars.

**Contribution des suies** La participation des suies est prise en compte selon le modèle semi-empirique présenté dans la section 1.4.4. La diffusion des particules solides de suies est donc négligée dans ces travaux.

i

Chacun de ces modèles a été implémenté spécifiquement au sein du code Monte Carlo utilisé dans le cadre de ces travaux. Alors que l’algorithme d’entraînement des réseaux de neurones a été entièrement développé dans le cadre de cette thèse, nous avons en revanche collaboré activement avec la société spécialisée *Méso-star* pour le code Monte Carlo. Plus spécifiquement, nous sommes partis du code HTRDR [10] basé sur la librairie *star-engine*<sup>a</sup> que nous avons ensuite modifié en profondeur pour traiter nos problématiques propres.

<sup>a</sup><https://gitlab.com/meso-star>

### 4.1.2 Réseaux de neurones utilisés

Bien qu’une description détaillée du type de réseaux utilisés ait été faite dans le chapitre 2, nous proposons un résumé du paramétrage de cet outil appliqué à nos travaux. La topologie des réseaux utilisés est du type *FeedForward* à une seule couche cachée. Le nombre de neurones dans cette couche est initialisé ici à 10 unités élémentaires, toutes interconnectées avec les neurones d’entrée et de sortie. Les fonctions d’activation sont des tangentes hyperboliques.

L’ensemble de ces paramètres de configuration assure le caractère universel de l’interpolateur [67]. La méthode consiste à entraîner le réseau pour chaque champ thermophysique donné. Il est donc important de garantir un temps d’entraînement relativement faible pour conserver l’intérêt de l’outil sur des cas réels et à plus forte raison, dans le cadre de calculs couplés. La topologie à une unique couche cachée et le faible nombre de paramètres à optimiser sont justifiés par cet objectif de performance.

Le nombre de variables d’entrée dépend du cas traité. Dans les exemples qui suivent, la fraction volumique de suie n’est en l’occurrence jamais un paramètre du réseau dans la mesure où celle-ci ne varie pas. Leur contribution est donc implicitement prise en compte dans le calcul Monte Carlo et donc dans les données d’entraînement. Dans le cas homogène à température, pression et concentration des espèces constantes, le réseau ne prend que les coordonnées du point sur lequel la quantité radiative d’intérêt est calculée ou interpolée. Dans les cas hétérogènes et pour le calcul du champ de divergence du flux radiatif, les concentrations en espèces rayonnantes et la température sont ajoutées aux coordonnées spatiales. La pression reste constante et n’est donc pas utilisée comme paramètre.

L’algorithme d’entraînement suit la procédure décrite dans le chapitre 13 de HAGAN et al. [69] et le chapitre 2 de ce manuscrit. Nous utilisons un algorithme d’optimisation de Levenberg-Marquardt dans le paradigme bayésien. La matrice hessienne relative aux données d’entraînement est calculée puis utilisée à la fois pour mettre à jour les poids et biais du réseau à chaque itération, mais également pour calculer les valeurs les plus pertinentes des hyperparamètres de régularisation  $\alpha$ ,  $\beta$ . La trace de cette matrice permet aussi de donner une indication sur le nombre de paramètres effectifs  $\gamma$  et de contrôler le dimensionnement du réseau par rapport au besoin imposé par la base de données utilisée. Toutes les 500 époques, la valeur de  $\gamma$  est contrôlée automatiquement : si elle appartient à l’intervalle d’acceptabilité donné dans le chapitre 2, l’entraînement continue ; sinon, le nombre de neurones est ajusté. La topologie à une couche cachée rend immédiate cette étape dans la mesure où la question de savoir sur quelle couche il faut ajouter ou supprimer un neurone ne se pose pas.

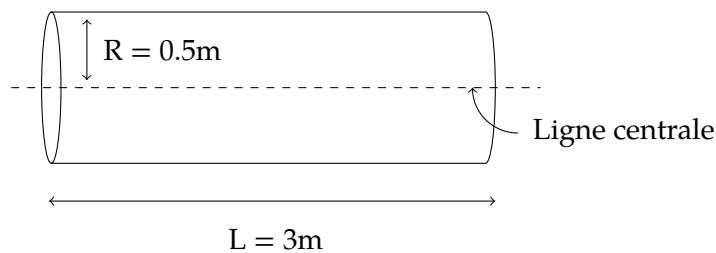


FIG. 4.1 : Géométrie étudiée - Cas homogène

## 4.2 Etude d'un cas axisymétrique homogène de Benchmark

Comme évoqué plus haut, les premiers cas homogène et hétérogène sont directement issus des travaux de benchmark de COELHO, PEREZ et EL HAFI [2]. L'objectif est de valider à la fois le code de calcul Monte Carlo, le modèle SNB-CK utilisé et le principe méthodologique développé dans ces travaux de thèse. Deux grandeurs radiatives sont étudiées ici, à savoir le flux à la paroi  $q$  exprimé en  $kW \cdot m^{-2}$  et la divergence du flux  $\nabla \cdot q$  en  $kW \cdot m^{-3}$  à l'intérieur du domaine (voir section 1.2).

La géométrie du domaine d'étude est une cavité axisymétrique de 3m de long et de 0.5m de rayon (voir figure 4.1). Les parois sont noires et leur température est fixée à 300 K. Le milieu est composé d'un mélange homogène composé de 20% de  $H_2O$ , de 10% de  $CO_2$  et de 70% de gaz transparent (ici du  $N_2$ ). La fraction volumique de suie est constante et égale à  $10^{-7} m^3 \cdot m^{-3}$ . La température est fixée à 1200 K. Les points de référence sont localisés par deux coordonnées spatiales au sein du cylindre à savoir leur abscisse  $l$  et leur distance à l'axe de révolution  $r$ .

### 4.2.1 Validation du code radiatif en milieu homogène

Dans un premier temps, l'algorithme de calcul par Monte Carlo est validé par comparaison des résultats avec les données du cas de benchmark de COELHO, PEREZ et EL HAFI [2] dans un cas homogène, sans formulation en collision nulle. Chaque point de référence a été calculé de telle sorte que l'écart type sur chaque résultat soit inférieur à 1%. Les graphiques présentés sur la figure 4.2 permettent de comparer les données de références calculées par Ray Tracing et modèle SNB de Malkmus [42] et les résultats obtenus par l'algorithme de Monte Carlo développé dans nos travaux avec un modèle SNB-CK. Une comparaison similaire de ces méthodes avait déjà été menée dans le papier de benchmark utilisé comme support et nous pouvons constater que les résultats obtenus se superposent pour un cas homogène. Notre méthode de calcul des points de références est donc validée dans un cas homogène à la fois pour le calcul du flux radiatif à la paroi et pour la divergence du flux dans le volume.

### 4.2.2 Validation du concept méthodologique : entraînement et interpolation par réseau de neurones

L'étude des capacités d'interpolation des réseaux de neurones dans deux cas homogènes est un bon exercice pour la validation du concept méthodologique proposé. Nous souhaitons à la fois démontrer le potentiel de cette méthode sur un cas simple ainsi qu'obtenir un premier ordre de grandeur des différents paramètres du réseau. L'objectif est notamment d'obtenir une première approximation sur le nombre de points d'entraînement nécessaires et sur la topologie optimale du réseau.

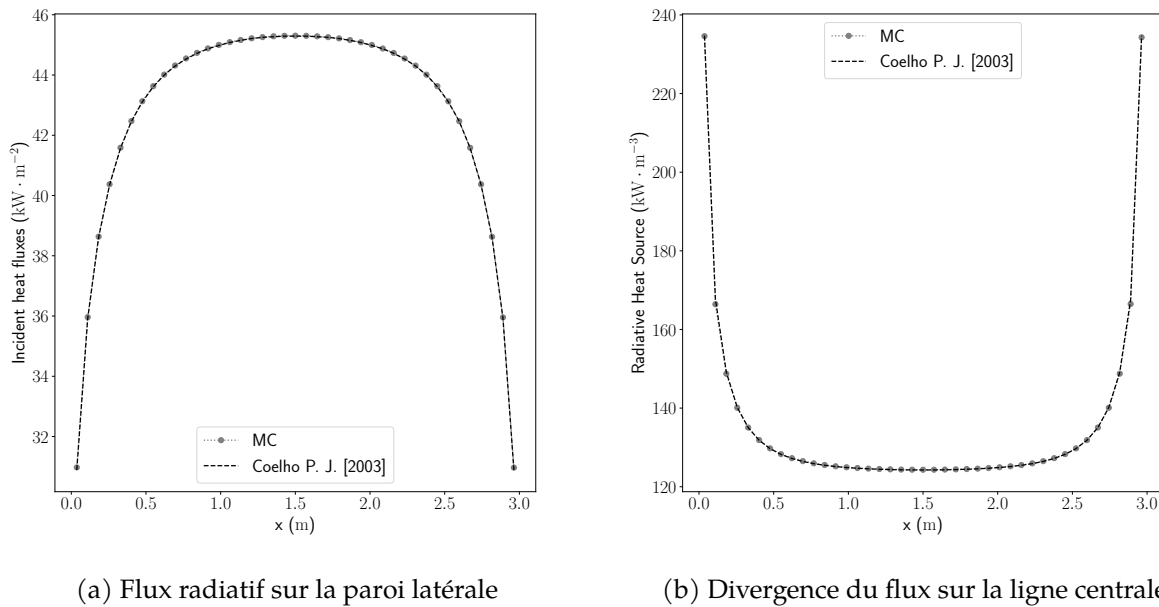


FIG. 4.2 : Validation de l'algorithme de calcul par Monte Carlo et SNB-CK sur cas homogène à une température de 1200 K- Comparaison avec les données de référence [2]

Dans ce cas, les points d'entraînement ont été échantillonnés aléatoirement et uniformément dans le domaine. Tous les paramètres thermophysiques étant égaux, les seules entrées du réseau sont les coordonnées spatiales  $l$  et  $r$  des points sondes. Le nombre d'époques maximal a été fixé à 10 000 dans ce cas de figure afin de permettre un certain nombre de corrections sur le nombre de neurones dans la couche cachée. L'autre critère d'arrêt utilisé porte sur la valeur du taux d'apprentissage  $\mu$  fixé à 1000. C'est en pratique un ordre de grandeur préconisé par HAGAN et al. [69] et utilisé par défaut dans différents logiciels spécialisés comme MATLAB. Au delà de cette valeur, nous pouvons considérer que l'algorithme est convergé.

Sans que la MSE (*Mean Squared Error*, voir équations (2.8) et (2.9)) ou les hyperparamètres  $\alpha$  et  $\beta$  (voir chapitre 2) soient des critères d'arrêt à proprement parler, un suivi de leurs valeurs est assuré pendant toute la phase d'entraînement des différents réseaux de neurones. Notons que pour tous les cas présentés dans ce manuscrit, une convergence sur ces paramètres a été atteinte dans la mesure où ils conditionnent le calcul du nombre effectif de paramètres  $\gamma$  (voir équation (2.59)).

Points d'entraînement	Neurones	Paramètres	$\gamma$	Temps d'entraînement	$\varepsilon_{moy}\%$	$\varepsilon_{max}\%$
100	11	45	36	26.44 s	0.86%	39.74%
200	11	45	37	49.93 s	1.53%	23.51%
500	13	53	43	132.41 s	0.23%	1.69%

TAB. 4.2 : Prédiction des réseaux de neurones à T = 1200 K

Le tableau 4.2 regroupe les erreurs moyennes  $\varepsilon_{moy}$  et maximales  $\varepsilon_{max}$  sur la divergence de flux obtenue par interpolation grâce à un réseau de neurones. Trois bases de données de 100, 200 et 500 points ont été testées. L'écart type sur les données d'entraînement calculées par méthode de Monte Carlo a été

contrôlé sur chaque point afin qu'il soit inférieur à 1%. Comme cela a déjà été mentionné précédemment, la topologie du réseau, et plus spécifiquement, le nombre de neurones dans la couche cachée est ajusté au fur et à mesure de l'entraînement grâce à la valeur du nombre de paramètres effectifs  $\gamma$ .

Nous pouvons par ailleurs noter que le besoin en nombre de neurones augmente dans ce cas avec le nombre de points d'entraînement. Ce comportement semble assez intuitif dans la mesure où les relations implicites entre les différents cas d'entraînement se complexifient avec le nombre de points de références. Cette tendance est d'ailleurs assez commune en matière de méthodes d'interpolation. Un parallèle peut par exemple être fait avec l'interpolation polynomiale langrangienne où le degré du polynôme d'interpolation augmente avec le nombre de points de passage. L'augmentation du nombre de points de référence implique donc l'augmentation rapide du temps et des ressources de calcul nécessaires car il est à la fois dépendant de la taille de la base de données et du nombre de neurones dans la couche cachée. La construction de la matrice Jacobienne, de la matrice Hessienne et le calcul de son inverse (voir section 2.3.2) sont en effet les opérations les plus chronophages et les plus lourdes de l'algorithme d'entraînement et sont immédiatement dépendantes du nombre de jeux d'entraînement et du nombre de paramètres à optimiser dans le réseau. Ainsi, pour garantir la performance de l'outil proposé, une attention toute particulièrement devra être portée au nombres de points de référence utilisés. Un compromis doit être trouvé entre temps d'entraînement et précision des résultats. Certaines applications requièreront certainement une résolution très fine du rayonnement (signature infra-rouge par exemple) auquel cas, un coût de calcul supplémentaire pourra se justifier. Dans d'autres situations, la recherche trop extrême de finesse sur une grandeur comme la divergence du flux tiendra plutôt de la « sur-qualité ».

Sur la figure 4.3, la divergence du flux radiatif a été estimée sur la ligne centrale du cylindre d'étude ( $r = 0$ ) par différents réseaux de neurones entraînés selon les trois bases de données disponibles dans le tableau 4.2. L'estimation de l'outil informatique est comparée aux résultats de référence donnés par COELHO, PEREZ et EL HAFI [2]. Les trois différentes courbes traduisent visuellement les informations contenues dans le tableau précédent. Nous constatons tout d'abord que la base de données contenant 100 points de référence n'est pas suffisamment fournie pour permettre une interpolation de qualité de la divergence du flux radiatif. La courbe relative à la base de données de 200 points donne en revanche de meilleurs résultats bien qu'un léger écart puisse être constaté au plus près des parois ( $l \approx 0m$  et  $l \approx 3m$ ). L'erreur relative maximale encore élevée s'explique par ces différences en comparaison de l'erreur moyenne très satisfaisante autour de 1.5%. Enfin, les prédictions de l'ANN<sup>1</sup> entraîné sur la base de 500 points sont en revanche excellentes sur l'ensemble de la ligne centrale avec des indicateurs de qualité moyen et maximal tout à fait satisfaisants. Nous atteignons globalement en effet une erreur moyenne très inférieure à 1% et une erreur maximale de 1.69%. Cette première approche sur cas homogène apporte une preuve de concept forte et encourageante pour le traitement de cas plus complexes.

## 4.3 Cas hétérogène : étude d'une flamme de diffusion simplifiée

---

### 4.3.1 Paramétrisation

---

Le second cas d'étude est également issu de l'article de COELHO, PEREZ et EL HAFI [2] mais est hétérogène en température et en concentration en espèces rayonnantes. Bien que le domaine d'étude soit toujours axisymétrique, les dimensions ne sont pas les mêmes que précédemment. Nous considérons ici un cylindre de 1.2 m de longueur et de 0.3 m de rayon (voir figure 4.4).

---

<sup>1</sup>Artificial Neural Network



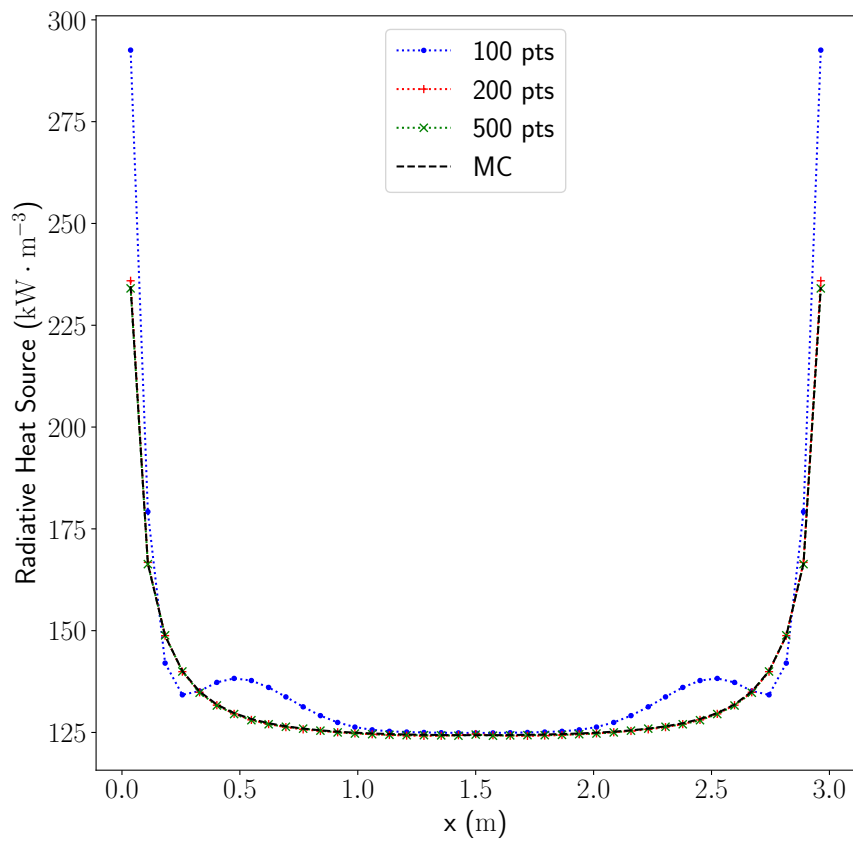


FIG. 4.3 : Prédiction de la divergence du flux radiatif sur l'axe de rotation du domaine par réseau de neurones - Cas homogène à  $T = 1200$  K

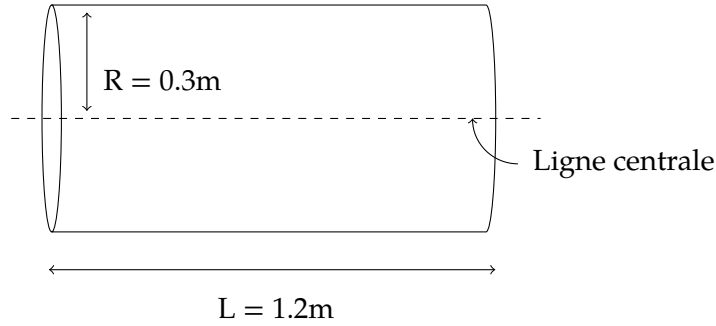


FIG. 4.4 : Géométrie étudiée - Cas hétérogène

Les parois sont à une température de 800 K à l'exception du mur droit ( $l = 1.2$ ) fixé à 300 K. La fraction volumique de suie reste constante et égale à  $10^{-7}\text{m}^3.\text{m}^{-3}$ . Les distributions paramétriques des autres grandeurs thermophysiques sont données par les équations suivantes et représentées graphiquement sur les figures 4.5a à 4.5c :

$$T(l, r) = 800 + 1200 \left(1 - \frac{r}{R}\right) \left(\frac{l}{L}\right) \quad (4.1)$$

$$x_{\text{H}_2\text{O}}(l, r) = 0.05 \left[1 - 2 \left(\frac{l}{L} - 0.5\right)^2\right] \left(2 - \frac{r}{R}\right) \quad (4.2)$$

$$x_{\text{CO}_2}(l, r) = 0.04 \left[1 - 3 \left(\frac{l}{L} - 0.5\right)^2\right] \left(2.5 - \frac{r}{R}\right) \quad (4.3)$$

Ce cas académique de benchmark a été choisi dans la mesure où il se rapproche de celui d'une flamme de diffusion simplifiée. Nous conserverons cette base pour l'ensemble des études à venir en y intégrant des perturbations ponctuelles (voir section 4.4.1), en bruitant artificiellement les données d'entraînement (voir section 4.4.2) ou en passant aux hautes pressions (voir section 4.3.5).

Dans le prochain paragraphe, le code de Monte Carlo en collision nulle est validé. Nous présentons ensuite les résultats obtenus après entraînement et interpolation par réseaux de neurones sur le même schéma que dans le cas homogène précédent. Une étude de sensibilité à la méthode de tirage des points de référence (voir section 3.2.3) a par ailleurs été menée. Enfin, une comparaison avec d'autres méthodes d'interpolations plus classiques est proposée pour démontrer l'intérêt et la puissance des réseaux de neurones pour les tâches de regression de ce type.

Dans chacun de ces cas, le réseau de neurones a cette fois 5 entrées : les 2 coordonnées spatiales  $l$  et  $r$ , la température et les fractions molaires d'espèces rayonnantes au point d'interpolation (ou d'entraînement). La pression et la fraction volumique étant constantes, ces grandeurs n'ont pas besoin d'être prises explicitement en compte en tant que données d'entrée du réseau.

### 4.3.2 Validation du code de Monte Carlo sur cas hétérogène

Dans cette configuration, la formulation en collisions nulles de notre algorithme de Monte Carlo est utilisée (voir section 1.4.2). Avant de lancer la génération d'une base d'entraînement, une validation de la méthode stochastique a été menée par comparaison avec les données fournies dans les travaux de COELHO, PEREZ et EL HAFI [2]. Au même titre que dans le cas homogène, nous pouvons constater sur

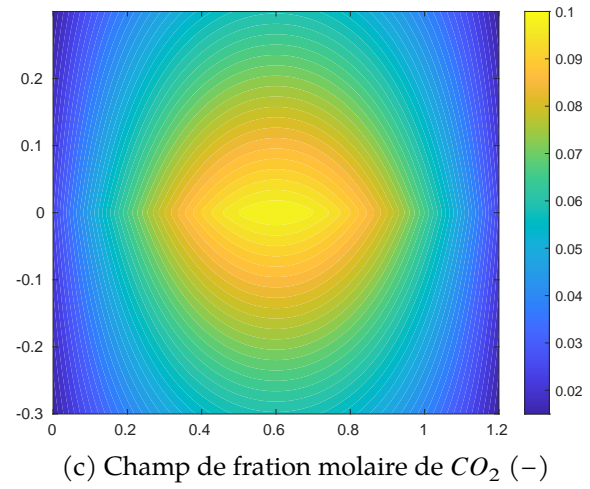
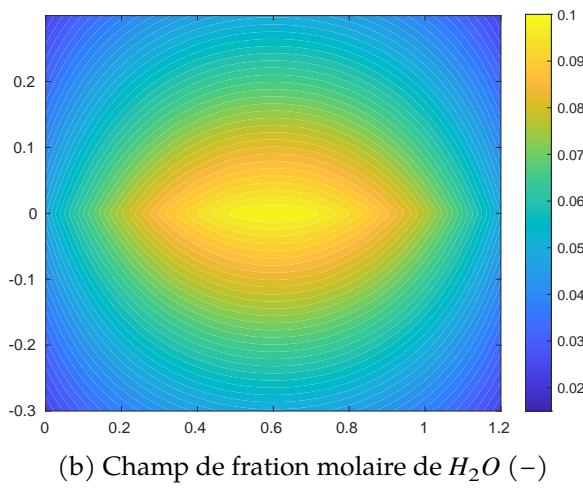
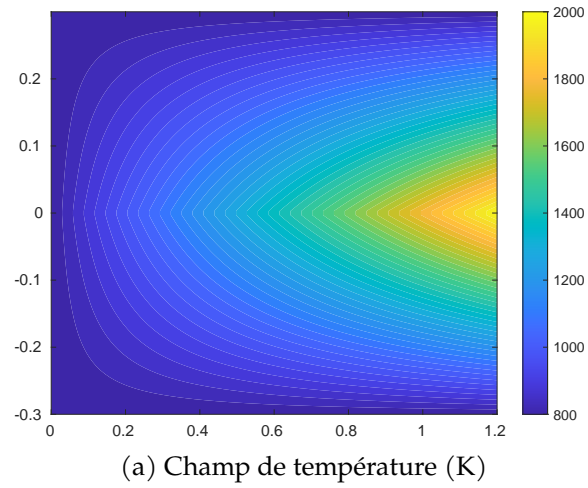


FIG. 4.5 : Représentation des champs thermophysiques du cas de benchmark hétérogène [2]

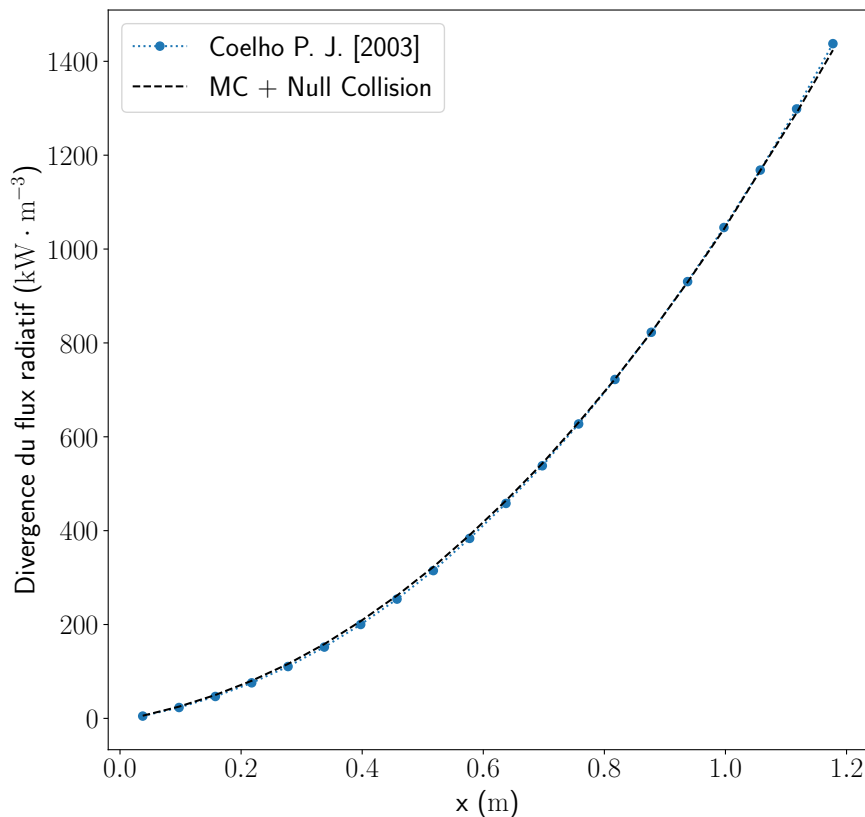


FIG. 4.6 : Validation du code Monte Carlo en collision nulle - Divergence du flux radiatif sur l'axe de révolution du domaine ( $r = 0$ ) en  $kW.m^{-3}$

la figure 4.6 que les données générées par notre algorithme se superposent aux données de référence fournies. Cette correspondance se retrouve sur l'ensemble du domaine avec le même niveau de précision. Le code utilisé dans sa formulation en collision nulle pour traiter des milieux hétérogènes est donc validé.

### 4.3.3 Influence de la méthode de tirage des points de référence

Nous présentons dans cette section les résultats obtenus par réseaux de neurones et comparons la qualité de l'interpolation en fonction de la méthode de tirage des points de référence dans le domaine. Dans la section 3.2.3, nous avons présenté deux méthodes dont l'une consiste à simplement tirer aléatoirement et uniformément les coordonnées spatiales des points d'entraînement dans le domaine tandis que l'autre respecte un échantillonnage selon une suite de Halton à 2 dimensions (en base 2 et 3).

Dans ce cas, les indicateurs d'erreur relative utilisés dans le cas précédent ne sont plus pertinents pour juger de la fiabilité et de la performance de l'outil utilisé. En effet, la valeur de la divergence de flux peut très fortement varier entre les zones les plus chaudes et les plus froides du domaine. Par conséquent, une erreur de quelques  $kW.m^{-3}$  sur la valeur calculée dans les zones froides et pauvres en espèces

rayonnantes peut mener à des erreurs relatives très importantes alors que l'impact sur la thermique et la physique globale du milieu sera très limité. Nous préférons donc présenter des erreurs moyennes et maximales absolues, plus représentatives ici. Une comparaison relative sera faite en commentaire mais par rapport à la valeur moyenne de la divergence du flux dans l'ensemble du domaine. Trois bases de données de 200, 500 et 1000 points d'entraînement ont été étudiées. Nous comparons les résultats obtenus par interpolation avec les réseaux de neurones à un champ de référence calculé par méthode de Monte Carlo dont les points n'appartiennent pas aux bases de données d'entraînement.

**Tirage aléatoire et uniforme** Les points de référence ont été tirés dans ce cas aléatoirement dans le milieu selon une loi uniforme. Les erreurs absolues obtenues avec les différentes bases sont regroupées dans le tableau 4.3. Tout d'abord, nous pouvons remarquer que le nombre de neurones nécessaires et le nombre de paramètres sont bien plus faibles que pour le cas homogène (voir tableau 4.2). Cette conclusion est assez inattendue dans la mesure où le cas d'étude est plus complexe. La différence sur le nombre de variables d'entrée et l'ajout des informations physiques du champ semblent ici guider le réseau de neurones dans ses interpolations et diminuer de fait la difficulté de la modélisation.

Points d'entraînement	Neurones	Paramètres	$\gamma$	Temps d'entraînement	Erreur absolue moyenne	Erreur absolue maximale
200	3	22	18	8.89 s	1.08 kW.m <sup>-3</sup>	34.42 kW.m <sup>-3</sup>
500	5	36	30	51.97 s	0.697 kW.m <sup>-3</sup>	12.88 kW.m <sup>-3</sup>
1000	8	57	50	133.01 s	0.592 kW.m <sup>-3</sup>	12.87 kW.m <sup>-3</sup>

TABLE. 4.3 : Qualité des interpolations sur l'ensemble du domaine de la divergence de flux par réseaux de neurones en milieu hétérogène - Tirage uniforme

Sur la figure 4.7, nous traçons le champ de divergence de flux produit par le réseau de neurones entraîné sur 500 jeux de données. Nous faisons également apparaître la position des points d'entraînement dans le milieu. Qualitativement, la zone où la divergence du flux est la plus forte correspond assez logiquement à la zone où les températures sont les plus élevées. Ce facteur est prédominant au vu des niveaux assez élevés de température et des niveaux de fraction molaire des espèces plutôt bas. La valeur de divergence de flux maximale atteinte est de 1473 kW.m<sup>-3</sup> tandis que la valeur minimale est de -15 kW.m<sup>-3</sup>. Cette valeur négative traduit le fait que les gaz absorbent plus qu'ils n'émettent. Enfin, la figure 4.7 permet également de visualiser le travail complet de reconstruction du champ de divergence de flux effectué par le réseau de neurones à partir de quelques points sondes dans l'ensemble du milieu. Cette première étude laisse ainsi présager du potentiel de ce type d'outil pour la génération de champs continus à partir de données ponctuelles.

La valeur moyenne de la divergence de flux sur l'ensemble de ce champ est égale à 155.34 kW.m<sup>-3</sup>. D'après le tableau 4.3, les erreurs moyennes absolues représentent donc chacune moins de 1% de cette valeur quelle que soit la base de données utilisée. En revanche, l'erreur maximale est encore relativement élevée pour la base de données de 200 points. L'augmentation du nombre de points à 500 et 1000 permet de réduire cet écart d'un facteur 3 environ. Nous constatons également que l'utilisation de 1000 points n'est pas nécessaire car le gain est minime sur la qualité de l'interpolation. Notons enfin que les erreurs absolues maximales apparaissent dans les zones chaudes où la valeur de la divergence de flux est très élevée (voir figure 4.8). En conséquence, si la comparaison à la valeur moyenne de cette grandeur est sévère, l'erreur relative en ces points particuliers est très faible. Typiquement, l'erreur absolue de 12.88 kW.m<sup>-3</sup> est obtenue aux points de coordonnées ( $l = 1.19m, r = -0.030$ ). La valeur de la divergence du

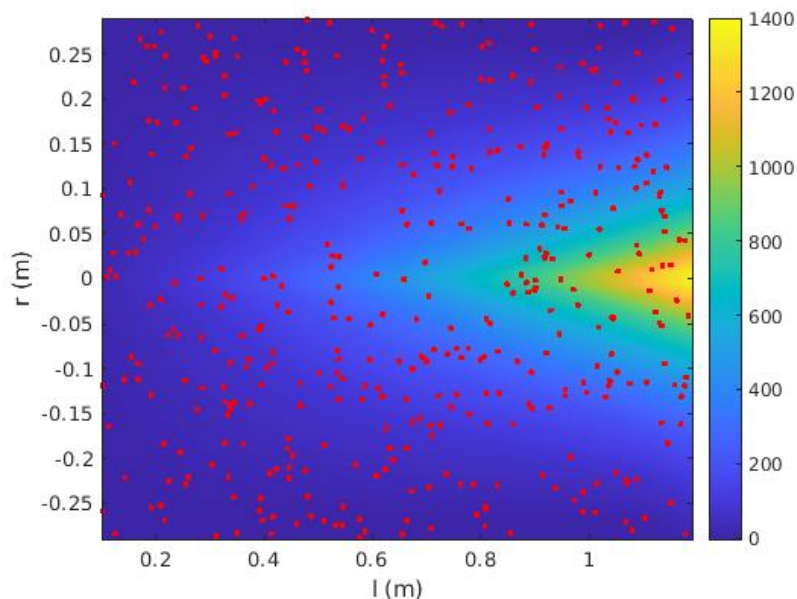


FIG. 4.7 : Construction du champ de divergence de flux par réseaux de neurones - Cas hétérogène, tirage aléatoire de 500 points de références

flux calculée par Monte Carlo en ce point est de  $1101 \text{ kW.m}^{-3}$  soit un écart relatif de 1.17 % seulement dans les zones les plus chaudes de la flamme.

Pour finir, il est intéressant de remarquer sur la figure 4.8 que les zones où les erreurs absolues sont élevées par rapport au reste du champ concordent avec des zones pauvres en points de référence (voir cercles bleu), *a fortiori* lorsque la divergence de flux est élevée. Cette remarque relève de la tendance et confirme l'importance de la méthode de tirage des points de référence.

**Tirage selon une suite de Halton** Dans ce paragraphe, nous souhaitons comparer les résultats obtenus avec une méthode de tirage selon une suite de Halton à discrédance faible (voir section 3.2.3). Nous utilisons une suite à deux dimensions en base 2 et 3 pour déterminer les coordonnées spatiales des points de référence dans le milieu. Théoriquement, une suite de halton de dimension  $N$  est construite à l'intérieur d'un hypercube de côté égal à 1 et de même dimension. Dans ce cas, les coordonnées selon  $l$  et  $r$  ont été tirées entre 0 et 1 puis normalisées entre les bornes du domaine d'étude.

Points d'entraînement	Neurones	Paramètres	$\gamma$	Temps d'entraînement	Erreur absolue moyenne	Erreur absolue maximale
200	3	22	17	32.06 s	$0.784 \text{ kW.m}^{-3}$	$26.86 \text{ kW.m}^{-3}$
500	6	43	37	79.00 s	$0.577 \text{ kW.m}^{-3}$	$9.433 \text{ kW.m}^{-3}$
1000	8	57	51	283.22 s	$0.544 \text{ kW.m}^{-3}$	$9.432 \text{ kW.m}^{-3}$

TAB. 4.4 : Qualité des interpolations par réseaux de neurones en milieu hétérogène - Tirage de Halton

Si l'on compare les résultats des tableaux 4.3 et 4.4, nous pouvons constater que les tendances sont les mêmes avec le tirage aléatoire et celui de Halton. Néanmoins, la qualité de l'interpolation augmente sensiblement grâce à la suite à discrédance faible, notamment au niveau de l'erreur maximale absolue. Le

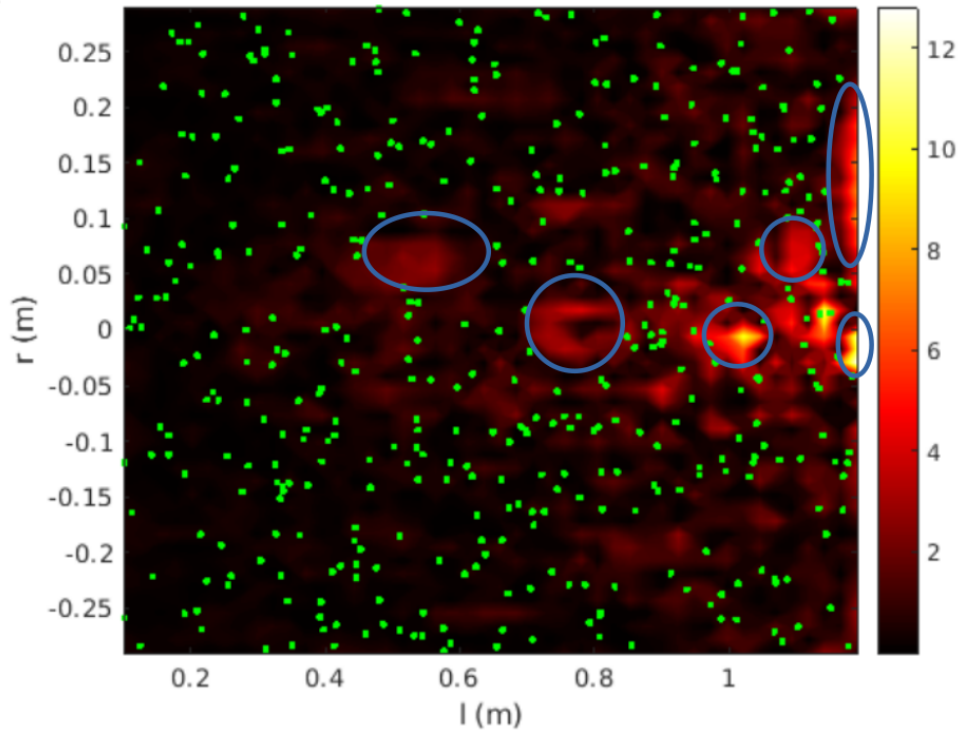


FIG. 4.8 : Soustraction du champ interpolé à celui de référence et visualisation des 500 points d'entraînement aléatoires (échelle en  $kW.m^{-3}$ )

champ soustrait présenté dans la figure 4.9 représente la distribution des erreurs absolues en  $kW.m^{-3}$  dans le cas d'un tirage des points de référence selon une suite de Halton. Les cercles bleu sont les mêmes que ceux identifiés dans la figure 4.8. Nous pouvons remarquer que la distribution plus uniforme des points de référence permet en partie de réduire les zones identifiées avec les plus fortes erreurs précédemment. Cet effet se constate notamment sur le cercle le plus à gauche ainsi que sur les quatre zones les plus proches de la paroi de droite. Des pics sont encore présents dans la région la plus chaude de la flamme mais la qualité globale de l'interpolation est améliorée. Cette conclusion démontre ainsi l'importance de la méthode de tirage des points d'entraînement. Cela dit, gardons malgré tout en tête que dans les deux cas, les résultats obtenus restent très satisfaisants et démontrent le potentiel de cette méthode sur des champs très hétérogènes en température et en espèces.

**Tracé de la divergence de flux sur la ligne centrale du cylindre** Enfin et pour clore cette section, nous proposons de tracer sur la figure 4.10 les résultats obtenus par interpolation sur la ligne centrale du cylindre et de les comparer aux données du cas de Benchmark de COELHO, PEREZ et EL HAFI [2]. Les bases de données d'entraînement pour la production de ces résultats sont celles contenant 500 points de référence. Les trois courbes se superposent quasiment toutes les trois, ce qui vient confirmer les assertions précédentes. Nous pouvons également constater la légère amélioration de la qualité d'interpolation dans la zone la plus chaude avec le tirage de Halton. En conclusion de cette étude comparative, nous pouvons donc donner une préférence pour cette méthode de sélection des points d'entraînement. De plus, il apparaît assez clairement qu'une base de 500 jeux de données est suffisante pour une prédiction précise du champ de divergence de flux dans son ensemble.

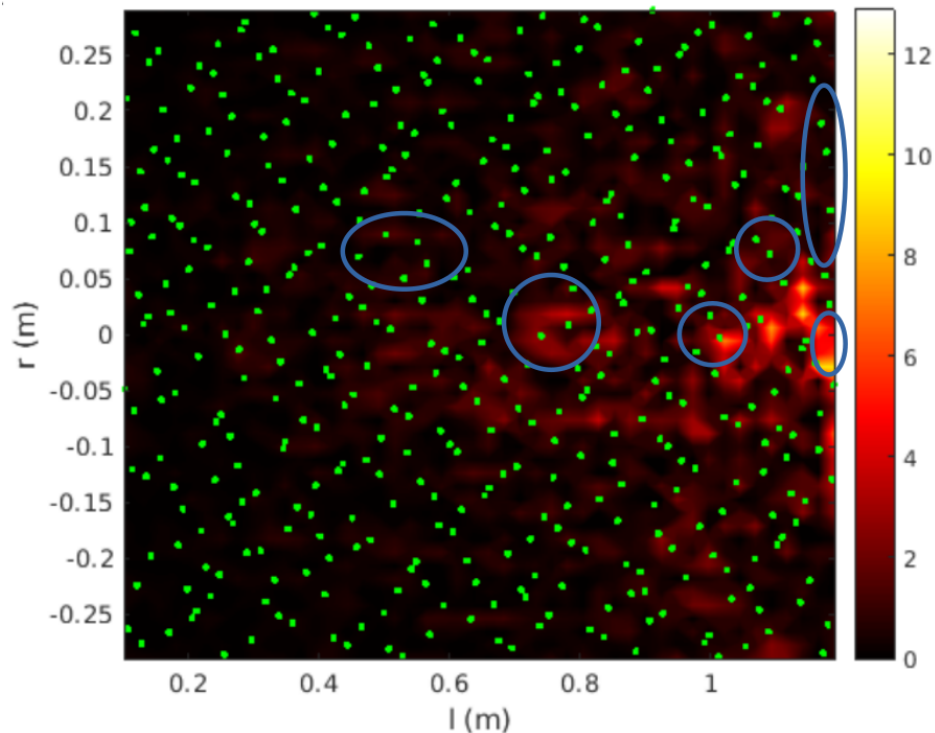


FIG. 4.9 : Soustraction du champ interpolé à celui de référence et visualisation des 500 points d'entraînement selon une suite Halton (Echelle en  $kW.m^{-3}$ )

#### 4.3.4 Comparaison avec d'autres méthodes d'interpolation

Comme nous avons déjà pu l'évoquer à plusieurs reprises, nous utilisons ici les réseaux de neurones comme des interpolateurs universels complexes. Sur la base de ce constat, nous pouvons alors nous demander légitimement l'intérêt que peut représenter l'utilisation de ce type d'outil en comparaison d'autres méthodes d'interpolation plus simples et bien connues.

Ce paragraphe se propose de répondre à cette question. Néanmoins, nous pouvons d'ores et déjà apporter quelques éléments de réponses conceptuels. Il existe en effet une différence fondamentale entre les réseaux de neurones et les autres méthodes d'interpolations vis à vis de la nature des données d'entrées pouvant être prises en compte. Les ANN se distinguent en effet par leur capacité à pouvoir travailler avec des données d'entrées hétérogènes. L'interpolation se fait alors sur l'ensemble de l'espace des phases de la fonction à modéliser et pas seulement sur les coordonnées spatiales du point inconnu comme c'est le cas avec des méthodes d'interpolation linéaire ou de krigeage par exemple [102, 103]. De ce point de vue et bien qu'un réseau de neurones reste un outil purement mathématique, il est en un sens plus proche de la physique réelle du problème à modéliser.

Nous comparons dans ce paragraphe les réseaux de neurones à 4 autres méthodes plus classiques : l'interpolation par splines cubique (SC), par voisins naturels (VN) et l'interpolation bi-linéaire (BL) [104, 105]. Plusieurs comparaisons entre ces différentes méthodes et les réseaux de neurones existent déjà dans la littérature, en particulier dans les domaines de la cartographie, de la géologie et de l'imagerie [106, 107, 102]. Néanmoins, en fonction du cas d'application, la supériorité des réseaux de neurones dans ces études est questionable en comparaison des autres méthodes d'interpolation plus directes.



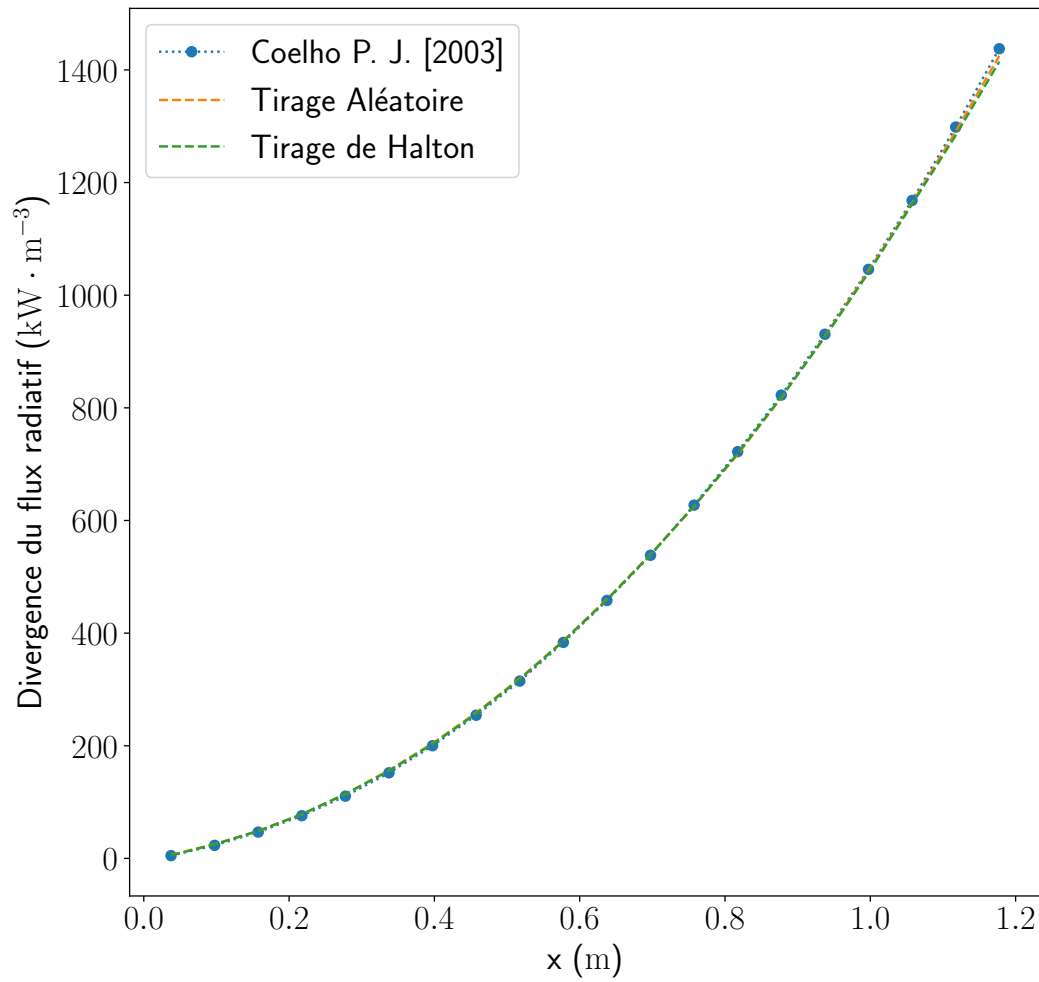


FIG. 4.10 : Comparaison entre les données de benchmark et les interpolations par réseaux de neurones sur la ligne centrale ( $r = 0$ ) - Base de données de 500 points

Cette limite provient du fait que les données d'entrées sont les mêmes quelle que soit la méthode, à savoir uniquement des coordonnées spatiales. Ici, les réseaux de neurones sont enrichis de données thermophysiques permettant de guider les interpolations faites par l'outil en fonction de la physique du problème et plus seulement de la position du point interpolé.

Nous montrons sur la figure 4.11 un comparatif entre les résultats obtenus grâce aux différentes méthodes d'interpolation. Nous utilisons dans les 4 situations la base de données de 500 points tirés selon une suite de Halton. Nous précisons que le logiciel MATLAB version R2020a a été utilisé pour l'interpolation bi-linéaire, par Spline Cubique et par proche voisin. Nous montrons alors clairement que dans cette situation, un réseau de neurones adapté, nourri par la physique du problème surpasse de très loin les méthodes plus classiques et a donc toute sa pertinence dans cette étude. Le tableau 4.5 vient également compléter ce constat en comparant la valeur de la RMSE sur l'ensemble d'un champ interpolé par ces différentes méthodes. Si les erreurs produites par les méthodes classiques sont proches, il existe en revanche un ordre de grandeur de différence avec les résultats obtenus par réseau de neurones.

Méthode	RMSE ( $kW.m^{-3}$ )
ANN	1.0453
Splines cubiques	15.79
Proche voisin	15.21
Bi-linéaire	14.06

TABLE. 4.5 : Comparaison de la valeur de la RMSE calculée pour les différentes méthodes d'interpolation

Pour finir, nous pouvons constater de manière plus qualitative les différences entre les champs interpolés et soustraits pour chaque méthode dans l'annexe B. Nous retrouvons les comportements constatés sur la figure 4.11, à savoir des oscillations sur la divergence du flux pour les 3 méthodes d'interpolation. La comparaison à l'échelle entre les champs soustraits corrobore également les conclusions tirées d'après le tableau 4.5 et les valeurs de RMSE. Nous constatons en effet que les erreurs produites par le réseau de neurones sont très faibles, et ce sur l'ensemble du champ, par rapport à celle produite par les méthodes standards d'interpolation.

### 4.3.5 Etude du cas de référence à haute pression

Dans cette section, nous reprenons le cas de Benchmark hétérogène présenté précédemment mais dans une configuration à haute pression et en intégrant une nouvelle espèce participante. La pression est conservée constante mais égale à 22 bars afin de nous rapprocher des conditions d'opération classiques d'une chambre de combustion de turbomachine. Dans le même état d'esprit, nous ajoutons du  $CO$  dans le domaine, distribué spatialement selon l'équation suivante :

$$x_{CO}(l, r) = 0.004 \left[ 1 - 3 \left( \frac{l}{L} - 0.5 \right)^2 \right] \left( 2.5 - \frac{r}{R} \right) \quad (4.4)$$

Cette équation est identique à celle régissant la distribution du  $CO_2$  (voir équation (4.3)) au facteur multiplicatif près. Il existe donc un facteur 10 entre la distribution de ces deux espèces, ce qui est une approximation correcte des proportions constatées dans des flammes de diffusion standards. Le modèle de gaz utilisé est le modèle HPBM à 367 bandes (voir section 1.4.4) développé par PIERROT [1] pour les applications Safran.

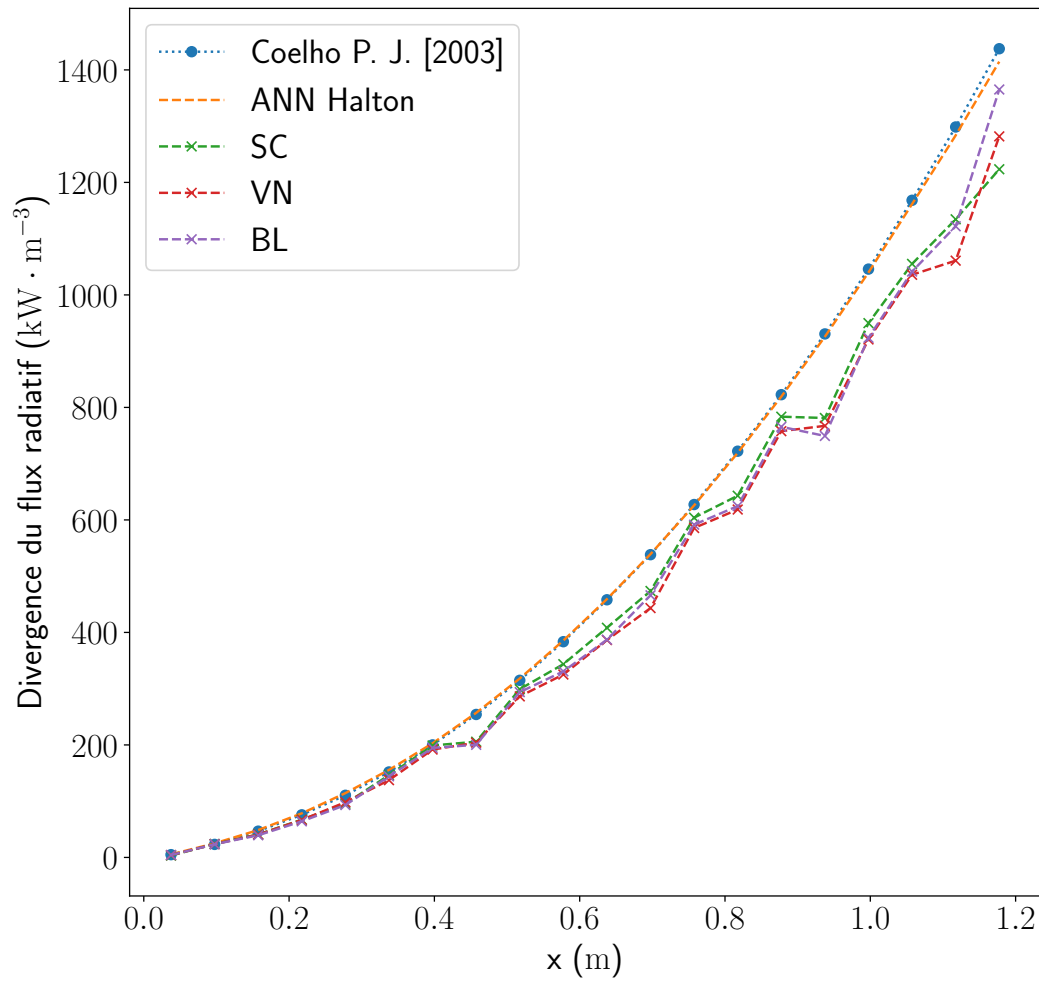


FIG. 4.11 : Comparaison entre les différentes méthodes d'interpolation - Résultats sur la ligne centrale du cylindre ( $r = 0$ )

i

Le code de Monte Carlo utilisé pour la génération des points d'entraînement et de comparaison avec l'ANN est le même que précédemment. Néanmoins, la méthode de modélisation du gaz a été modifiée. L'implémentation du modèle HPBM n'a ainsi pas été rigoureusement validée faute de temps et de résultats de référence consolidés sur ce cas et à cette pression.

Nous traçons sur la figure 4.12 la divergence du flux radiatif sur la ligne centrale du volume ( $r = 0$ ) dans les conditions exposées ci-dessus. Les valeurs numériques obtenues sont référencées dans le tableau C.1 de l'annexe C pour une éventuelle validation à venir ou pour reproduction des résultats.

Encore une fois, le réseau de neurone utilisé ici a été entraîné sur une base de 500 points tirés selon une suite de Halton. Le nombre de neurones nécessaires est ici faible (voir tableau 4.6) par rapport au cas à pression atmosphérique pour un même nombre de points d'entraînement. Il est difficile d'interpréter qualitativement ce résultat. Il semble que les liens entre les différents points de référence soient plus simples que dans un cas à pression atmosphérique. Néanmoins, un parallèle peut éventuellement être fait entre l'allure des spectres d'absorption des gaz à ces différentes pressions et donc à l'élargissement des raies d'absorption. Le modèle HPBM lui-même est de fait plus simple à manipuler qu'un modèle spectral de type SNB-CK. Peut-être pouvons-nous simplement voir ici une prise en compte implicite de la difficulté variable à modéliser le rayonnement des gaz en fonction de la pression par le réseau de neurones.

Malgré l'ajout d'un nouveau gaz rayonnant, la simplicité d'implémentation du modèle HPBM et son efficacité permettent également de réduire drastiquement le temps de calcul de la base de données, à tout moyen informatique égal par ailleurs. Avec notre configuration <sup>2</sup>, un point de Monte Carlo à 500 000 réalisations est calculé en 5.6 s en moyenne contre 31.7 s avec le modèle SNB-CK à pression atmosphérique. Notons cependant que les valeurs des coefficients d'absorption majorants peuvent varier à la baisse d'une pression à l'autre et ainsi influencer l'efficacité de l'algorithme à collisions nulles.

Paramètre	Valeur
Base de données	500 points
Méthode de tirage	Halton
Temps moyen de calcul d'un point de référence	5.6 s
Nombre de neurones	4
Nombre de paramètres	33
Valeur de $\gamma$	28
Nombre d'époques	10000
Temps d'entraînement	44.65 s
RMSE	$3.725 \text{ kW} \cdot \text{m}^{-3}$

Tab. 4.6 : Paramètres et caractéristiques de l'entraînement par ANN - Cas à haute pression

## 4.4 Prise en compte de perturbations

Sur la base des bons résultats obtenus sur le cas hétérogène de benchmark, nous avons ensuite mis à l'épreuve notre méthode sur deux cas plus complexes. L'idée est de tester la robustesse de l'outil

<sup>2</sup>Ordinateur portable équipé d'un processeur Intel Core i5-8265U et de 15.5 GB de RAM

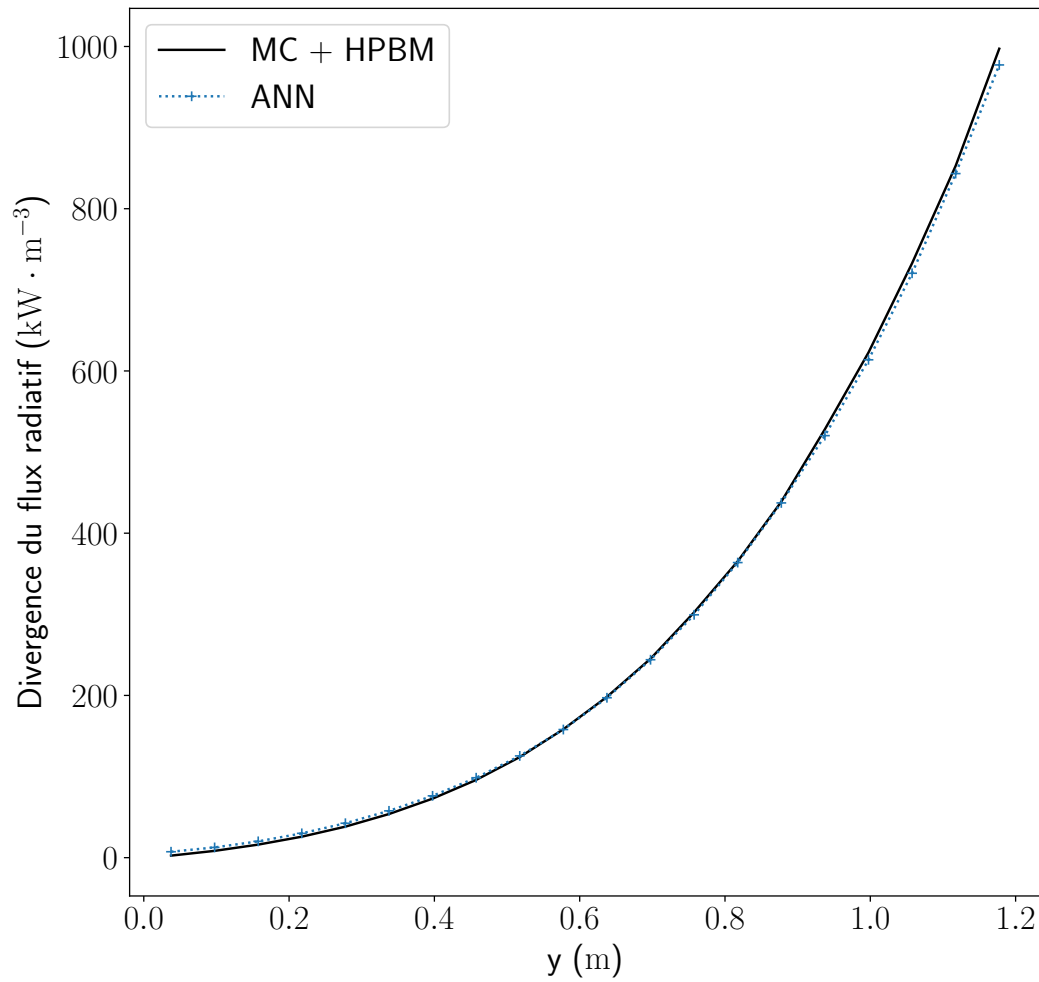


FIG. 4.12 : Divergence du flux radiatif sur la ligne centrale ( $r = 0$ ) à haute pression ( $P = 22 \text{ bars}$ ) et en présence de  $CO$

développé face à des problématiques pouvant être rencontrées dans des contextes industriels.

#### 4.4.1 Point chaud

Le premier cas d'étude concerne la prise en compte d'un point chaud dans le domaine. Ce type de phénomène intense et très ponctuel se traduit par un pic de température et d'espèces rayonnantes et donc par des gradients très importants aux alentours du centre du point chaud. Il n'est pas rare de constater cette singularité dans les applications de combustion, dans le cas par exemple d'auto-inflammation non désirée du mélange air-essence (phénomène de cliquetis par exemple).

Les champs thermophysiques en arrière plan du point chaud ne changent pas et sont toujours définis par les équations (4.1) à (4.3). Nous ajoutons par superposition le point chaud de manière paramétrique grâce aux équations suivantes (voir figure 4.13) :

$$T_{PC} = 1200 \exp \left\{ \frac{-\sqrt{(l_{PC} - l)^2 + (r_{PC} - r)^2}}{0.04} \right\} \quad (4.5)$$

$$x_{H_2O,PC} = 0.05 \exp \left\{ \frac{-\sqrt{(l_{PC} - l)^2 + (r_{PC} - r)^2}}{0.04} \right\} \quad (4.6)$$

$$x_{CO_2,PC} = 0.1 \exp \left\{ \frac{-\sqrt{(l_{PC} - l)^2 + (r_{PC} - r)^2}}{0.04} \right\} \quad (4.7)$$

Avec  $l_{PC} = 0.1$  et  $r_{PC} = 0.15$ , les coordonnées du centre du point chaud. En accord avec les conclusions tirées dans la section précédente, nous avons utilisé ici une base de données contenant 500 points de référence. Le réseau a été entraîné sur 10 000 époques. Ses caractéristiques sont résumées dans le tableau 4.7. Comme cela a déjà été évoqué à plusieurs reprises dans ce manuscrit (voir notamment la figure 3.2), un pré-traitement a été absolument nécessaire dans ce cas afin d'identifier les extrema potentiels pour la normalisation des données d'entraînement. Ainsi, le point de coordonnées (0.1, 0.15) et les données thermophysiques qui s'y rapportent ont été intégrées à la base d'entraînement.

Sans cet ajout, la borne supérieure pour la normalisation des données aurait potentiellement été très éloignée de la divergence de flux maximale pouvant être rencontrée dans le milieu. Dans cette situation, si nous souhaitons alors estimer par le réseau une valeur de divergence supérieure à cette borne, la sortie attendue sur le neurone de la dernière couche devrait théoriquement être supérieure à 1 (nous rappelons que les données sont normalisées entre -1 et 1 dans ce cas). Or, la fonction d'activation utilisée pour le réseau est une tangente hyperbolique qui tend asymptotiquement vers 1 (voir section 2.2). Autrement dit, le réseau de neurones ne sera pas capable (ou alors très difficilement par une valeur extrêmement marquée de son biais) de dépasser cette valeur maximale de sortie. On dit alors qu'il est saturé. Physiquement, cette limite aura pour conséquence de ne pas pouvoir estimer une divergence de flux très supérieure à la borne de normalisation utilisée. Ce phénomène est mis en exergue sur la figure 4.14 où l'on compare la réponse du réseau avec et sans traitement préalable du champ. Nous pouvons en effet constater que sur le point chaud, la réponse du réseau ne peut pas aller au delà de  $1\,617\text{ kW}\cdot\text{m}^{-3}$  environ qui est une valeur proche de la borne supérieure de normalisation dans ce cas fixée à  $1\,470\text{ kW}\cdot\text{m}^{-3}$ .

Nous pouvons remarquer que le nombre de neurones nécessaires est plus faible que pour le cas hétérogène standard. Cette particularité semble curieuse de prime abord, dans la mesure où l'apparition

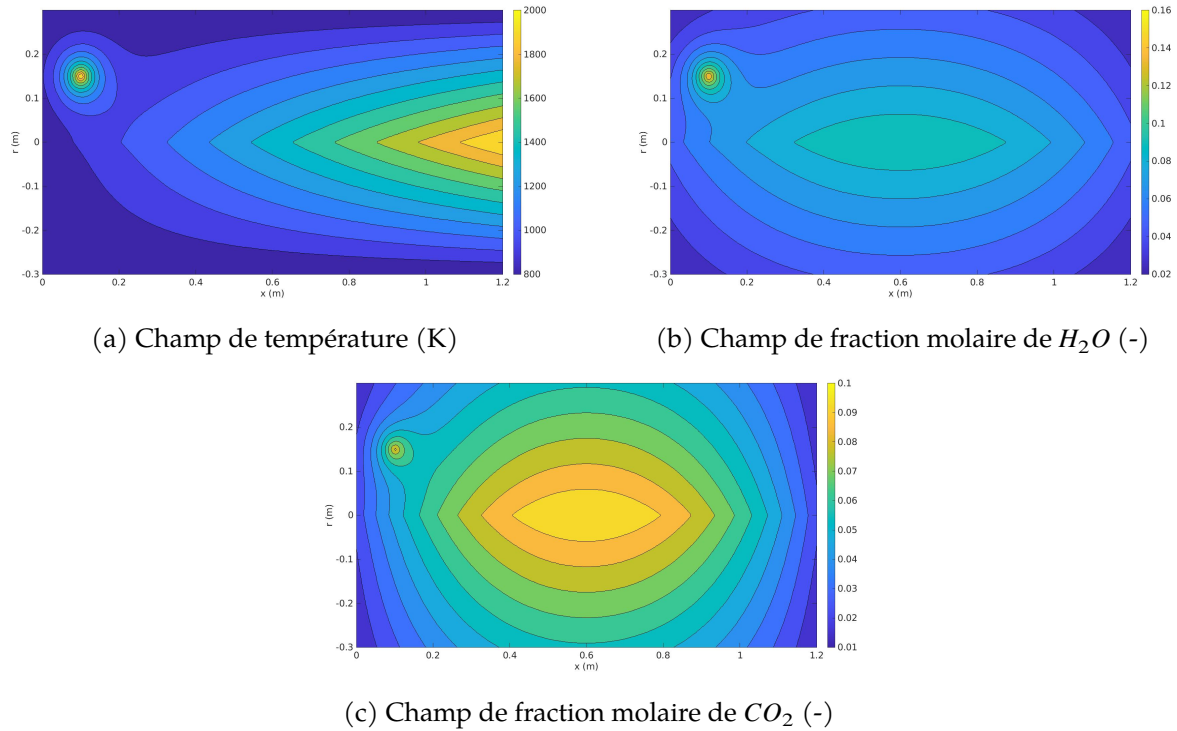


FIG. 4.13 : Champs thermophysiques - Cas « Point Chaud »

Paramètre	Valeur
Base de données	500 points
Méthode de tirage	Halton
Temps moyen de calcul d'un point de référence	21.3 s
Nombre de neurones	4
Nombre de paramètres	29
Valeur de $\gamma$	25
Nombre d'époques	10000
Temps d'entraînement	49.03 s
RMSE	$5.740 \text{ kW.m}^{-3}$

TAB. 4.7 : Caractéristiques du réseau de neurones et de l'entraînement - Cas « Point chaud »

d'un point chaud complexifie le problème à modéliser. Aussi, l'ajout de ce type de perturbation devrait intuitivement plutôt nécessiter plus de neurones pour sa bonne prise en compte. On peut se rendre compte par cet exemple que l'estimation *a priori* du nombre de neurones nécessaires pour la modélisation d'une fonction est très difficile. Un algorithme capable d'ajuster ce paramètre en fonction du problème est donc d'autant plus pertinent. Une explication possible pour justifier ce faible nombre de paramètres pourrait se trouver dans l'écart relatif qui existe entre les valeurs de divergence de flux très élevées à proximité du point chaud et celles dans le reste du domaine. Dans le cas hétérogène de benchmark, il est possible que le réseau de neurones aille chercher à estimer les subtilités entre valeurs proches dans les zones chaudes, ce qui implique un besoin de paramètres libres important. Nous pouvons imaginer ici qu'au regard du pic au point chaud et après normalisation, le lien fait entre les valeurs de divergence de

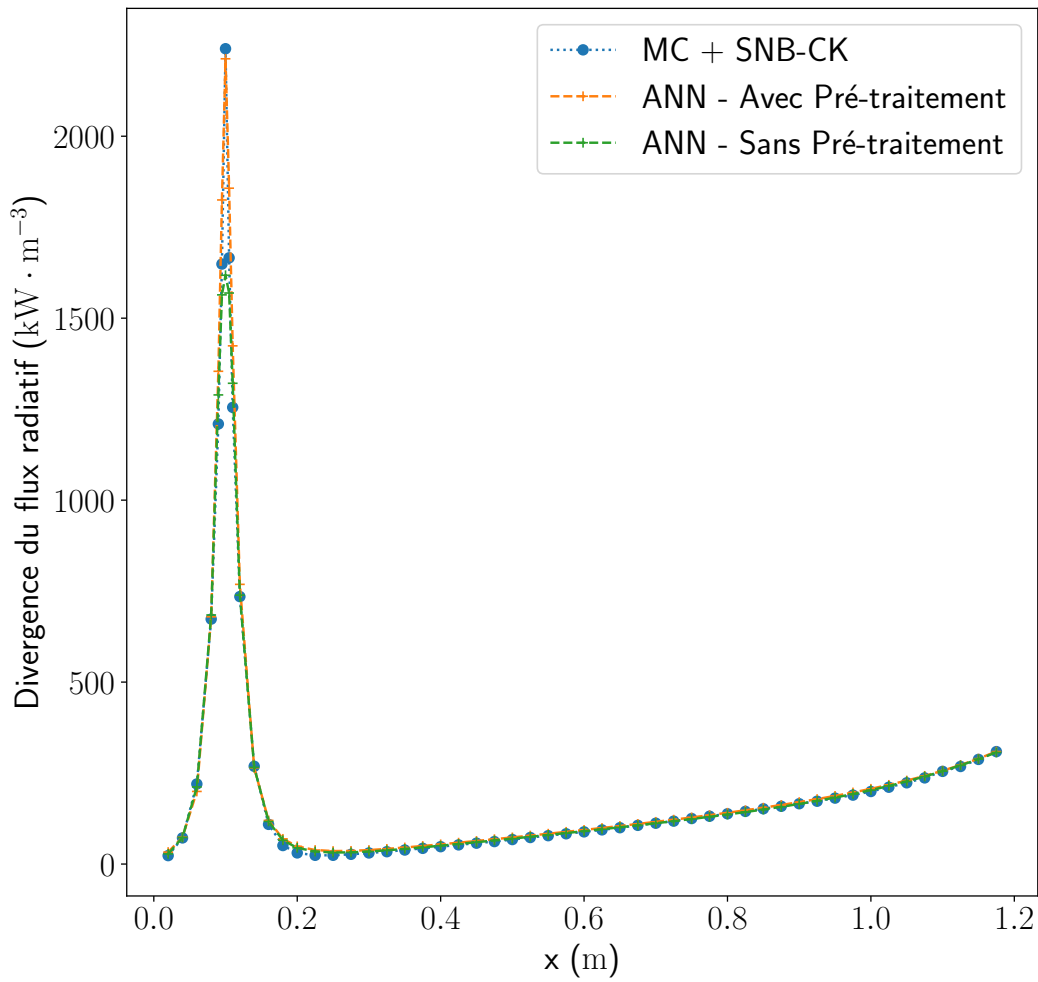


FIG. 4.14 : Saturation du réseau en l'absence de pré-traitement du champ thermophysique - Tracé de la divergence de flux ( $kW.m^{-3}$ ) sur la ligne axiale ( $r = 0.15$ )



flux dans la zone primaire de la flamme soit plus simple et plus lisse.

Quoi qu'il en soit, nous pouvons constater d'après les figures 4.15a à 4.15c que l'utilisation d'aussi peu de neurones est tout à fait justifiée dans la mesure où les résultats sur les lignes  $r = 0$  (ligne centrale),  $r = 0.15$  (ligne axiale passant par le point chaud) et  $l = 0.1$  (ligne radiale passant par le point chaud) sont encore une fois très proches de la référence calculée par Monte Carlo. Notons malgré tout que la valeur de RMSE est plus élevée que dans les autres cas d'étude. Cependant, les valeurs extrêmes de divergence de flux le sont également, il est donc compréhensible que l'erreur moyenne absolue soit plus forte. Les tableaux C.2 à C.4 de l'annexe C regroupent les valeurs numériques ayant permis le tracé de ces courbes. L'outil est donc tout à fait capable de prendre en considération ce type de perturbation ponctuelle et très localisée si les recommandations pour la construction de la base de données d'entraînement sont respectées. C'est un résultat encourageant vis à vis des capacités de l'outil à travailler dans des milieux très hétérogènes issus de simulation de systèmes réels.

## 4.4.2 Gestion du bruit numérique

Pour terminer ce chapitre, nous proposons une étude mettant à l'épreuve la robustesse de l'outil en présence de bruit numérique. L'idée est de tester si le réseau de neurones est capable d'absorber des erreurs potentielles sur les points d'entraînement et de proposer malgré tout une estimation correcte de la grandeur de référence. Ce cas présente deux intérêts. Tout d'abord, un réseau robuste et peu sensible au bruit sera capable de travailler correctement sur des jeux de données pouvant contenir des erreurs comme cela peut être le cas à la fois lorsqu'on utilise une base numérique ou issue d'essais expérimentaux. Ensuite, si le réseau est capable de s'accomoder de données bruitées, alors il sera possible de diminuer fortement le nombre de réalisations de Monte Carlo afin d'optimiser le temps de calcul des points de référence.

Encore une fois, les distributions de température et de fraction molaire des espèces ne diffèrent pas du cas de Benchmark (se référer aux équations (4.1) à (4.3)). Nous ajoutons cependant en chaque point un bruit numérique artificiel tel que :

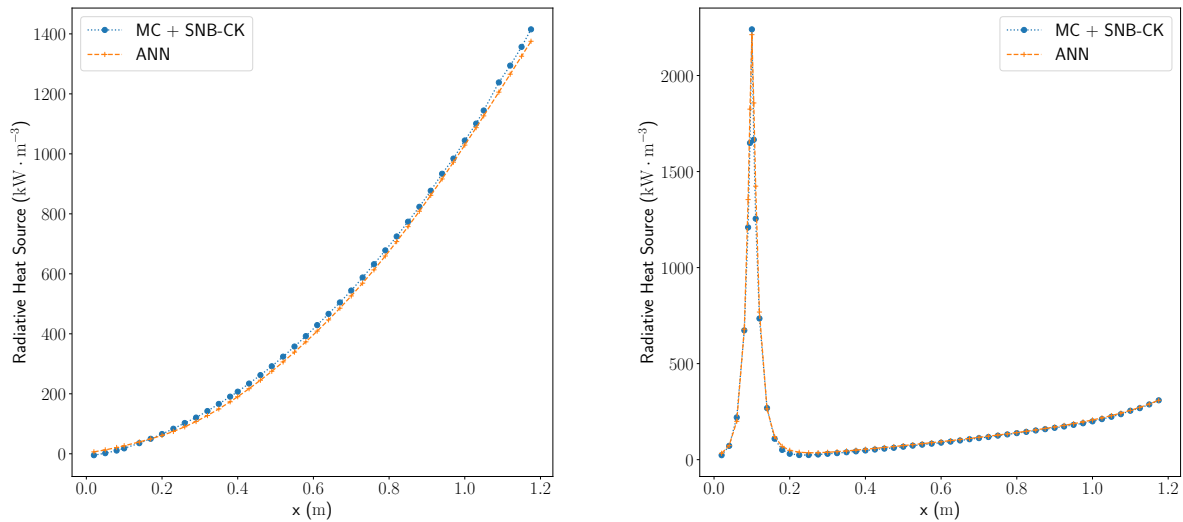
$$T_{noisy} = T + \mathcal{N}(0, 20) \quad (4.8)$$

$$x_{H_2O, noisy} = x_{H_2O} + \mathcal{N}(0, 0.005) \quad (4.9)$$

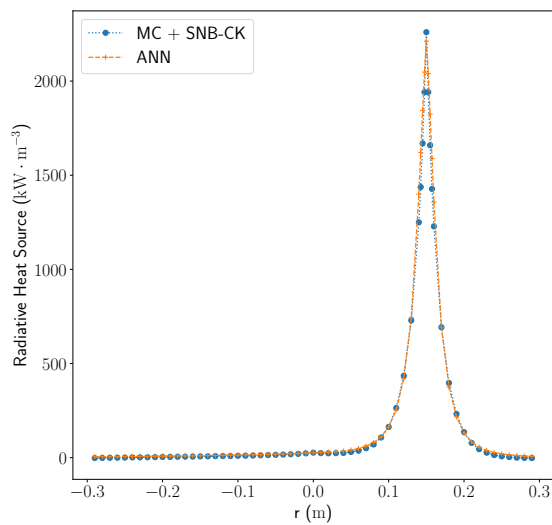
$$x_{CO_2, noisy} = x_{CO_2} + \mathcal{N}(0, 0.005) \quad (4.10)$$

Nous notons  $\mathcal{N}(\mu_n, \sigma)$ , une loi normale centrée en  $\mu_n$  et d'écart type  $\sigma$ . Ainsi, nous ajoutons par exemple au champ de température paramétrique et en chaque point, un bruit aléatoire selon une gaussienne centrée en 0 et d'écart type égal à 20 K. Les champs générés sont présentés sur la figure 4.16.

Dans la mesure où le bruit a été intégré aux paramètres thermophysiques du champ, la divergence de flux correspondante est également sujette à une erreur mais plus difficilement estimable. Le réseau de neurones a encore une fois été entraîné sur une base de 500 points bruités et tirés par méthode de Halton (voir tableau 4.8). Nous cherchons à approcher ces résultats grâce au réseau afin de conclure sur sa capacité à gérer le bruit numérique. Notons que la valeur de la RMSE présentée dans le tableau 4.8 est issue de la comparaison entre le champ estimé par le réseau et le champ théorique non-bruité calculé par Monte Carlo. On constate que, bien que la RMSE soit moins bonne qu'avec un entraînement sur une base de données précise (voir tableau 4.5), elle reste toujours faible devant les valeurs significatives de divergence de flux calculées dans le milieu. Sur la figure 4.17, nous présentons les résultats de



(a) Divergence du flux radiatif sur la ligne centrale  $r = 0$  ( $kW.m^{-3}$ ) (b) Divergence du flux radiatif sur la ligne longitudinale  $r = 0.15$  ( $kW.m^{-3}$ )



(c) Divergence du flux radiatif sur la ligne radiale  $x = 0.10$  ( $kW.m^{-3}$ )

FIG. 4.15 : Comparaison entre les interpolations du réseau de neurones et les valeurs de référence sur les lignes de contrôle - Cas « Point chaud »

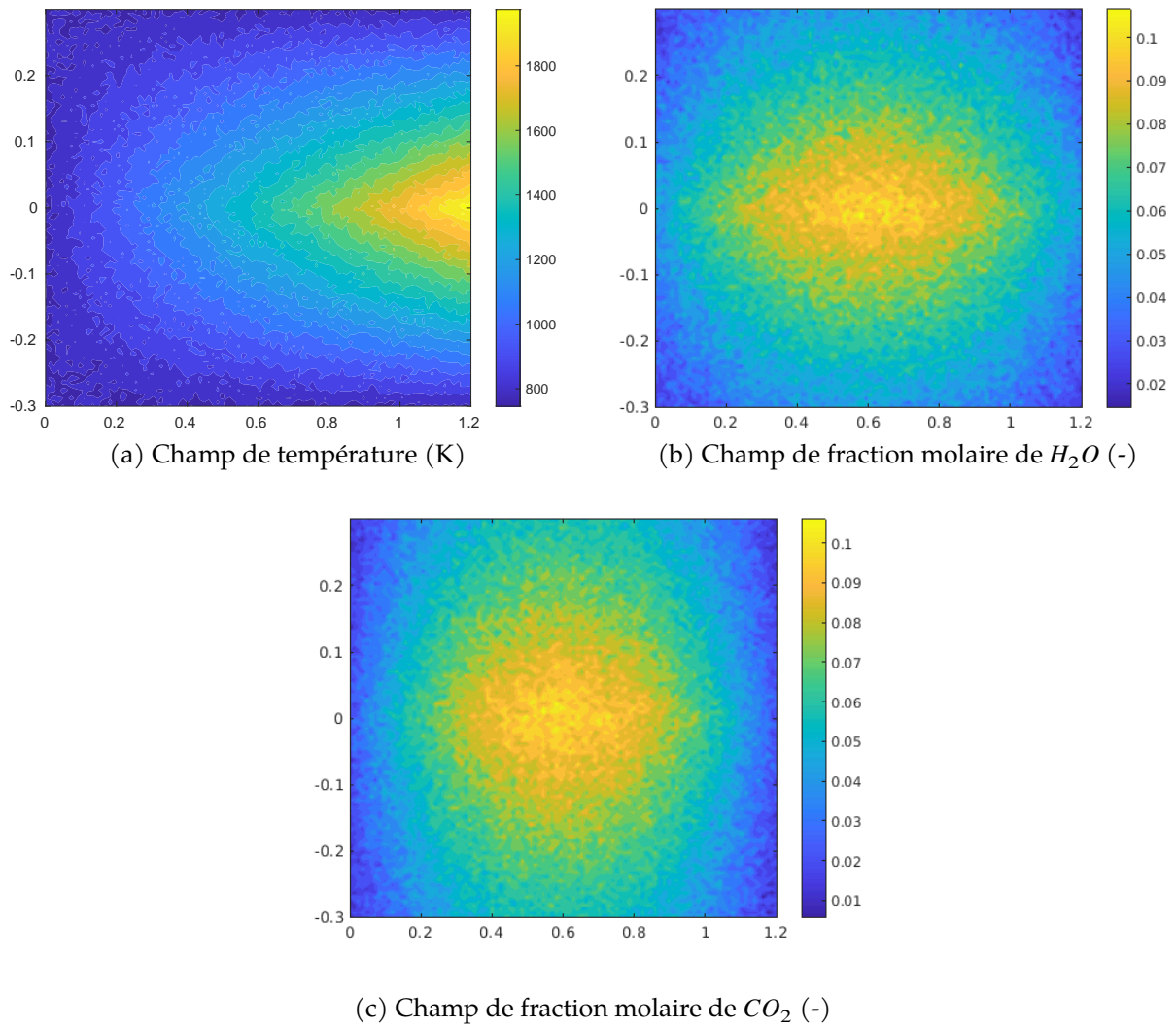


FIG. 4.16 : Champs thermophysiques - Cas bruité

Paramètre	Valeur
Base de données	500 points
Méthode de tirage	Halton
Temps moyen de calcul d'un point de référence	23.4 s
Nombre de neurones	7
Nombre de paramètres	50
Valeur de $\gamma$	44
Nombre d'époques	10000
Temps d'entraînement	138.7 S
RMSE	3.781 $kW.m^{-3}$

TAB. 4.8 : Caractéristiques du réseau de neurones et de l'entraînement - Cas bruité

l'interpolation par ANN et la courbe de référence visée, calculée sans bruit numérique. La courbe de divergence de flux bruitée est également tracée.

La première conclusion que nous pouvons tirer de l'ensemble de ces résultats est l'efficacité du réseau à lisser sa réponse malgré le bruit numérique fort sur les données d'entraînement. La divergence du flux le long de la ligne centrale est en effet particulièrement proche de la référence sans bruit alors même que les calculs Monte Carlo d'entraînement sont assez éloignés de la fonction objectif. Cette capacité à généraliser une solution malgré des erreurs dans les données d'entraînement est une des grandes forces des réseaux de neurones. Plus spécifiquement, la technique de régularisation adaptative (voir chapitre 2) utilisée ici fonctionne particulièrement bien alors que le nombre d'époques est élevé. Le réseau ne tombe donc pas dans le sur-apprentissage (voir section 2.2.5) et ne copie pas le bruit.

## 4.5 Ouverture et perspective

### 4.5.1 Passage aux applications industrielles et couplage

Ces travaux de thèse s'inscrivent globalement dans une démarche d'amélioration continue des méthodes de modélisation de la thermique des chambres au sein de chez Safran. L'ensemble de ce manuscrit recense ainsi les travaux menés sur cas académiques afin de valider, tester et mettre à l'épreuve la nouvelle méthode de modélisation du rayonnement développée ces trois dernières années. Nous avons montré par le biais de plusieurs cas d'étude variés dans leurs particularités, que les réseaux de neurones avaient le potentiel pour remplacer ou tout du moins suppléer les méthodes de résolution déterministe ou stochastique présentées dans ce manuscrit.

La souplesse et l'adaptabilité de l'outil permettent également de l'utiliser sans prérequis théorique et avec très peu d'hypothèses de départ. Quels que soient les cas traités, le nombre de neurones nécessaires à la modélisation précise du flux ou de la divergence de flux reste faible. Un entraînement pleinement intégré à la chaîne de modélisation est donc envisageable pour un coût de calcul global en temps et en RAM accessible.

Les contraintes posées par le contexte industriel autour de ces travaux ont donc été respectées. Les essais sur cas réels peuvent donc être envisagés à l'avenir. Sans avoir la prétention d'être exhaustifs, les cas académiques traités ici ont permis de lever un certain nombre de questions sur la robustesse et les capacités d'adaptation des réseaux de neurones. Néanmoins, le passage à des cas industriels représente

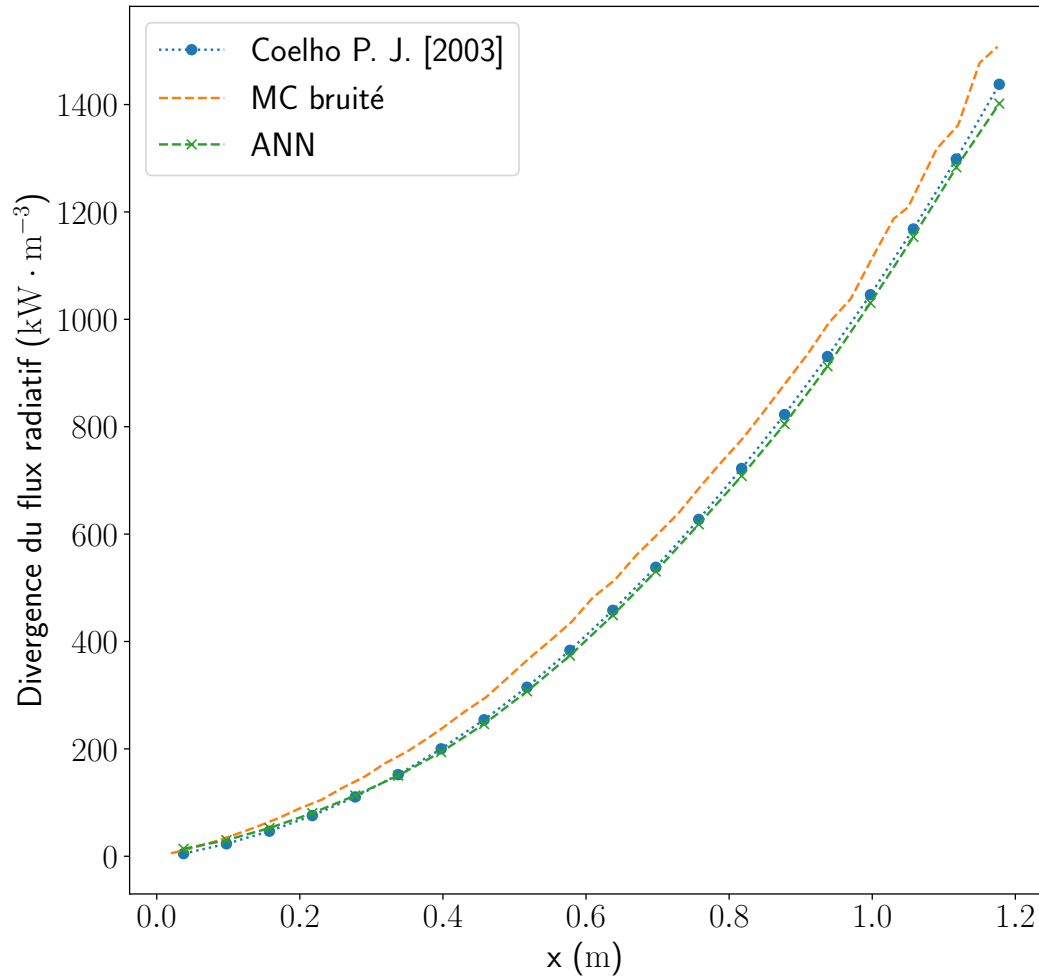


FIG. 4.17 : Divergence du flux radiatif sur la ligne centrale ( $r = 0$ ) en cas de bruit numérique gaussien

un réel tremplin, notamment vis à vis de la sensibilité à la géométrie. Ce cap n'a malheureusement pas été franchi pendant cette thèse faute de temps. Néanmoins, un réel travail préliminaire a été mené pour préparer le passage à un cas industriel de chez Safran. La chambre d'étude envisagée est celle du moteur TP400 équipant notamment l'A400M. Les différents séjours en entreprise ont déjà permis de préparer l'utilisation de nos codes sur le calculateur de l'entreprise. Le code Monte Carlo, celui dédié à l'entraînement des réseaux de neurone et les scripts de couplage ont pu être compilés et exécutés sur le matériel de l'entreprise. Passé ce stade, des problèmes informatiques de comptabilité entre les fichiers de géométries et le code de rayonnement se sont présentés et n'ont pas pu être levés par manque de temps. Nous précisons malgré tout que ces difficultés ne sont pas du tout insurmontables mais nécessitent quelques opérations chronophages sur les formats de fichier.

Deux couplages sont envisageables. Dans un premier temps, l'outil pourra être testé sur un champ thermophysique fixe mais non paramétrique. Les données de température, de pression et de fraction molaire des espèces seront obtenues par lecture du maillage. Notons que sur un cas réel, la pression et la fraction volumique de suie ne sont pas constantes et qu'il faudra donc les intégrer comme variable d'entrée dans l'ANN. Ce couplage faible est donc proche dans la démarche des études menées ici. Un réseau de neurones spécifique devra être utilisé pour la modélisation du flux aux parois et de la divergence de flux dans le domaine.

Le second type de couplage consiste à intégrer pleinement l'outil dans la chaîne de calcul. La résolution du champ de divergence de flux sera menée pour chaque itération du solveur CFD et ce terme source sera alors injecté dans l'équation de conservation de l'énergie afin de prendre en compte l'influence du rayonnement sur le champ thermophysique. Nous sommes alors ici sur un couplage fort, permettant d'étudier par exemple les interactions entre la turbulence et le rayonnement thermique du fluide. En pratique, la résolution du rayonnement n'est menée que toutes les 10 itérations environ. Il sera alors nécessaire de reconstituer à chaque fois une base de données par méthode de Monte Carlo et d'entraîner à nouveau le réseau de neurones. A mesure que les paramètres tendront de plus en plus vers une solution globale, l'entraînement sera alors de plus en plus rapide si le champ thermophysique change peu entre deux itérations.

### 4.5.2 Un interpolateur complexe global

---

Comme cela a été évoqué à plusieurs reprises dans l'ensemble du manuscrit, l'algorithme développé ici peut fondamentalement être considéré comme une méthode d'interpolation complexe au sens large. Contrairement à la plupart des méthodes bien connues, les réseaux de neurones autorisent de prendre en données d'entrée des grandeurs complètement hétérogènes entre elles et de construire des liens entre chacun des points de référence. De ce point de vue, le potentiel du couplage entre méthode de Monte Carlo et réseaux de neurones bayésiens ne se limite pas à la modélisation du rayonnement thermique dans les chambres de combustion mais peut s'étendre à d'autres horizons. Plus globalement, les réseaux de neurones utilisés de la sorte peuvent avoir un intérêt dès lors que l'on cherche à passer d'un ensemble discret au domaine continu. Ils deviennent donc un complément très intéressant aux algorithmes de Monte Carlo dont la force mais également la faiblesse se trouvent justement dans son fonctionnement en « points sondes ».

Dans cette idée, la méthode proposée a été appliquée à un cas d'étude complètement différent en collaboration avec des chercheurs spécialisés dans la thermique du bâtiment et les technologies photovoltaïques. Bien que la thématique abordée soit éloignée de nos propres applications, les mécanismes dans le processus de modélisation sont proches.

L'objectif est de disposer d'un outil permettant l'estimation des besoins en énergie d'un bâtiment simplifié (climatisation en été et chauffage en hiver), équipé ou non de panneaux solaires et/ou d'une toiture végétalisée. En se basant sur une température cible de 19°C à atteindre dans le bâtiment, le besoin en chauffage ou en climatisation a été estimé par méthode de Monte Carlo à différentes périodes de l'année, à différents endroits du globe et dans plusieurs configurations. L'approche en espace de chemin et la double randomisation permettent de remonter par un seul calcul de Monte Carlo à des estimations de besoins en énergie. Une approche similaire a notamment été proposée dans les travaux de VILLEFRANQUE et al. [108].

Une base de données a été construite en fonction des différentes variables (paramètres météorologiques, position géographique, configuration du bâtiment, date, etc.) de chaque point de fonctionnement. Un réseau de 25 neurones dans la couche cachée a ensuite été entraîné sur une base de 6500 points de référence incluant 5 villes du monde, avec ou sans panneaux photovoltaïques. L'outil a permis de modéliser sur l'ensemble d'une année type la puissance en chauffage/refroidissement nécessaire pour respecter la température de consigne d'un bâtiment. L'énergie globale a ensuite été calculée par intégration temporelle. Après entraînement, l'outil était ainsi finalement capable de donner une estimation de l'énergie nécessaire à 5 % près par rapport à un calcul complet par Monte Carlo en quelques microsecondes malgré des fluctuations très importantes de la fonction cible au cours des heures. Cette étude a fait l'objet de la rédaction d'un papier proposé à la conférence IHTC 17<sup>3</sup> organisée à Cap Town en Août 2023. Il a été co-écrit avec Thomas Villemin, Olivier Farges, Gilles Parent et Rémy Claverie. Le script de l'article est proposé dans l'annexe D.

Au delà de cet exemple, le couplage entre méthode de Monte Carlo et réseaux de neurones bayésiens est pertinent dès lors que l'on souhaite étendre l'utilisation des méthodes de Monte Carlo à des grandeurs continues et dépendantes de beaucoup de paramètres. En physique, ce type d'outil a donc toute sa place, notamment dans le cadre de réduction de modèle par exemple.

## Références

- [1] L. PIERROT. « Développement, étude critique et validation de modèles de propriétés radiatives infrarouges de CO<sub>2</sub> et H<sub>2</sub>O à haute température. Applications au calcul des transferts dans des chambres aéronautiques et à la télédétection. » Thèse de doct. 1997.
- [2] P. J. COELHO, P. PEREZ et M. EL HAFI. « Benchmark numerical solutions for radiative heat transfer in two-dimensional axisymmetric enclosures with nongray sooting media ». In : *Numerical Heat Transfer, Part B : Fundamentals* 43.5 (2003), p. 425-444.
- [67] M. LESHNO et al. « Multilayer feedforward networks with a nonpolynomial activation function can approximate any function ». en. In : *Neural Networks* 6.6 (jan. 1993), p. 861-867. DOI : [10.1016/S0893-6080\(05\)80131-5](https://doi.org/10.1016/S0893-6080(05)80131-5).
- [69] M. T. HAGAN et al. *Neural Network Design*. 2<sup>e</sup> éd. Boston : PWS Publishing Co., 1996.
- [10] N. VILLEFRANQUE et al. « A path tracing Monte Carlo library for 3-D radiative transfer in highly resolved cloudy atmospheres ». In : *Journal of Advances in Modeling Earth systems* 11 (2019).
- [42] W. MALKMUS. « Random Lorentz Band Model with Exponential-Tailed S<sup>-1</sup> Line-Intensity Distribution Function\* ». In : *Journal of the Optical Society of America* 57.3 (1<sup>er</sup> mars 1967), p. 323. DOI : [10.1364/JOSA.57.000323](https://doi.org/10.1364/JOSA.57.000323).
- [43] P. RIVIÈRE et A. SOUFIANI. « Updated band model parameters for H<sub>2</sub>O, CO<sub>2</sub>, CH<sub>4</sub> and CO radiation at high temperature ». In : *International Journal of Heat and Mass Transfer* 55.13 (2012), p. 3349-3358. DOI : [10.1016/j.ijheatmasstransfer.2012.03.019](https://doi.org/10.1016/j.ijheatmasstransfer.2012.03.019).

<sup>3</sup>International Heat Transfer Conference

- [47] R. GOODY et al. «The correlated-k method for radiation calculations in nonhomogeneous atmospheres». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 42.6 (1989), p. 539-550.
- [50] F. LIU et G. J. SMALLWOOD. «Application of the statistical narrow-band correlated-k method to low-resolution spectral intensity and radiative heat transfer calculations Effects of the quadrature scheme». In : *Int. J. Heat Mass Transfer* 43 (2000).
- [53] A. A. LACIS et V. OINAS. «A description of the correlated-k distribution method for modeling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres». In : *J. Geophysical Research* 96 (1991).
- [54] V. GOUTIERE, F. LIU et A. CHARETTE. «An assessment of real-gas modelling in 2D enclosures». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 64.3 (2000), p. 299-326. DOI : [10.1016/S0022-4073\(99\)00102-8](https://doi.org/10.1016/S0022-4073(99)00102-8).
- [93] V. EYMET et al. «Null-collision meshless Monte-Carlo - Application to the validation of fast radiative transfer solvers embedded in combustion simulators». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 129 (2013), p. 145-157.
- [101] J. TAINE et A. SOUFIANI. «Gas IR Radiative Properties : From Spectroscopic Data to Approximate Models». In : *Advances in heat transfer* 33 (1999), p. 295-414.
- [102] J. P. RIGOL, C. H. JARVIS et N. STUART. «Artificial neural networks as a tool for spatial interpolation». In : *International Journal of Geographical Information Science* 15.4 (2001), p. 323-343.
- [103] V. NEVTIPILOVA. «Testing Artificial Neural Network (ANN) for Spatial Interpolation». In : *Journal of Geology & Geosciences* 03.2 (2014). DOI : [10.4172/2329-6755.1000145](https://doi.org/10.4172/2329-6755.1000145).
- [104] N. S.-N. LAM. «Spatial Interpolation Methods : A Review». In : *The American Cartographer* 10.2 (1983), p. 129-150.
- [105] L. DINIS, R. NATAL JORGE et J. BELINHA. «Analysis of 3D solids using the natural neighbour radial point interpolation method». In : *Computer Methods in Applied Mechanics and Engineering* 196.13 (2007), p. 2009-2028.
- [106] N. PLAZIAC. «Image interpolation using neural networks». In : *IEEE Transactions on Image Processing* 8.11 (1999), p. 1647-1651.
- [107] J. GO, K. SOHN et C. LEE. «Interpolation using neural networks for digital still cameras». In : *IEEE Transaction on Consumer Electronics* 46.3 (2000), p. 610-616.
- [108] N. VILLEFRANQUE et al. «The “teapot in a city” : A paradigm shift in urban climate modeling». In : *Science Advances* 8 (2022).



# Conclusion et Perspectives

## Une solution nouvelle pour répondre à des besoins croissants

---

Les travaux de recherche menés pendant cette thèse ont permis le développement d'une nouvelle méthode de modélisation du rayonnement. En combinant les méthodes de Monte Carlo avec un modèle de gaz fin pour la résolution de l'ETR et les réseaux de neurones artificiels, nous avons pu tirer parti du meilleur de ces deux outils. Les contraintes fortes imposées par le contexte dans lequel s'inscrivent ces travaux ont ainsi pu être respectées. La formulation de la méthode de Monte Carlo a notamment permis de supprimer toute nécessité de stocker les nombreuses informations des différentes bandes spectrales pour chaque point calculé. Le calcul de la grandeur intégrée est immédiat et sans intermédiaire. De la même manière, l'utilisation de modèles spectraux précis et adaptés aux conditions d'opération est désormais accessible pour un coût de calcul très faible. La méthode développée offre donc une alternative crédible et efficace pour la modélisation en champ du rayonnement.

Plus précisément, nous avons démontré que les réseaux de neurones bayésiens étaient capables de traiter des cas très hétérogènes en température et en espèces. Nous avons pu reconstruire fidèlement des champs de divergence de flux radiatif à partir de quelques points de référence échantillonnés dans le milieu dans des configurations différentes. L'algorithme d'entraînement de Levenberg-Marquardt couplé à la régularisation bayésienne permet d'automatiser l'apprentissage et l'adaptation du réseaux au cas d'étude présenté. Il est d'ailleurs intéressant et assez surprenant de remarquer le peu de neurones nécessaires à la modélisation précise du rayonnement malgré la complexité physique de ce phénomène. Moins de dix neurones sont utilisés dans les différents cas hétérogènes traités lorsque les conditions physiques sont utilisées comme données d'entrée. Cette *simplicité* au sens du nombre de degrés de liberté est une application directe du principe de parcimonie du rasoir d'Ockham et est fondamentalement lié au paradigme bayésien. Les avantages sont nombreux car ils permettent de diminuer le temps d'entraînement du réseau à son minimum et de fortement limiter les tendances au sur-apprentissage. Cette dernière caractéristique est d'ailleurs mise en exergue dans l'étude du cas artificiellement bruité. L'outil a été capable de prendre en compte les erreurs sur les données d'entrée et de les corriger, jusqu'à reproduire presque exactement les résultats du champs hétérogène de référence.

Bien qu'il n'ait pas été possible d'appliquer notre méthode sur un cas réel de chambre par faute de temps, nos codes de calculs ont pu être ajoutés aux outils Safran et compilés sur les serveurs de l'entreprise. Nous espérons que des tests de validation à l'échelle industrielle pourront être menés par Safran en commençant probablement par l'étude de la chambre du moteur TP400 équipant l'avion de transport A400M. Ce cas d'étude était celui qui avait été envisagé lors de la définition de ce projet de thèse.

## De nouvelles applications envisagées pour l'industrie aéronautique

---

Au delà des très bons résultats obtenus et du potentiel de la méthode présentée dans cette étude, cette thèse a permis de développer un ensemble de codes en langage C fonctionnels et indépendants. Ils sont capables sans s'y limiter, de produire une base de données solide puis d'entraîner un réseau de neurones adapté à la situation pour modéliser un champ d'une grandeur radiative intégrée. L'autonomie de ces codes autorisent ainsi le déploiement et le couplage de ces méthodes avec la plupart des logiciels libres ou commerciaux de CFD ou de combustion. Bien qu'à l'origine conçues et développées pour être couplées à Fluent, il est tout à fait envisageable d'utiliser ces bibliothèques en complément de logiciels comme AVBP<sup>4</sup> par exemple. Les champs produits par ce type de logiciel ne sont pas moyennés comme ceux obtenus par méthode RANS. Ils présentent donc des hétérogénéités très fortes et à faible échelle dans l'ensemble du domaine. L'utilisation de notre méthode sur ce type de cas pourrait donc se révéler très instructive sur la manière de réagir du réseau.

Les différentes présentations de ces travaux au sein du réseau combustion du groupe Safran ont également suscité la curiosité d'équipes travaillant sur des thématiques proches, notamment du côté de la filiale Safran Helicopter Engines. Leurs problématiques étant similaires à celles des moteurs d'avion, une extension de nos méthodes aux chambres de combustion d'hélicoptères serait donc pertinente.

Enfin et compte tenu de la précision atteinte sur cas académiques, les outils développés pourraient être pertinents pour l'étude de signature infrarouge de jet en sorti de tuyère. Ces applications requièrent en générale une très haute résolution du rayonnement si bien que des modèles de type RPR sont souvent utilisés. Jusqu'à présent, ces simulations, bien que non couplées la plupart du temps, se heurtaient aux mêmes limites dues à la modélisation en champ par méthode des Ordonnées Discrètes. Les méthodes de Monte Carlo permettent d'utiliser les modèles RPR sans contrainte particulière, voire même de travailler dans l'absolu sur les spectres réels des différentes espèces rayonnantes. Le réseau de neurones serait alors utilisé de la même manière que dans ces travaux pour reconstruire un champ complet et précis à partir des données calculées.

## Ouverture : une méthode d'interpolation aux possibilités vastes

---

Fondamentalement, la méthode développée dans ces travaux ne se limite pas aux applications citées précédemment. Les réseaux de neurones et les méthodes de Monte carlo sont des outils qui dépassent les applicatifs. Cette méthode peut également être envisagée dès lors que l'on souhaite généraliser des résultats ponctuels obtenus par méthode de Monte Carlo à un domaine continu, qu'il soit spatial ou temporel par exemple. Pour illustrer nos propos, nous avons par exemple utilisé le concept faisant l'objet de ces travaux pour l'estimation des besoins en énergie d'un bâtiment simplifié. L'objectif était de mettre en évidence l'incidence de la présence de panneaux solaires sur la consommation énergétique annuelle d'un bâtiment. Des simulations Monte Carlo ont été menées afin de calculer la puissance consommée par le bâtiment à des dates échantillonnées tout au long d'une année type. L'ensemble de ces données a ensuite été utilisé pour l'entraînement d'un réseau de neurones dont le rôle était d'estimer les puissances en tout temps du bâtiment d'après des données d'entrée variées (conditions météorologiques, localisation, configuration du bâtiment, etc.). Cet exemple étudié en marge de ce sujet de thèse et en collaboration avec d'autres chercheurs spécialisés a permis de démontrer le potentiel de cette méthode et ainsi étendre le champ des possibles.

Pour clore ce manuscrit, nous ne pouvons qu'inciter le lecteur à expérimenter les réseaux de neurones

---

<sup>4</sup>Logiciel spécialisé dans la modélisation LES d'écoulement turbulents réactifs, développé par le CERFACS

pour leurs propres travaux sans forcément se limiter à leur utilisation telle que présentée ici. Néanmoins, il est important de garder à l'esprit que ces objets mathématiques ne sont pas la solution à tous les problèmes malgré l'intérêt sans garde fou qu'ils suscitent parfois. Nous avons tenté au cours de ce manuscrit de transmettre au lecteur des intuitifs lui permettant d'appréhender et de conceptualiser ces outils dans le but de les utiliser de manière raisonnée et raisonnable. Malgré la sémantique entourant les réseaux de neurones et le caractère *intelligent* qu'on leur prête, ils ne font que reproduire leur propre réalité. Si les données qu'on leur présente dans le processus d'apprentissage sont erronées, biaisées ou faussées, les résultats produits ne pourront qu'en être une copie. Finalement, cette tendance à reproduire les erreurs qu'on leur montre sont ce que les réseaux de neurones artificiels ont de plus terriblement humain.



# Cycle thermodynamique d'un turboréacteur

Dans cette annexe, nous nous intéressons au cycle thermodynamique d'une turbine à gaz aéronautique. L'objectif est de mettre en équation dans un cas idéal, le fonctionnement d'une turbomachine afin de pouvoir écrire l'expression du rendement thermique de ce type de système. Nous nous plaçons dans le cas simple présenté sur la figure 1.1 en considérant les transformations (compression et détente) comme isentropiques et les gaz comme parfaits. Sous cette condition, les équations de Laplace s'appliquent :

$$PV^\gamma = C_1 \quad (\text{A.1})$$

$$TV^{\gamma-1} = C_2 \quad (\text{A.2})$$

$$T^\gamma P^{1-\gamma} = C_3 = C_1^{1-\gamma} C_2^\gamma \quad (\text{A.3})$$

Avec  $C_1$ ,  $C_2$  et  $C_3$  des constantes dépendantes du gaz parfait étudié et  $\gamma$  le coefficient isentropique égal au rapport entre les capacités thermiques à pression et volume constant.

Le cycle idéal d'une turbomachine (voir figure A.1) comporte quatre phases distinctes :

- (a) Une compression isentropique qui se traduit par une augmentation de la température du gaz parfait suivant l'équation (A.3). En connaissant la pression initiale en entrée du système au point (1) et le taux de compression du compresseur, il est possible de calculer la température au point (2)
- (b) Une augmentation isobare de la température du point (2) au point (3) du gaz. Dans la pratique, cette étape fait référence à l'injection du carburant et à la combustion du mélange. Néanmoins ce processus n'est théoriquement pas nécessaire en soit. Un apport de chaleur externe par effet Joule par exemple suffit en principe pour le fonctionnement d'une turbine à gaz.
- ( $c_1$ ) Une détente isentropique dans la turbine entre les points (3a) et (3b), suffisante pour l'alimentation en énergie du compresseur. De manière analogue au premier point, en connaissant le taux de détente de la turbine, il est possible de calculer la température au point (3b) grâce à l'équation (A.3).
- ( $c_2$ ) Une seconde détente également supposée isentropique au sein de la tuyère pour augmenter la vitesse des gaz et produire la poussée nécessaire à l'aéronef.

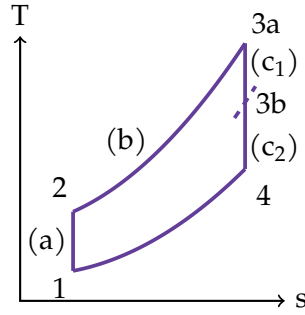


FIG. A.1 : Cycle idéal de Brayton d'une turbomachine simplifiée

Le premier principe de la thermodynamique nous permet d'écrire le bilan énergétique sur chacun des chemins du cycle. Nous identifions par les lettres  $W$  et  $Q$  les différents travaux et transferts thermiques pendant les transformations. Nous avons donc :

$$W_{12} = -c_p(T_2 - T_1) \quad (\text{A.4})$$

$$Q_{23} = c_p(T_3 - T_2) \quad (\text{A.5})$$

$$W_{34} = c_p(T_3 - T_4) \quad (\text{A.6})$$

Nous considérons sur les trajets (a) et (c) les transformations comme adiabatiques. Les pertes  $Q_{12}$  et  $Q_{34}$  sont donc nulles. De même, l'augmentation de température sur le trajet (b) se fait à pression et volume constant. Le travail  $W_{23a}$  est alors également considéré comme nul.

Sur la base de ces trois relations, nous pouvons ainsi écrire la formulation du rendement thermique du cycle  $\eta_{th}$ , égal au rapport entre le travail net récupéré en sortie et la quantité d'énergie apportée par la combustion sur le chemin (b) :

$$\eta_{th} = \frac{\text{Travail net récupéré}}{\text{Apport de chaleur}} = \frac{W_{34} + W_{12}}{Q_{23}} = \frac{c_p(T_3 - T_4) - c_p(T_2 - T_1)}{c_p(T_3 - T_2)} \quad (\text{A.7})$$

Nous notons  $r$  le taux de compression du compresseur, pouvant également être défini dans ces conditions idéales comme l'inverse du taux de dilution de l'ensemble Turbine-Tuyère. Nous avons ainsi :

$$r = \frac{P_2}{P_1} = \frac{P_3}{P_4} \quad (\text{A.8})$$

La relation de Laplace explicitée dans l'équation (A.3) nous permet d'exprimer les températures  $T_1$  et  $T_4$  en fonction de ce rapport et du coefficient isentropique  $\gamma$ . Ainsi, nous pouvons écrire :

$$T_1 = T_2 \cdot \left(\frac{P_2}{P_1}\right)^{\frac{1-\gamma}{\gamma}} = T_2 \cdot r^{\frac{1-\gamma}{\gamma}} \quad (\text{A.9})$$

$$T_4 = T_3 \cdot \left(\frac{P_3}{P_4}\right)^{\frac{1-\gamma}{\gamma}} = T_3 \cdot r^{\frac{1-\gamma}{\gamma}} \quad (\text{A.10})$$

En injectant les équations (A.9) et (A.10) dans l'équation (A.7), nous montrons que :

---

$$\eta_{th} = 1 - r^{\frac{1-\gamma}{\gamma}} = 1 - \left(\frac{1}{r}\right)^{\frac{\gamma-1}{\gamma}} \quad (\text{A.11})$$

Nous montrons également que le taux de compression peut être exprimé en fonction des rapports de température  $\frac{T_2}{T_1}$  et  $\frac{T_3}{T_4}$  ainsi que du coefficient isentropique :

$$\frac{T_2}{T_1} = \frac{T_3}{T_4} = r^{\frac{\gamma-1}{\gamma}} \quad (\text{A.12})$$

Enfin, en injectant l'équation (A.12) dans l'équation (A.11), on obtient finalement une expression du rendement thermique d'une turbine à gaz idéale, en fonction des températures du cycle :

$$\eta_{th} = 1 - \frac{T_4}{T_3} = 1 - \frac{T_1}{T_2} \quad (\text{A.13})$$





# Champs soustraits : comparaisons des méthodes d'interpolation

Dans cette annexe, nous présentons les différents champs de divergence de flux radiatif construits à partir d'un réseau de neurones entraîné sur une base de Halton de 500 points (voir figure [B.1a](#)) et par différentes méthodes d'interpolation (voir figures [B.1b](#) à [B.1d](#)). Chacun de ces champs a ensuite été soustrait à un champ de référence calculé par méthode de Monte Carlo afin de mettre en évidence les erreurs produites par les méthodes testées. Tous les champs de divergence de flux et les champs soustraits sont à la même échelle entre eux et en  $kW.m^{-3}$  afin de pouvoir les comparer.

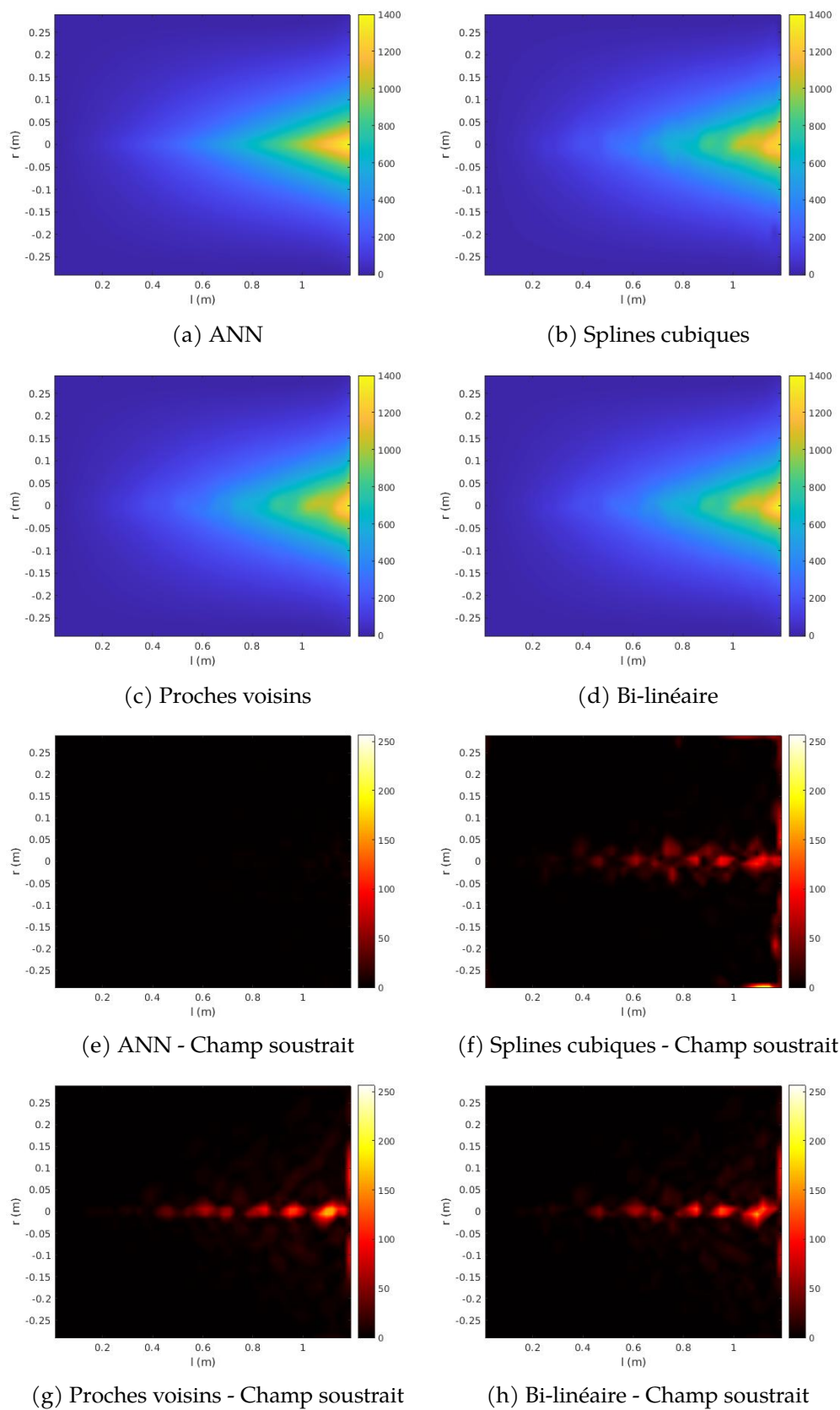


FIG. B.1 : Champs de divergence de flux en  $kW.m^{-3}$  calculés par méthodes d'interpolation et champs soustraits

## Tableaux de valeurs

Cette annexe regroupe les différents tableaux de valeurs permettant de tracer les courbes du chapitre 4 afin qu'elles puissent être reprises ou comparées dans de futurs travaux éventuels. Le tableau C.1 est plus précisément relié à la figure 4.12 tandis que les tableaux C.2 à C.4 font référence au cas «Point Chaud» de la section 4.4.1.

$x$ (m)	MC ( $kW.m^{-3}$ )	ANN ( $kW.m^{-3}$ )
0.0375	2.52	11.38
0.0975	8.38	18.50
0.1575	16.10	25.87
0.2175	25.87	34.09
0.2775	38.25	45.22
0.3375	53.79	59.93
0.3975	72.763	78.43
0.4575	95.86	101.10
0.5175	124.00	128.45
0.5775	158.27	161.25
0.6375	198.25	200.34
0.6975	245.22	246.53
0.7575	302.42	300.62
0.8175	365.34	363.38
0.8775	438.83	435.54
0.9375	528.13	518.45
0.9975	622.71	613.51
1.0575	732.76	722.05
1.1175	852.87	842.92
1.1775	997.23	966.80

ТАВ. С.1 : Valeurs de la divergence du flux sur la ligne centrale  $r = 0$  - Cas à haute pression

---

$x$ (m)	MC ( $kW.m^{-3}$ )	ANN ( $kW.m^{-3}$ )
0.02	-4.72	6.47
0.05	1.98	12.95
0.08	10.81	20.89
0.10	18.08	26.80
0.14	35.41	39.46
0.17	50.02	49.72
0.20	65.93	61.19
0.23	83.54	74.46
0.26	102.64	89.90
0.29	120.71	107.61
0.32	142.51	127.54
0.35	166.23	149.53
0.38	190.46	173.42
0.40	207.11	190.32
0.43	234.20	217.06
0.46	262.34	245.30
0.49	291.84	275.06
0.52	323.92	306.30
0.55	357.62	339.00
0.58	392.37	373.22
0.61	429.14	408.99
0.64	466.50	446.40
0.67	504.63	485.48
0.70	544.34	526.30
0.73	587.80	568.89
0.76	632.09	613.30
0.79	678.17	659.51
0.82	724.60	707.52
0.85	773.80	757.29
0.88	823.47	808.75
0.91	877.04	861.80
0.94	933.56	916.35
0.97	984.28	972.24
1.00	1044.62	1029.32
1.03	1100.96	1087.41
1.05	1144.55	1126.60
1.09	1238.19	1205.79
1.12	1294.35	1265.63
1.15	1356.71	1325.55
1.175	1415.11	1375.34

---

ТАВ. С.2 : Divergence du flux sur la ligne centrale  $r = 0$  - Cas «Point chaud»

$x$ (m)	MC ( $kW.m^{-3}$ )	ANN ( $kW.m^{-3}$ )
0.02	23.175	32.86
0.04	71.85	73.83
0.06	219.99	199.54
0.08	673.12	679.37
0.09	1209.33	1354.06
0.095	1648.75	1824.88
0.1	2240.43	2212.76
0.105	1665.8	1857.17
0.11	1254.87	1423.99
0.12	734.945	768.57
0.14	268.632	264.40
0.16	109.04	120.05
0.18	50.88	69.23
0.20	30.62	48.63
0.225	24.05	38.55
0.25	24.08	35.72
0.275	26.77	36.21
0.30	30.47	38.45
0.325	33.94	41.67
0.35	38.53	45.49
0.375	43.19	49.69
0.40	47.73	54.14
0.425	52.68	58.75
0.45	57.37	63.49
0.475	61.76	68.33
0.50	67.33	73.25
0.55	77.89	83.34
0.60	88.46	93.78
0.65	99.84	104.64
0.70	112.21	116.02
0.75	125.07	128.06
0.80	138.13	140.94
0.85	151.91	154.88
0.90	165.76	170.19
0.95	181.21	187.28
1.00	199.67	206.75
1.05	223.26	229.43
1.1	254.67	256.51
1.15	287.532	289.65
1.175	308.977	309.14

TAB. C.3 : Divergence du flux sur la ligne longitudinale  $r = 0.15$  - Cas «Point Chaud»

$r$ (m)	MC ( $kW.m^{-3}$ )	ANN ( $kW.m^{-3}$ )
-0.29	-2.12	4.17
-0.28	-1.37	4.56
-0.26	-0.18	5.36
-0.24	1.06	6.20
-0.22	1.86	7.09
-0.20	3.62	8.02
-0.18	4.65	9.01
-0.16	5.66	10.06
-0.14	6.87	11.18
-0.12	8.34	12.40
-0.10	9.60	13.75
-0.08	11.14	15.27
-0.06	13.52	17.08
-0.04	16.21	19.33
-0.02	20.03	22.34
0	26.488	26.67
0.02	23.20	29.18
0.04	25.17	35.21
0.06	36.96	48.66
0.08	69.85	80.12
0.1	163.55	162.91
0.12	434.91	419.14
0.13	728.93	744.74
0.14	1249.92	1398.83
0.1425	1437.35	1619.83
0.145	1669.19	1843.50
0.1475	1941.14	2047.57
0.15	2259.74	2211.64
0.1525	1941.05	2039.74
0.155	1659.51	1824.21
0.1575	1427.20	1588.42
0.16	1228.44	1357.54
0.17	692.407	693.79
0.18	396.407	375.27
0.20	135.584	131.21
0.22	44.98	55.37
0.24	13.72	27.23
0.26	3.38	14.88
0.28	-1.43	8.62
0.29	-2.12	6.60

Tab. C.4 : Divergence du flux sur la ligne radiale  $x = 0.1$  - Cas «Point Chaud»





# Script du papier présenté à la conférence IHTC-17

Cette annexe contient le script de notre article de conférence présenté pour l'IHTC 2023. Il a été co-écrit avec Thomas villemain en tant que premier auteur et Olivier Farges, Gilles Parent et Rémy Claverie en partenariat avec le LEMTA et le laboratoire Cerema Est de Tomblaine. A l'heure de la rédaction de cette annexe, le papier n'a pas encore été publié dans les *proceedings* de la conférence. La version présentée ici peut donc varier avec le papier officiel publié par les organisateurs de l'évènement.

La problématique abordée dans cette communication concerne l'étude de la thermique d'un bâtiment simplifié dans différentes configurations. En particulier, l'objectif est d'en estimer les besoins énergétiques pour le chauffage ou la climatisation sur une année complète en fonction d'une température de consigne fixée à 19°C. Nous souhaitons ainsi évaluer l'impact de la présence d'une toiture végétalisée et/ou de panneaux photovoltaïques sur la consommation et la puissance nécessaires. Cette étude a été menée dans différents endroits du globe d'après des relevés météorologiques s'étendant sur plusieurs années. Le deuxième objectif de cette étude visait à produire un outil permettant l'estimation rapide des besoins énergétiques d'un bâtiment, en se basant sur les calculs stochastiques effectués. Nous avons donc appliqué la méthodologie développée pendant cette thèse à ce cas très concret. Nous avons ainsi montré que la méthode développée dans le cadre de ces travaux de thèse ne se limite pas à son application originelle mais peut s'étendre à tout travaux de réduction de modèles basés sur des points sondes calculés par Monte Carlo. La particularité de ce cas d'étude par rapport aux travaux menés dans ce manuscrit provient du fait que les données météorologiques d'entrée (hygrométrie, angle du soleil, couverture nuageuse, température extérieure, etc.) proviennent de relevés réels et donc potentiellement bruités. Les données sont également très hétérogènes entre elles. Par exemple, l'entrée relative à la date est une valeur entière exprimée en seconde tandis que la présence de panneaux est une valeur booléenne. Nous démontrons donc par ailleurs la souplesse et la robustesse de l'outil face à ce type d'application.

## *Simulation of a rooftop photovoltaic system : a focus on the energy performance of the building*

### Résumé

---

Les systèmes photovoltaïques appliqués ou intégrés aux bâtiments constituent des possibilités prometteuses de développement des technologies renouvelables en milieu urbain. Cependant, le champ

photovoltaïque agit comme un masque solaire et aéraulique pour le bâtiment sous-jacent, ce qui a un impact sur les besoins énergétiques du bâtiment en matière de climatisation et de chauffage. Dans ce travail, le bilan énergétique d'un bâtiment simplifié avec un champ photovoltaïque intégré est modélisé à l'aide d'une approche de Monte Carlo. Cette méthode est appliquée pour évaluer les besoins de climatisation ou de chauffage du bâtiment pendant plusieurs années. Deux types de toitures sont étudiés : une toiture en béton avec une seule couche et une toiture verte avec deux couches supplémentaires (substrat et végétation). Les résultats de la simulation de Monte Carlo sont discutés et exploités avec un réseau neuronal artificiel de type *FeedForward* pour interpoler la consommation du bâtiment sur une année type en fonction des paramètres météorologiques.

## Abstract

---

Building Applied Photovoltaics and Building Integrated Photovoltaic systems are promising opportunities to develop renewable technologies in urban environment. However, the photovoltaic field acts as a solar and aeraulic mask for the underlying building that impacts the building's energy needs for air conditioning and heating. In this work, the energy balance of a simplified building with integrated photovoltaic field is modelled using a Monte Carlo approach. This method is applied to evaluate the needs for cooling or heating of the building for several years. Two kinds of roof decks are studied : concrete roof with a single layer and green roof with two additional layers (substrate and vegetation layers). The results from the Monte Carlo simulation are discussed and exploited with a FeedForward Artificial Neural Network to interpolate the building consumption according to the meteorological parameters.

## Introduction

---

In recent decades, energy needs have increased significantly and the exploitation of solar energy through photovoltaic panels (PV panels) is particularly interesting. In 2021, the cumulative installed solar capacity worldwide was approximately 843 GW [109]. In urban areas, the installation of PV panels on building roofs is generally favoured because they are very well exposed to solar radiation and urban space can be saved. According to BÓDIS et al. [110], exploitation of rooftops in European Union could produce up to 680 TWh of electricity per year that represents 24.4 % of EU's annual electricity consumption. Moreover, the problem of energy consumption of buildings is also a major issue in urban environments [111]. Thus, in recent years, the installation of combined systems, Building Applied Photovoltaics (BAPV) and Building Integrated PV (BIPV), has intensified [112]. However, the radiative and aeraulic mask effects created by such a PV field will disturb the energy balance of the underlying building and therefore its energy consumption, especially for heating and cooling. The findings are variable depending on the type of climate where the building is located [113, 114, 115]. The general observation is that PV modules lead to a reduction in cooling demand in hot periods and an increase in energy demand in cold periods due to the shading effect created by the PV field. Nevertheless, it is important to perform simulations to validate these statements for different configurations.

In the present study, the energy balance of a simplified building with integrated rooftop PV field is modelled using Monte Carlo approaches [116]. This method has proven to be a successful approach to urban issues and has many advantages : renewed physical images, insensitivity of computation times to the dimension of the integration domain and to the number of coupled phenomena, insensitivity to the geometrical complexity at temporal and spatial scales [108, 117, 118]. The objective of the model is to estimate the building's energy demand for cooling and heating to maintain a target temperature for the air inside the building. Then, the building's energy needs are found by the integration of the

Thermal parameters	Vegetation layer	Substrate layer	Concrete layer
$\lambda$ ( $\text{W m}^{-1} \text{K}^{-1}$ )	0.3	1	1
$\rho$ ( $\text{kg m}^{-3}$ )	582	1200	1700
$C_p$ ( $\text{J kg}^{-1} \text{K}^{-1}$ )	4800	840	840
Thickness (m)	0.06	0.1	0.15

TABLE D.1 : Thermal properties of the green roof structure.

power over time using double randomization technique [119, 120]. Two configurations are analyzed : (i) the roof is a classic concrete slab, (ii) the concrete slab is overhung by an extensive green roof with substrate and vegetation layers. The simulations are conducted for different locations in the world using Typical Meteorological Year (TMY) data for weather parameters (solar irradiation, air temperature and wind speed) : Reykjavik (Iceland), Las Vegas (USA), Cape Town (South Africa), Melbourne (Australia), Nancy (France).

Finally, a Feedforward Artificial Neural Network (FANN) is implemented to exploit the information obtained from the Monte Carlo random paths. For each path, the estimated power and the meteorological parameters are stored, then used to train a neural network. Time evolution of the total power for heating and cooling of the building is obtained by a multivariate regression function implicitly determined by the trained neural network according to the meteorological parameters. As a final step, the energy is calculated by time integration of the power. The results obtained by the neural network are compared with the simulation points obtained by the Monte Carlo method.

## Studied configuration

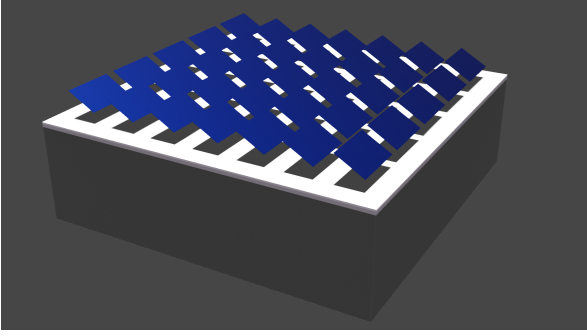
The studied geometry is a simplified building with a flat roof. The height of the building is 3 m, the base is square and 10 m long inside the building. The thickness of the walls is set at 0.2 m. The walls and the floor are supposed to be made of concrete. For the roof two cases are studied, the first one with only a concrete roof (figure D.1a) and the second one with three layers (figure D.1b). The two additional layers correspond to a situation where an extensive green roof is installed on the building. The top layer represents a vegetation layer and the middle one is the substrate layer where the plants take root. The thermal conductivities, material densities and heat capacities of the different layers of green roof are provided in tableau D.1. Moreover, the building is studied for two different rooftop configurations : (i) without PV field, (ii) with rooftop PV field.

When present, the photovoltaic field almost entirely covers the roof and the PV field acts as a solar mask and aeraulic obstacle for the underlying building. This selected situation should provide insights about how the energy requirements of a building change with the presence or the absence of a rooftop PV field.

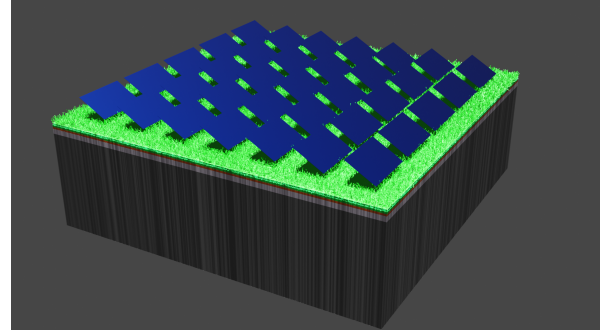
The temperature at the bottom of the ground is set at  $17^\circ\text{C}$ <sup>1</sup>. The target temperature for the air inside the building is  $19^\circ\text{C}$ . For convective heat transfer between indoor air and indoor walls, the convective heat transfer coefficient is set at  $10 \text{ W m}^{-2}$ . For outside convective heat transfer coefficient, the formula from McADAMS [121] is used.

The dimensions of the PV panels are quite standard with a length of 1.6 m and width of 1 m. The

<sup>1</sup>Strictly speaking, this temperature should depend on the location of the building and the depth in the ground, this could be discussed further. As this temperature is not a major issue in this work, it is set at  $17^\circ\text{C}$  regardless of the geographical location studied.



(a) Building geometry with concrete configuration.



(b) Building geometry with green roof configuration.

FIG. D.1 : Building geometry and rooftop photovoltaic field.

distance between the bottom of the PV panels and the top of the roof is set at 0.5 m which is rather usual for rooftop mounting configuration. The PV panels are facing south direction and the tilt angle is  $30^\circ$ . The PV field is composed of 7 rows of 5 panels each, for a total of 35 panels.

## Methods

In the present work, the Monte Carlo statistical method is used to evaluate physical quantities of interest at a probe point and a feedforward artificial neural network is used as a nonlinear multivariate interpolator based on data obtained with Monte Carlo simulations. The proposed model, based on the Monte Carlo method, has many advantages including the handling of the 3D geometric complexity of the building and its environment. Moreover, the Monte-Carlo method has already proven its ability to handle complex geometries [116] as well as to take into account transient phenomena with high variability as solar energy [122].

## Monte Carlo modelling of heat transfer

In this framework, the physical quantity is expressed as the expectation of a random variable. In the context of this paper, the physical quantity is the heating or cooling energy requirement of the building to maintain a given air temperature. The total energy needs are computed from the integration over time of the power  $P(t)$  (see équation (D.1)), both for cooling and heat requirements. The power to be supplied to the building to meet its energy needs is calculated as the power exchanged between the air inside the room ( $T_{\text{target}}$ ) and the wall temperature ( $T_w(\mathbf{x}, t)$ ) estimated from the Monte Carlo algorithm.

$$P(t) = \mathbb{E}[\mathcal{P}(t)] = \int_{\partial S} dS p_{\mathbf{X}}(\mathbf{x}) \left[ h_{\text{cv}}^{\text{in}} \left[ (T_w(\mathbf{x}, t) - T_{\text{target}}) \mathcal{H}(T_w(\mathbf{x}, t) > T_{\text{target}}) + (T_{\text{target}} - T_w(\mathbf{x}, t)) \mathcal{H}(T_w(\mathbf{x}, t) > T_{\text{cooling}}) \right] \right] \quad (\text{D.1})$$

where  $T_{\text{target}}$  is the target temperature for the air inside the building (considering a perfectly-mixed air),  $T_{\text{cooling}}$  is the limit temperature at which it is considered necessary to have air-conditioning,  $T_w(\mathbf{x}, t)$  is the temperature of the wall of the room,  $p_{\mathbf{X}}$  is the density probability function to sample a position on a wall inside the room and  $h_{\text{cv}}^{\text{in}}$  is the indoor convective coefficient.  $T_w(\mathbf{x}, t)$  is also expressed as the expectation of a random variable  $\Theta(\mathbf{x}, t)$  and thanks to double randomization [119, 120], the quantities

$P(t)$  and  $E_{\text{total}}$  can be obtained using a single Monte Carlo algorithm. The Monte Carlo algorithm for the evaluation of  $T_w(\mathbf{x}, t)$  is described in ?? 1 and the corresponding figure is shown in figure D.2 where conductive, convective and radiative heat transfers have a statistical description.

$$E_{\text{total}} = \mathbb{E}[\mathcal{E}] = \int_{t_I}^{t_F} d\tau p_\tau(\tau) \mathcal{P}(\tau) \quad (\text{D.2})$$

where  $p_\tau$  is the density probability function to sample the time over the interval  $[t_I, t_F]$  (this function is chosen uniformly in this work).

To estimate  $T_w(\mathbf{x}, t)$ , a path starts at a random position  $\mathbf{x}$  on one of the walls of the room (including roof and ground surfaces) at a given time. Then the path goes backward in time until the initial or a known temperature is reached. It has to be noted that walls of the building are not considered isothermal. The Monte Carlo algorithm accounts for fully coupled heat transfers in the building, including conduction in the wall, convection and radiation which occur at the surfaces.

For the transient conductive heat transfer inside the solid domains, diffusive random walks are used. At each step of the random walk in conduction, the current time is decreased of a quantity depending on the material thermal inertia. When surfaces of the solid domain are encountered, the flux continuity equation is written and the path couples conductive, convective and radiative heat transfers as written in équation (D.3). Only for these interfaces with the outdoor environment, the meteorological data are needed (air temperature, solar irradiation, wind speed). For solar irradiation, depending on the position of the sun and the shading effect created by the photovoltaic field, its contribution to the wall temperature can be zero.

$$T(\mathbf{y}, t) = \mathbb{P}_{\text{cd} \rightarrow \text{cd}} T(\mathbf{y} - \delta_r \mathbf{n}, t) + \mathbb{P}_{\text{cd} \rightarrow \text{cv}} T_{\text{air}}(t) + \mathbb{P}_{\text{cd} \rightarrow \text{rad}} \int_{2\pi} dp_\Omega(\omega) T_{\text{rad}, -\omega}(\mathbf{y}, t) + \frac{1}{\lambda/\delta + h_{\text{cv}} + h_{\text{rad}}} \int_{2\pi} dp_\Omega(\omega) |\omega \cdot \mathbf{n}| \mathcal{L}(\mathbf{Y} \notin \text{PV}) \phi_{\text{sun}}(\mathbf{y}, t) \quad (\text{D.3})$$

In équation (D.3),  $\mathbf{n}$  is the normal vector at the interface oriented towards the outside,  $\mathbb{P}_{\text{cd} \rightarrow \text{cd}}$ ,  $\mathbb{P}_{\text{cd} \rightarrow \text{cv}}$  and  $\mathbb{P}_{\text{cd} \rightarrow \text{rad}}$  are probabilities depending on material properties and thermal exchange coefficients. The solar flux is taken into account when the sampled ray in the direction  $\omega$  does not intercept the photovoltaic field (for the roof surface) and when the sun is facing the building's surface for the other walls.

## FANN interpolation

Based on the results of the Monte Carlo calculations, we aim to develop a tool that can quickly and accurately estimate the required heating or cooling capacity according to meteorological parameters over a typical year. A feedforward artificial neural network (FANN) is used and trained on a database built from the probe point calculations of the stochastic algorithm described above. It is used as a nonlinear multivariate interpolator. It has been shown that this type of network is capable of modelling any continuous function over a given interval [67, 66]. The neural network, in the same way as a more classical correlation, is then comparable to a reduced model of the treated problem which depends on many independent parameters. The network is composed of elementary computing units (see figure D.3a) interconnected and organized in successive layers (see figure D.3b). In this work, we will use ANN with only one hidden layer.

The output of each neuron can be calculated as follows :

---

**Algorithm 1** : Monte Carlo algorithm to estimate the temperature  $T_w(\mathbf{x}, t)$ .

---

```
1 Sample a wall and a position  $\mathbf{x}_1$  to estimate  $T_w(\mathbf{x}_1, t_1)$ 
2  $\mathbf{x} = \mathbf{x}_1$ 
3  $t = t_1$ 
4  $w = 0$ 
5  $A = 0$  (accumulation term to account for solar irradiance)
6 while  $t > t_I$  do
7   Sample  $\tau$  according to exponential law according to the material properties  $\lambda, \rho, C_p$ 
8    $t = t - \tau$ 
9   if  $t < t_I$  then
10     $w = w + T_I(\mathbf{x}) + A$ 
11  else
12    if  $\mathbf{x} == \mathbf{x}_1$  or  $\mathbf{x} == \mathbf{y}_1$  then
13      Calculate  $\mathbb{P}_{cd \rightarrow cd}$  and  $\mathbb{P}_{cd \rightarrow cv}$  according to the continuity equation (équation (D.3))
14      Sample  $r$  according to uniform law between 0 and 1
15      if  $r < \mathbb{P}_{cd \rightarrow cd}$  then
16         $\mathbf{x}_1 = \mathbf{x}_1 + \delta \mathbf{n}$  (random walk continues into conduction in the solid)
17      else if  $r < \mathbb{P}_{cd \rightarrow cd} + \mathbb{P}_{cd \rightarrow cv}$  then
18        Evaluate  $A$  for solar irradiance contribution according to équation (D.3)
19         $w = w + T_{cv}(t) + A$ 
20      else
21        Evaluate  $A$  for solar irradiance contribution according to équation (D.3)
22         $w = w + T_{rad}(\mathbf{x}, t) + A$ 
23      end
24    else
25      Conduction in the solid domain using a random walk on  $\delta$ -sphere
26    end
27  end
28 end
```

---

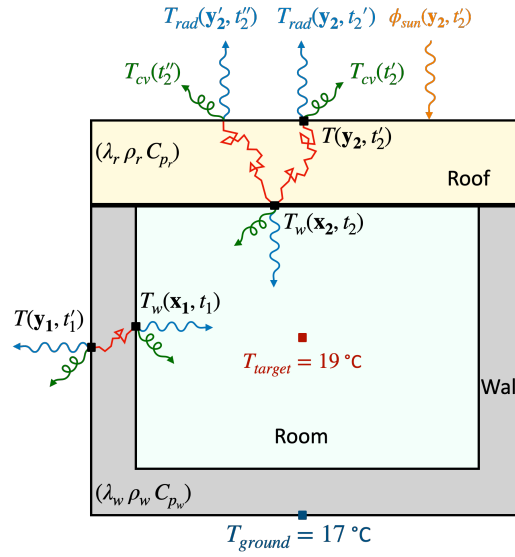


FIG. D.2 : Building geometry and conducto-convective-radiative transfers at interfaces

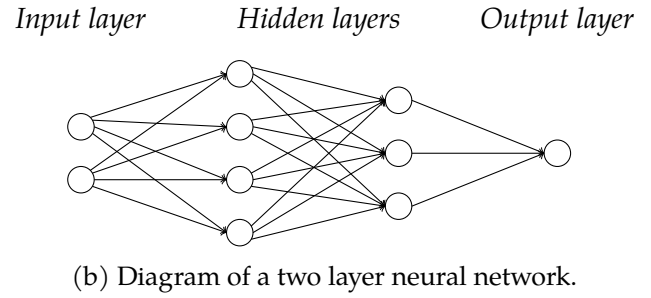
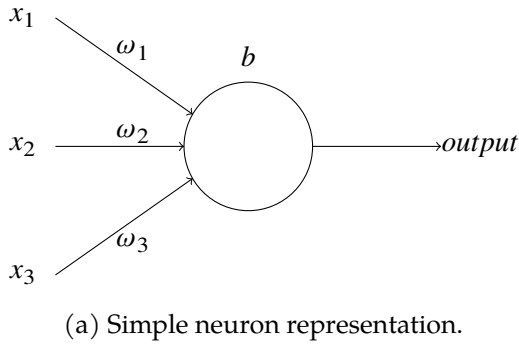


FIG. D.3 : Neuron representation and two layer neural network.

$$a = \sigma \left( \sum_j \omega_j x_j + b \right) = \sigma(z) \quad (\text{D.4})$$

Where  $\sigma$  is the activation function of the neuron,  $\omega_i$  and  $x_i$  are respectively the  $i^{\text{th}}$  weight and input and  $b$  the bias. For regression tasks, functions of the sigmoid family are classically used. In the present study, the hyperbolic tangent is implemented. The network used consists of only one hidden layer of 25 neurons.

To train the neural network, we use a Levenberg Marquardt algorithm [69] combined with Bayesian regularisation [90, 68]. This method is known to converge particularly quickly but requires the calculation of the Hessian matrix for the database used. Therefore, it is relatively computationally intensive compared to other methods and that limits the size of the database that can be used. Bayesian regularisation is used to *guide* the learning process and to penalise network weights  $\omega_j$  that are too high. If we do not use regularisation, the cost function, here the MSE, can be expressed as follows :

$$C = E_d = \frac{1}{2N} \sum_i [t(i) - a(i)]^2 \quad (\text{D.5})$$

Where,  $N$  is the number of training set,  $t(i)$  the  $i^{\text{th}}$  theoretical output (from the database) and  $a(i)$ , the  $i^{\text{th}}$  response of the ANN. With the bayesian regularisation, we have :

$$C = \alpha E_d + \beta E_w \quad (\text{D.6})$$

Where  $E_w$  is the sum of squares of the network weights. The hyperparameters  $\alpha$  and  $\beta$  are computed during the algorithm as follows :

$$\alpha = \frac{\gamma}{2E_w} \quad (\text{D.7})$$

$$\beta = \frac{N_{out} - \gamma}{2E_D} \quad (\text{D.8})$$

$$\gamma = n - 2\alpha \text{Tr}\{He^{-1}\} \quad (\text{D.9})$$

With  $n$  is the total number of parameters and  $N_{out}$ , the number of training output data. In these equations,  $\gamma$  is the effective number of parameters. This quantity is also calculated during the algorithm thanks to the Bayesian paradigm and gives us information on the number of parameters useful for modelling the target function. In other words, this coefficient informs us about the relevance of the topology used and implicitly, about the number of neurons on the hidden layer.

The value of  $\gamma$  is calculated according to the trace of the Hessian matrix  $\text{Tr}\{He^{-1}\}$  and is thus adapted to the database used. The  $\alpha$  and  $\beta$  coefficients are also modified during the training according to the database in an optimal way. These characteristics mean that this method has particularly powerful generalisation capabilities and is very little sensitive to overfitting.

Finally, in the theoretical development of Bayesian regularisation, it is assumed that the theoretical responses from the database are subject to Gaussian error. Although this assumption is not generally correct, the Monte Carlo algorithm used to generate the training data implies by construction this type of numerical error. This training method is therefore particularly suitable for the case under study. This methodology, combining a monte carlo algorithm and Bayesian regularisation, has already been successfully used to calculate the field thermal radiation in aeronautical combustion chambers [94]. We can summarise the whole algorithm used here by the following procedure :

The maximal number of epochs is set at 1000. The data set was normalised before training between -1 and 1 depending on the minimum and maximum of each parameter. To train the neural network, we use a code developed in C language and based on the algorithms described in Hagan et al. [69]. The training of the neural networks is conducted on a four cores standard desktop computer equipped with an Intel Core i5-8265U processor and 15.5 GB of RAM.

## Results

---

Weather data are extracted from the National Solar Radiation Database (NSRDB<sup>2</sup>) from the NREL laboratory and from the Photovoltaic Geographical Information System (PVGIS<sup>3</sup>). The global horizontal irradiance ( $\text{W m}^{-2}$ ), the wind speed ( $\text{m s}^{-1}$ ) and the air temperature ( $^{\circ}\text{C}$ ) are available at hourly time step. The solar elevation ( $^{\circ}$ ) and the solar azimuth ( $^{\circ}$ ) angles are calculated using the Python's tool pvlib<sup>4</sup>. The three main weather data are plotted for the cities of Las Vegas, Nancy and Cape Town in figure D.4.

---

<sup>2</sup><https://nsrdb.nrel.gov/>

<sup>3</sup>[https://re.jrc.ec.europa.eu/pvg\\_tools/en/](https://re.jrc.ec.europa.eu/pvg_tools/en/)

<sup>4</sup><https://pvlib-python.readthedocs.io/en/stable/>



---

**Algorithm 2 : FAAN**

---

```
1 Training data normalization
2 Random initialization of the parameters
3 while  $Nb_{epoch} < criterion$  do
4   foreach training set of the batch do
5     Feedforward pass :
6     Compute the individual network output, the associated error and the sensivity vector
7      $\delta^L$ 
8     on the last layer
9     Backward pass :
10    Retropropagation of the sensitivities  $\delta^l$  in each layer  $l$ 
11    Calculation of the Jacobien and Hessian matrices
12  end
13  foreach Parameter in each layer  $l$  do
14    Update the weights and the biases
15    Update the hyperparameters  $\alpha$ ,  $\beta$  and  $\gamma$ 
16  end
17 Repeat the procedure between points 3 and 4 for each epoch until a stopping criterion is
    reached
18 end
```

---



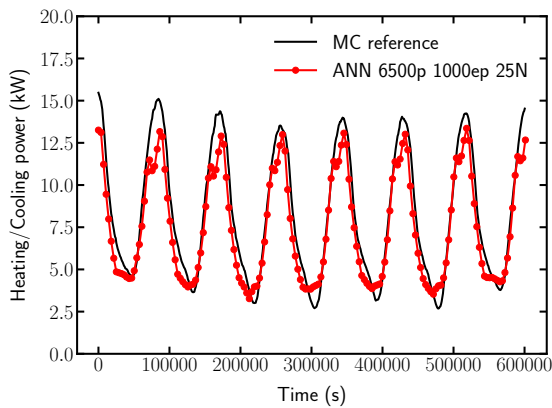
FIG. D.4 : Weather data (global horizontal irradiance, air temperature and wind speed) for 2019 for Las Vegas, Cape Town and Nancy. For the radiation data the daily maximum is displayed and for the air temperature the daily average is plotted.

City		Monte Carlo		FANN	
		Energy (MWh) $\pm\sigma$	Comp. time (s)	Energy (MWh)	Comp. time <sup>5</sup> (s)
Las Vegas	No PV	53.04 $\pm$ 0.23	10.192	50.43	74 $\cdot$ 10 <sup>-6</sup>
	PV	45.04 $\pm$ 0.19	10.260	43.76	74 $\cdot$ 10 <sup>-6</sup>
Cape Town	No PV	26.95 $\pm$ 0.22	9.717	25.75	65 $\cdot$ 10 <sup>-6</sup>
	PV	20.70 $\pm$ 0.20	9.599	19.61	63 $\cdot$ 10 <sup>-6</sup>

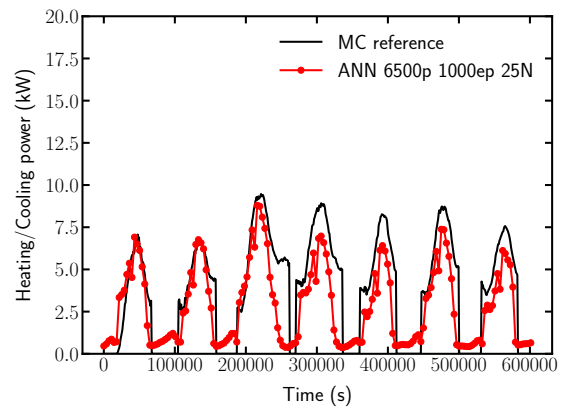
TABLE D.2 : Energy evaluation of the building's energy needs for the concrete configuration without photovoltaic field in 2019 and comparison between the Monte Carlo algorithm (10<sup>6</sup> realizations) and the FANN.

<sup>5</sup> Interpolation time for 8760 points over the whole year (every hour) and for numerical integration.

The total power for the building needs is estimated using the Monte Carlo algorithm for random times during the 2019 year. Then the Feedforward Artificial Neural Network is used to interpolate these points to simulate the power over a summer week for Las Vegas and Cape Town. The two selected weeks correspond to a summer and a winter season for both cities. The neural network was also tested with a training using only TMY data instead of 2019 data to estimate the power for the 2019 year. The obtained results demonstrate the robustness of the FANN to predict the power simulated with the Monte Carlo method and are shown in figure D.5 and figure D.6 for the concrete rooftop configuration (one layer only) with and without photovoltaic field for the summer week. The result for the green roof configuration is given in figure D.7 for Las Vegas during the summer week and demonstrates similar behaviour for this rooftop configuration.



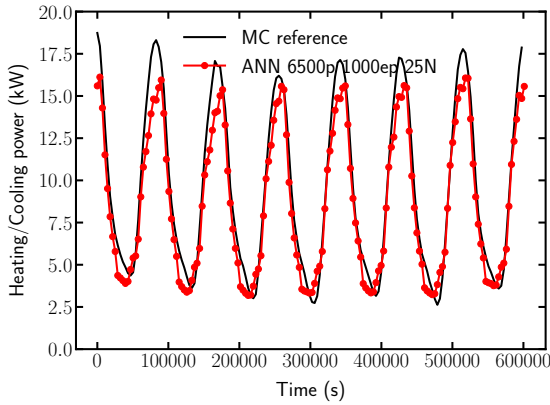
(a) Las Vegas, 2019 Summer week : Monte Carlo simulation in black color (reference solution), interpolation using TMY with the FANN in red color. Configuration with photovoltaic field and concrete roof.



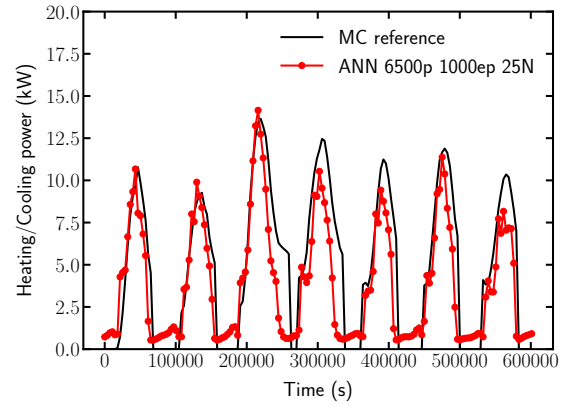
(b) Cape Town, 2019 Summer week : Monte Carlo simulation in black color (reference solution), interpolation using TMY with the FANN in red color. Configuration with photovoltaic field and concrete roof.

FIG. D.5 : Monte Carlo simulation points and FANN interpolation for Las Vegas and Cape Town in 2019 summer periods for the concrete roof configuration with photovoltaic field.

The database used to train the neural network consists of 6500 parameter sets from 5 different cities in both PV and non-PV configurations. In other words, on average, the neural network only *saw* less



(a) Las Vegas, 2019 Summer week : Monte Carlo simulation in black color (reference solution), interpolation using TMY with the FANN in red color. Configuration without photovoltaic field and concrete roof.



(b) Cape Town, 2019 Summer week : Monte Carlo simulation in black color (reference solution), interpolation using TMY with the FANN in red color. Configuration without photovoltaic field and concrete roof.

FIG. D.6 : Monte Carlo simulation points and FANN interpolation for Las Vegas and Cape Town in 2019 summer periods for the concrete roof configuration without photovoltaic field.

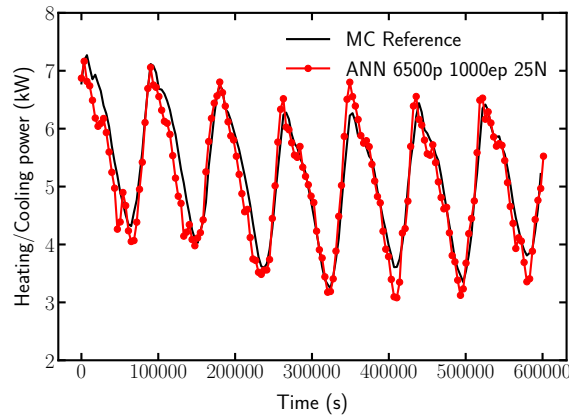


FIG. D.7 : Las Vegas, 2019 Summer week : Monte Carlo simulation in black color (reference solution), interpolation using TMY with the FANN in red color. Configuration with photovoltaic field, green roof layer.

than 2 points per day per configuration during training to be able to interpolate with as much accuracy the power needed for heating or cooling at each moment of the year 2019. In figures D.5 à D.7, the network was trained on an average year and therefore different from the one on which it interpolated its results. Under these conditions, the integration of the power calculated by the network over the whole year makes it possible to calculate an energy requirement that is very close, to within a few percent, to that obtained by the Monte Carlo method. This flexibility and this robustness demonstrate the great capabilities of this type of tool for the development of accurate and efficient scale models. Results for winter periods are not presented here but our simulations suggest that the presence of the photovoltaic field is advantageous in reducing the needs of the building. In the same way, in the summer period, it

seems that the air conditioning consumption is higher without the photovoltaic field.

## Conclusion

---

The results show that the energy required to heat and cool the building is lower for a rooftop PV configuration for Las Vegas and Cape Town in 2019 with a difference of 8 MWh for Las Vegas and 6 MWh for Cape Town. Moreover, the configuration with a green roof tends to favour this effect with an additional thermal insulation brought by the two additional layers (vegetation and substrate layers). Finally, the neural network interpolation, using the FANN model, proved to be particularly efficient with only 6500 training Monte Carlo points for all 5 cities studied and the configurations with and without photovoltaic panels. In a future work, it is envisaged to conduct a sensitivity study on the building parameters : number of layers, thermal properties etc.

## Acknowledgement

---

This research work was carried out within the scope of the PROOF Project, funded by the French National Research Agency (ANR) under the grant number no ANR-19-CE22-008.

## Références

---

- [66] K. HORNIK, M. STINCHCOMBE et H. WHITE. «Multilayer feedforward networks are universal approximators». en. In : *Neural Networks* 2.5 (1989), p. 359-366. DOI: [10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- [67] M. LESHNO et al. «Multilayer feedforward networks with a nonpolynomial activation function can approximate any function». en. In : *Neural Networks* 6.6 (jan. 1993), p. 861-867. DOI: [10.1016/S0893-6080\(05\)80131-5](https://doi.org/10.1016/S0893-6080(05)80131-5).
- [68] D. J. C. MACKAY. «Bayesian Interpolation». In : *Neural Computation* 4.3 (1992), p. 415-447.
- [69] M. T. HAGAN et al. *Neural Network Design*. 2<sup>e</sup> éd. Boston : PWS Publishing Co., 1996.
- [90] D. F. FORESEE et M. HAGAN. «Gauss-Newton approximation to Bayesian learning». In : *Proceedings of International Conference on Neural Networks (ICNN'97)*. International Conference on Neural Networks (ICNN'97). T. 3. Houston, TX, USA : IEEE, 1997, p. 1930-1935. DOI: [10.1109/ICNN.1997.614194](https://doi.org/10.1109/ICNN.1997.614194).
- [94] A. ROYER et al. «A new method for modeling radiative heat transfer based on Bayesian artificial neural networks and Monte Carlo method in participating media». In : *International Journal of Heat and Mass Transfer* 201 (2023), p. 123610.
- [108] N. VILLEFRANQUE et al. «The “teapot in a city” : A paradigm shift in urban climate modeling». In : *Science Advances* 8 (2022).
- [109] Irena. <https://irena.org/solar>.
- [110] K. BÓDIS et al. «A high-resolution geospatial assessment of the rooftop solar photovoltaic potential in the European Union». In : *Renewable and Sustainable Energy Reviews* 114 (2019), p. 109309. DOI: <https://doi.org/10.1016/j.rser.2019.109309>.
- [111] Z. LI et al. «A review of operational energy consumption calculation method for urban buildings». In : *Building Simulation* 13.4 (2020), p. 739-751. DOI: [10.1007/s12273-020-0619-0](https://doi.org/10.1007/s12273-020-0619-0).
- [112] A. GHOSH. «Potential of building integrated and attached/applied photovoltaic (BIPV/BAPV) for adaptive less energy-hungry building's skin : A comprehensive review». In : *Journal of Cleaner Production* 276 (2020), p. 123343. DOI: <https://doi.org/10.1016/j.jclepro.2020.123343>.

- 
- [113] Y. WANG, D. WANG et Y. LIU. «Study on Comprehensive Energy-saving of Shading and Photovoltaics of Roof Added PV Module». In : *Energy Procedia* 132 (2017). 11th Nordic Symposium on Building Physics, NSB2017, 11-14 June 2017, Trondheim, Norway, p. 598-603. DOI : <https://doi.org/10.1016/j.egypro.2017.09.672>.
- [114] Y. A. SHEIKH et al. «Impact of rooftop photovoltaic on energy demand of a building in a hot semi-arid climate». In : *Journal of Renewable and Sustainable Energy* 13.6 (2021), p. 065101.
- [115] H. ALASADI, J.-K. CHOI et R. B. MULFORD. «Influence of Photovoltaic Shading on Rooftop Heat Transfer, Building Energy Loads, and Photovoltaic Power Output». In : *Journal of Solar Energy Engineering* 144.6 (2022), p. 061011.
- [116] M. SANS et al. «Solving transient coupled conductive and radiative transfers in porous media with a Monte Carlo Method : Characterization of thermal conductivity of foams using a numerical Flash Method». In : *International Journal of Thermal Sciences* 179 (2022), p. 107656.
- [117] L. IBARRART, S. BLANCO, C. CALIOT et al. «Advection, diffusion and linear transport in a single path-sampling Monte-Carlo algorithm : getting insensitive to geometrical refinement». In : <https://hal.science/hal-03818899> (2022).
- [118] J.-M. TREGAN, J.-L. AMESTOY, M. BATI et al. «Coupling radiative, conductive and convective heat-transfers in a single Monte Carlo algorithm : a general theoretical framework for linear situations». In : <https://hal.science/hal-03819157> (2022).
- [119] S. MAIRE et G. NGUYEN. «Stochastic finite differences for elliptic diffusion equations in stratified domains». In : *Mathematics and Computers in Simulation* 121 (2016), p. 146-165.
- [120] M. BOSSY et al. «Monte Carlo methods for linear and non-linear Poisson-Boltzmann equation». In : *ESAIM : Proceedings and Surveys* 48 (2015), p. 420-446.
- [121] W. McADAMS. *Heat Transmission*. 1954.
- [122] O. FARGES et al. «Life-time integration using Monte Carlo Methods when optimizing the design of concentrated solar power plants». In : *Solar Energy* 113 (2015), p. 57-62.



# Bibliographie

- [1] L. PIERROT. « Développement, étude critique et validation de modèles de propriétés radiatives infrarouges de CO<sub>2</sub> et H<sub>2</sub>O à haute température. Applications au calcul des transferts dans des chambres aéronautiques et à la télédétection. » Thèse de doct. 1997.
- [2] P. J. COELHO, P. PEREZ et M. EL HAFI. « Benchmark numerical solutions for radiative heat transfer in two-dimensional axisymmetric enclosures with nongray sooting media ». In : *Numerical Heat Transfer, Part B : Fundamentals* 43.5 (2003), p. 425-444.
- [3] H. I. H. SARAVANAMUTTOO et al. *Gas Turbine Theory*. Seventh Edition. Pearson education, 2017.
- [4] H. C. HOTTEL. « McAdams, Heat Transmission ». In : sous la dir. de MCGRAW-HILL. 3<sup>e</sup> éd. 1954. Chap. Radiant-Heat Transmission.
- [5] R. SIEGEL et J. R. HOWELL. *Thermal Radiation Heat Transfer*. 5<sup>e</sup> éd. CRC Press, 2010. 982 p.
- [6] M. F. MODEST. *Radiative heat transfer*. Sous la dir. d'A. PRESS. 2<sup>th</sup>. 2003.
- [7] M. GALTIER et al. « Integral formulation of null-collision Monte Carlo algorithms ». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 125 (2013), p. 57-68. DOI : [10.1016/j.jqsrt.2013.04.001](https://doi.org/10.1016/j.jqsrt.2013.04.001).
- [8] M. EL HAFI et al. « Three viewpoints on null-collision Monte Carlo algorithms ». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 260 (2021).
- [9] 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 ». PhD Thesis. Ecole Nationale Supérieure des Mines d'Albi-Carmaux, Université de Toulouse, 2020.
- [66] K. HORNIK, M. STINCHCOMBE et H. WHITE. « Multilayer feedforward networks are universal approximators ». en. In : *Neural Networks* 2.5 (1989), p. 359-366. DOI : [10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- [67] M. LESHNO et al. « Multilayer feedforward networks with a nonpolynomial activation function can approximate any function ». en. In : *Neural Networks* 6.6 (jan. 1993), p. 861-867. DOI : [10.1016/S0893-6080\(05\)80131-5](https://doi.org/10.1016/S0893-6080(05)80131-5).
- [68] D. J. C. MACKAY. « Bayesian Interpolation ». In : *Neural Computation* 4.3 (1992), p. 415-447.
- [69] M. T. HAGAN et al. *Neural Network Design*. 2<sup>e</sup> éd. Boston : PWS Publishing Co., 1996.
- [10] N. VILLEFRANQUE et al. « A path tracing Monte Carlo library for 3-D radiative transfer in highly resolved cloudy atmospheres ». In : *Journal of Advances in Modeling Earth systems* 11 (2019).
- [11] M. ROGER. « Modèles de sensibilité dans le cadre de la méthode de Monte-Carlo : illustrations en transfert radiatif ». Thèse de doct. Institut National Polytechnique de Toulouse, 2006.
- [12] T. F. SMITH, Z. F. SHEN et J. N. FRIEDMAN. « Evaluation of Coefficients for the Weighted Sum of Gray Gases model ». In : *ASME JHT* 104 (1982), p. 602-608.
- [13] A. COPPALLE et P. VERVISCH. « The total emissivities of high-temperature flames ». In : *Combustion and Flame* 49.1 (1983), p. 101-108. DOI : [10.1016/0010-2180\(83\)90154-2](https://doi.org/10.1016/0010-2180(83)90154-2).
- [14] M. F. MODEST. « The Weighted-Sum-of-Gray-Gases Model for Arbitrary Solution Methods in Radiative Transfer ». In : *Journal of Heat Transfer* 113.3 (1991), p. 650-656. DOI : [10.1115/1.2910614](https://doi.org/10.1115/1.2910614).

- [15] M. K. DENISON. «A spectral line-based weighted-sum-of-gray-gases model for arbitrary RTE solvers». Thèse de doct. Brigham Young University, 1994.
- [16] M. K. DENISON et B. W. WEBB. «k-distributions and Weighted-Sum-of-Gray-Gases-A hybrid model». In : *International Heat Transfer Conference Digital Library*. Begel House Inc. 1994.
- [17] «The spectral-line weighted-sum-of-gray-gases model for H<sub>2</sub>O/CO<sub>2</sub> mixtures». In : *Journal of heat transfer* 117.3 (1995), p. 788-792.
- [18] M. DENISON et B. WEBB. «An absorption-line blackbody distribution function for efficient calculation of total gas radiative transfer». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 50.5 (1993), p. 499-510. DOI : [10.1016/0022-4073\(93\)90043-H](https://doi.org/10.1016/0022-4073(93)90043-H).
- [19] M. K. DENISON et B. W. WEBB. «Development and application of an absorption- line blackbody distribution function for CO<sub>2</sub>». In : 38.10 (1995).
- [20] J. T. PEARSON. «The Development of Updated and Improved SLW Model Parameters and Its Application to Comprehensive Combustion Predictions». Thèse de doct. Brigham Young University, 2013.
- [21] F. ANDRÉ, V. P. SOLOVJOV et B. W. WEBB. «The  $\omega$ -absorption line distribution function for rank correlated SLW model prediction of radiative transfer in non-uniform gases». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 280 (2022). DOI : [10.1016/j.jqsrt.2022.108081](https://doi.org/10.1016/j.jqsrt.2022.108081).
- [22] M. GALTIER et al. «Assessment of narrow-band and full spectrum gas radiation methods in a real industrial glass furnace configuration». In : *Applied Thermal Engineering* 216 (2022), p. 119020. DOI : [10.1016/j.applthermaleng.2022.119020](https://doi.org/10.1016/j.applthermaleng.2022.119020).
- [23] J. M. HAMMERSLEY et D. C. HANDSCOMB. *Monte Carlo Methods*. London : Chapman et Hall, 1965.
- [24] J. DELATORRE et al. «Monte Carlo advances and concentrated solar applications». In : *Solar Energy* 103 (2014), p. 653-681. DOI : [10.1016/j.solener.2013.02.035](https://doi.org/10.1016/j.solener.2013.02.035).
- [25] M. GALTIER. «Approche statistique du rayonnement dans les milieux gazeux hétérogènes : de l'échantillonnage des transitions moléculaires au calcul de grandeurs radiatives». Thèse de doct. 2014.
- [26] W. L. DUNN et J. K. SHULTIS. *Exploring monte carlo methods*. Elsevier, 2022.
- [27] M. GALTIER et al. «Integral formulation of null-collision Monte Carlo algorithms». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 125 (2013), p. 57-68.
- [28] H. R. SKULLERUD. «The stochastic computer simulation of ion motion in a gas subjected to a constant electric field». In : *Journal of Physics D : Applied Physics* 1.11 (1968), p. 1567. DOI : [10.1088/0022-3727/1/11/423](https://doi.org/10.1088/0022-3727/1/11/423).
- [29] E. WOODCOCK et al. «Techniques used in the GEM code for Monte Carlo neutronics calculations in reactors and other systems of complex geometry». In : *Proc. Conf. Applications of Computing Methods to Reactor Problems*. T. 557. 2. Argonne National Laboratory. 1965.
- [30] E. BRUN et al. «TRIPOLI-4®, CEA, EDF and AREVA reference Monte Carlo code». In : *Annals of Nuclear Energy* 82 (2015), p. 151-160. DOI : [10.1016/j.anucene.2014.07.053](https://doi.org/10.1016/j.anucene.2014.07.053).
- [31] M. KERKER. *The scattering of light and other electromagnetic radiation*. Elsevier, 1969.
- [32] P. BOULET et al. «PROPRIETES RADIATIVES DES MILIEUX DIFFUSANTS». École thématique CNRS «Rayonnement thermique en milieux semi-transparents», 2014.
- [33] A. LIEMERT et A. KIENLE. «Radiative transfer in two-dimensional infinitely extended scattering media». In : *Journal of Physics A : Mathematical and Theoretical* 44.50 (2011). DOI : [10.1088/1751-8113/44/50/505206](https://doi.org/10.1088/1751-8113/44/50/505206).
- [34] N. JACQUINET-HUSSON et al. «The 2009 edition of the GEISA spectroscopic database». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 112.15 (oct. 2011), p. 2395-2445. DOI : [10.1016/j.jqsrt.2011.06.004](https://doi.org/10.1016/j.jqsrt.2011.06.004).



- [35] L. ROTHMAN et al. «HITEMP, the high-temperature molecular spectroscopic database». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 111.15 (oct. 2010), p. 2139-2150. DOI : [10.1016/j.jqsrt.2010.05.001](https://doi.org/10.1016/j.jqsrt.2010.05.001).
- [36] L. ROTHMAN et al. «The HITRAN2012 molecular spectroscopic database». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 130 (nov. 2013), p. 4-50. DOI : [10.1016/j.jqsrt.2013.07.002](https://doi.org/10.1016/j.jqsrt.2013.07.002).
- [37] S. TASHKUN et V. PEREVALOV. «CDS-4000 : High-resolution, high-temperature carbon dioxide spectroscopic databank». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 112.9 (juin 2011), p. 1403-1410. DOI : [10.1016/j.jqsrt.2011.03.005](https://doi.org/10.1016/j.jqsrt.2011.03.005).
- [38] S. TASHKUN et al. «CDS-296, high-resolution carbon dioxide spectroscopic databank : An update». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 228 (mai 2019), p. 124-131. DOI : [10.1016/j.jqsrt.2019.03.001](https://doi.org/10.1016/j.jqsrt.2019.03.001).
- [39] L. SPARKS. «Efficient line-by-line calculation of absorption coefficients to high numerical accuracy». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 57.5 (mai 1997), p. 631-650. DOI : [10.1016/S0022-4073\(96\)00154-9](https://doi.org/10.1016/S0022-4073(96)00154-9).
- [40] J. TAINE. «A line-by-line calculation of low-resolution radiative properties of CO<sub>2</sub>-CO-Transparent nonisothermal gases mixtures up to 3000K». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 30.4 (1983), p. 371-379.
- [41] Y. NYFFENEGGER-PERE et al. «Une estimation exacte du forçage radiatif par la méthode de Monte-Carlo». In : *SFT 2022-30<sup>ème</sup> Congrès Français de Thermique 2022*. 2022.
- [42] W. MALKMUS. «Random Lorentz Band Model with Exponential-Tailed S<sup>-1</sup> Line-Intensity Distribution Function\*». In : *Journal of the Optical Society of America* 57.3 (1<sup>er</sup> mars 1967), p. 323. DOI : [10.1364/JOSA.57.000323](https://doi.org/10.1364/JOSA.57.000323).
- [43] P. RIVIÈRE et A. SOUFIANI. «Updated band model parameters for H<sub>2</sub>O, CO<sub>2</sub>, CH<sub>4</sub> and CO radiation at high temperature». In : *International Journal of Heat and Mass Transfer* 55.13 (2012), p. 3349-3358. DOI : [10.1016/j.ijheatmasstransfer.2012.03.019](https://doi.org/10.1016/j.ijheatmasstransfer.2012.03.019).
- [44] A. SOUFIANI, J. M. HARTMANN et J. TAINE. «Validity of band-model calculations for CO<sub>2</sub> and H<sub>2</sub>O applied to radiative properties and conductive-radiative transfer». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 33.3 (1985), p. 243-257.
- [45] S. J. YOUNG. «Non-isothermal band model theory». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 18 (1977), p. 1-28.
- [46] R. M. GOODY et Y. L. YUNG. *Atmospheric radiation : theoretical basis*. Oxford university press, 1995.
- [47] R. GOODY et al. «The correlated-k method for radiation calculations in nonhomogeneous atmospheres». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 42.6 (1989), p. 539-550.
- [48] J. TAINE et A. SOUFIANI. «High temperature gas radiative property parameters of statistical narrow-band model for H<sub>2</sub>O, CO<sub>2</sub> and CO, and correlated-K model for H<sub>2</sub>O and CO<sub>2</sub>». In : 40 (1997), p. 987-991.
- [49] L. PIERROT, A. SOUFIANI et J. TAINE. «Accuracy of narrow-band and global models for radiative transfer in H<sub>2</sub>O, CO<sub>2</sub>, and H<sub>2</sub>O-CO<sub>2</sub> mixtures at high temperature». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 62 (1999), p. 523-548. DOI : [10.1016/S0022-4073\(98\)00125-3](https://doi.org/10.1016/S0022-4073(98)00125-3).
- [50] F. LIU et G. J. SMALLWOOD. «Application of the statistical narrow-band correlated-k method to low-resolution spectral intensity and radiative heat transfer calculations Effects of the quadrature scheme». In : *Int. J. Heat Mass Transfer* 43 (2000).
- [51] F. ANDRE et al. *Rayonnement des gaz : des spectres de raies aux modèles approchés*. Four Solaire d'Odeillo : Ecole thématique CNRS, 2014.
- [52] C. B. LUDWIG et al. *Handbook of infrared radiation from combustion gases*. National Aeronautics and space administration (NASA). 1973. 496 p.

- [53] A. A. LACIS et V. OINAS. «A description of the correlated-k distribution method for modeling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres». In : *J. Geophysical Research* 96 (1991).
- [54] V. GOUTIERE, F. LIU et A. CHARETTE. «An assessment of real-gas modelling in 2D enclosures». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 64.3 (2000), p. 299-326. DOI : [10.1016/S0022-4073\(99\)00102-8](https://doi.org/10.1016/S0022-4073(99)00102-8).
- [55] S. A. CLOUGH, F. X. KNEIZYS et R. W. DAVIES. «Line Shape and the Water Vapor Continuum». In : *Atmospheric Research* 23 (1989), p. 229-241.
- [56] M. Y. PERRIN et J. M. HARTMANN. «Temperature-dependent measurements and modeling of absorption by CO<sub>2</sub>-N<sub>2</sub> mixtures in the far line\_wings of the 4.3 μm CO<sub>2</sub> band». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 42.4 (1989), p. 311-317.
- [57] J. M. HARTMANN et al. «The infrared continuum of pure water vapor : calculations and high-temperature measurements». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 49.6 (1993), p. 675-691.
- [58] C. BRODBECK et al. «Absorption of radiation by gases from low to high pressures. II. Measurements and calculations of CO infrared spectra». In : *Journal de Physique II* 4.12 (1994), p. 2101-2118. DOI : [10.1051/jp2:1994249](https://doi.org/10.1051/jp2:1994249).
- [59] J. M. HARTMANN et al. «Absorption of radiation by gases from low to high pressures. I. Empirical line-by-line and narrow-band statistical models». In : *Journal de Physique II* 1.7 (1991), p. 739-762. DOI : [10.1051/jp2:1991107](https://doi.org/10.1051/jp2:1991107).
- [60] M. DENISON et B. WEBB. «The absorption-line blackbody distribution function at elevated pressures». In : *1st International Symposium on Radiative Heat Transfer, Kusadasi, Turkey*. 1995.
- [61] F. ANDRÉ et R. VAILLON. «The spectral-line moment-based (SLMB) modeling of the wide band and global blackbody-weighted transmission function and cumulative distribution function of the absorption coefficient in uniform gaseous media». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 109.14 (2008), p. 2401-2416. DOI : [10.1016/j.jqsrt.2008.05.002](https://doi.org/10.1016/j.jqsrt.2008.05.002).
- [62] P. S. CUMBER, M. FAIRWEATHER et H. S. LEDIN. «Application of wide band radiation models to non-homogeneous combustion systems». In : *Int. J. Heat Mass Transfer* 41.11 (1998), p. 1573-1584.
- [63] D. EDWARDS et A. BALAKRISHNAN. «Thermal radiation by combustion gases». In : *International Journal of Heat and Mass Transfer* 16.1 (1973), p. 25-40. DOI : [10.1016/0017-9310\(73\)90248-2](https://doi.org/10.1016/0017-9310(73)90248-2).
- [64] V. P. SOLOVJOV et B. W. WEBB. «An Efficient Method for Modeling Radiative Transfer in Multi-component Gas Mixtures With Soot». In : *Journal of Heat Transfer* 123.3 (2001), p. 450-457.
- [65] C. A. HOERLLE, F. H. R. FRANÇA et P. R. PAGOT. «Effects of radiation modeling on non-premixed sooting flames simulations under oxyfuel conditions». In : *Combustion and Flame* 217 (2020), p. 294-305.
- [70] W. S. McCULLOCH et W. PITTS. «A logical calculus of the ideas immanent in nervous activity». en. In : *Bulletin of Mathematical Biophysics* 5 (1943), p. 19.
- [71] A. C. NEVES et al. «A New Approach to Damage Detection in Bridges Using Machine Learning». en. In : *Experimental Vibration Analysis for Civil Structures*. Sous la dir. de J. P. CONTE et al. T. 5. Series Title : Lecture Notes in Civil Engineering. San Diego, USA : Springer International Publishing, 2018, p. 73-84. DOI : [10.1007/978-3-319-67443-8\\_5](https://doi.org/10.1007/978-3-319-67443-8_5).
- [72] F. ROSENBLATT. «The perceptron : A probabilistic model for information storage and organization in the brain.» en. In : *Psychological Review* 65.6 (1958), p. 386-408. DOI : [10.1037/h0042519](https://doi.org/10.1037/h0042519).
- [73] Z. LI et al. «A Survey of Convolutional Neural Networks : Analysis, Applications, and Prospects». en. In : *IEEE Transactions on Neural Networks and Learning Systems* 33.12 (déc. 2022), p. 6999-7019. DOI : [10.1109/TNNLS.2021.3084827](https://doi.org/10.1109/TNNLS.2021.3084827).
- [74] Z. C. LIPTON, J. BERKOWITZ et C. ELKAN. «A critical review of recurrent neural networks for sequence learning». arXiv preprint arXiv :1506.00019. 2015.

- [75] A. APICELLA et al. «A survey on modern trainable activation functions». In : *Neural Networks* 138 (2021), p. 14-32. DOI : [10.1016/j.neunet.2021.01.026](https://doi.org/10.1016/j.neunet.2021.01.026). arXiv : 2005.00817[cs, stat].
- [76] D. E. RUMELHART, G. E. HINTON et R. J. WILLIAMS. «Learning representation by back-propagating errors». In : *Nature* 323.9 (1986), p. 533-536.
- [77] M. NIELSEN. *Neural Networks and Deep Learning*. 2015.
- [78] S. K. KUMAR. «On weight initialization in deep neural networks». 2017.
- [79] L. LU et al. «Dying relu and initialization : Theory and numerical examples». 2019.
- [80] S. LAWRENCE, C. L. GILES et A. C. TSOI. «Lessons in neural network training : Overfitting may be harder than expected». In : *AAAI/IAAI*. 1997, p. 540-545.
- [81] Y. TIAN et Y. ZHANG. «A comprehensive survey on regularization strategies in machine learning». In : *Information Fusion* 80 (2022), p. 146-166. DOI : <https://doi.org/10.1016/j.inffus.2021.11.005>.
- [82] L. PRECHELT. «Early Stopping | but when?» In : *Neural Networks : Tricks of the Trade*. Second Edition. 2012, p. 53-67.
- [83] D. WHITLEY et al. «Genetic algorithms and neural networks». In : *Genetic algorithms in engineering and computer science* 3 (1995), p. 191-201.
- [84] L. BOTTOU. «Stochastic Gradient Descent Tricks». In : *Neural Networks : Tricks of the Trade*. Sous la dir. de G. MONTAVON, G. B. ORR et K.-R. MÜLLER. T. 7700. Series Title : Lecture Notes in Computer Science. Berlin, Heidelberg : Springer Berlin Heidelberg, 2012, p. 421-436. DOI : [10.1007/978-3-642-35289-8\\_25](https://doi.org/10.1007/978-3-642-35289-8_25).
- [85] T. TIELEMAN, G. HINTON et al. «Lecture 6.5-rmsprop : Divide the gradient by a running average of its recent magnitude». In : *COURSERA : Neural networks for machine learning* 4.2 (2012), p. 26-31.
- [86] D. P. KINGMA et J. BA. «Adam : A method for stochastic optimization». In : *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*. 2014.
- [87] D. W. MARQUARDT. «An algorithm for least-squares estimation of nonlinear parameters». In : *Journal of the society for Industrial and Applied Mathematics* 11.2 (1963), p. 431-441.
- [88] J. J. MORÉ. «The Levenberg-Marquardt algorithm : Implementation and theory». In : *Numerical Analysis*. Sous la dir. de G. A. WATSON. T. 630. Series Title : Lecture Notes in Mathematics. Berlin, Heidelberg : Springer Berlin Heidelberg, 1978, p. 105-116. DOI : [10.1007/BFb0067700](https://doi.org/10.1007/BFb0067700).
- [89] S. GRATTON, A. S. LAWLESS et N. K. NICHOLS. «Approximate Gauss–Newton Methods for Nonlinear Least Squares Problems». In : *SIAM Journal on Optimization* 18.1 (2007), p. 106-132. DOI : [10.1137/050624935](https://doi.org/10.1137/050624935).
- [90] D. F. FORESEE et M. HAGAN. «Gauss-Newton approximation to Bayesian learning». In : *Proceedings of International Conference on Neural Networks (ICNN'97)*. International Conference on Neural Networks (ICNN'97). T. 3. Houston, TX, USA : IEEE, 1997, p. 1930-1935. DOI : [10.1109/ICNN.1997.614194](https://doi.org/10.1109/ICNN.1997.614194).
- [91] M. F. MODEST. «Narrow-band and full-spectrum k-distributions for radiative heat transfer—correlated-k vs. scaling approximation». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 76.1 (2003), p. 69-83.
- [92] J. R. HOWELL et K. J. DAUN. «The Past and Future of the Monte Carlo Method in Thermal Radiation Transfer». In : *Journal of Heat Transfer* 143.10 (2021).
- [93] V. EYMET et al. «Null-collision meshless Monte-Carlo - Application to the validation of fast radiative transfer solvers embedded in combustion simulators». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 129 (2013), p. 145-157.
- [94] A. ROYER et al. «A new method for modeling radiative heat transfer based on Bayesian artificial neural networks and Monte Carlo method in participating media». In : *International Journal of Heat and Mass Transfer* 201 (2023), p. 123610.

- [95] L. PENAZZI et al. «Toward the use of Symbolic Monte Carlo for Conduction-Radiation Coupling in Complex Geometries». In : *RAD-19 - 9th International Symposium on Radiative Transfer*. Athens, Greece : Begellhouse, 2019, 8 p. DOI : [10.1615/RAD-19.380](https://doi.org/10.1615/RAD-19.380).
- [96] M. GALTIER et al. «A symbolic approach for the identification of radiative properties». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 196 (2017), p. 130-141. DOI : [10.1016/j.jqsrt.2017.03.026](https://doi.org/10.1016/j.jqsrt.2017.03.026).
- [97] L. KOCIS et W. J. WHITEN. «Computational investigations of low-discrepancy sequences». In : *ACM Transactions on Mathematical Software* 23.2 (1997), p. 266-294. DOI : [10.1145/264029.264064](https://doi.org/10.1145/264029.264064).
- [98] E. THIÉMARD. «Sur le calcul et la majoration de la discrédance à l'origine». Thèse de doct. Ecole Polytechnique Fédérale de Lausanne, 2000.
- [99] L. PALLUOTTO et al. «Assessment of randomized Quasi-Monte Carlo method efficiency in radiative heat transfer simulations». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 236 (2019). DOI : [10.1016/j.jqsrt.2019.07.013](https://doi.org/10.1016/j.jqsrt.2019.07.013).
- [100] J. H. HALTON. «On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals». In : *Numerische Mathematik* 2 (1960), p. 84-90. DOI : [10.1007/BF01386213](https://doi.org/10.1007/BF01386213).
- [101] J. TAINE et A. SOUFIANI. «Gas IR Radiative Properties : From Spectroscopic Data to Approximate Models». In : *Advances in heat transfer* 33 (1999), p. 295-414.
- [102] J. P. RIGOL, C. H. JARVIS et N. STUART. «Artificial neural networks as a tool for spatial interpolation». In : *International Journal of Geographical Information Science* 15.4 (2001), p. 323-343.
- [103] V. NEVTIPILOVA. «Testing Artificial Neural Network (ANN) for Spatial Interpolation». In : *Journal of Geology & Geosciences* 03.2 (2014). DOI : [10.4172/2329-6755.1000145](https://doi.org/10.4172/2329-6755.1000145).
- [104] N. S.-N. LAM. «Spatial Interpolation Methods : A Review». In : *The American Cartographer* 10.2 (1983), p. 129-150.
- [105] L. DINIS, R. NATAL JORGE et J. BELINHA. «Analysis of 3D solids using the natural neighbour radial point interpolation method». In : *Computer Methods in Applied Mechanics and Engineering* 196.13 (2007), p. 2009-2028.
- [106] N. PLAZIAC. «Image interpolation using neural networks». In : *IEEE Transactions on Image Processing* 8.11 (1999), p. 1647-1651.
- [107] J. GO, K. SOHN et C. LEE. «Interpolation using neural networks for digital still cameras». In : *IEEE Transaction on Consumer Electronics* 46.3 (2000), p. 610-616.
- [108] N. VILLEFRANQUE et al. «The "teapot in a city" : A paradigm shift in urban climate modeling». In : *Science Advances* 8 (2022).
- [109] Irena. <https://irena.org/solar>.
- [110] K. BÓDIS et al. «A high-resolution geospatial assessment of the rooftop solar photovoltaic potential in the European Union». In : *Renewable and Sustainable Energy Reviews* 114 (2019), p. 109309. DOI : <https://doi.org/10.1016/j.rser.2019.109309>.
- [111] Z. LI et al. «A review of operational energy consumption calculation method for urban buildings». In : *Building Simulation* 13.4 (2020), p. 739-751. DOI : [10.1007/s12273-020-0619-0](https://doi.org/10.1007/s12273-020-0619-0).
- [112] A. GHOSH. «Potential of building integrated and attached/applied photovoltaic (BIPV/BAPV) for adaptive less energy-hungry building's skin : A comprehensive review». In : *Journal of Cleaner Production* 276 (2020), p. 123343. DOI : <https://doi.org/10.1016/j.jclepro.2020.123343>.
- [113] Y. WANG, D. WANG et Y. LIU. «Study on Comprehensive Energy-saving of Shading and Photovoltaics of Roof Added PV Module». In : *Energy Procedia* 132 (2017). 11th Nordic Symposium on Building Physics, NSB2017, 11-14 June 2017, Trondheim, Norway, p. 598-603. DOI : <https://doi.org/10.1016/j.egypro.2017.09.672>.

- 
- [114] Y. A. SHEIKH et al. «Impact of rooftop photovoltaic on energy demand of a building in a hot semi-arid climate». In : *Journal of Renewable and Sustainable Energy* 13.6 (2021), p. 065101.
- [115] H. ALASADI, J.-K. CHOI et R. B. MULFORD. «Influence of Photovoltaic Shading on Rooftop Heat Transfer, Building Energy Loads, and Photovoltaic Power Output». In : *Journal of Solar Energy Engineering* 144.6 (2022), p. 061011.
- [116] M. SANS et al. «Solving transient coupled conductive and radiative transfers in porous media with a Monte Carlo Method : Characterization of thermal conductivity of foams using a numerical Flash Method». In : *International Journal of Thermal Sciences* 179 (2022), p. 107656.
- [117] L. IBARRART, S. BLANCO, C. CALIOT et al. «Advection, diffusion and linear transport in a single path-sampling Monte-Carlo algorithm : getting insensitive to geometrical refinement». In : <https://hal.science/hal-03818899> (2022).
- [118] J.-M. TREGAN, J.-L. AMESTOY, M. BATI et al. «Coupling radiative, conductive and convective heat-transfers in a single Monte Carlo algorithm : a general theoretical framework for linear situations». In : <https://hal.science/hal-03819157> (2022).
- [119] S. MAIRE et G. NGUYEN. «Stochastic finite differences for elliptic diffusion equations in stratified domains». In : *Mathematics and Computers in Simulation* 121 (2016), p. 146-165.
- [120] M. BOSSY et al. «Monte Carlo methods for linear and non-linear Poisson-Boltzmann equation». In : *ESAIM : Proceedings and Surveys* 48 (2015), p. 420-446.
- [121] W. McADAMS. *Heat Transmission*. 1954.
- [122] O. FARGES et al. «Life-time integration using Monte Carlo Methods when optimizing the design of concentrated solar power plants». In : *Solar Energy* 113 (2015), p. 57-62.





---

## Résumé

---

Les chambres de combustion aéronautiques sont des environnements contraints fonctionnant à des températures et des pressions en constante augmentation depuis les dernières décennies. Dans ces conditions, les transferts thermiques par rayonnement sont conséquents et leurs influences sur la structure de flamme et sur les températures de parois est sensible. Il est donc essentiel d'être capable de modéliser avec précision ce mode de transfert sur des cas industriels très hétérogènes et sur des maillages de plusieurs centaines de millions de cellules. Les méthodes de résolution courantes ne permettent pas d'atteindre une précision satisfaisante compte tenu des besoins croissants. Des méthodes de résolution fines et l'utilisation de modèles de gaz spectraux seraient nécessaires mais impliqueraient une consommation de ressources de calcul prohibitives.

Pour palier ce problème, nous proposons dans ces travaux une nouvelle méthodologie de modélisation du rayonnement basée sur les méthodes de Monte Carlo et les réseaux de neurones bayésiens. Le principe du concept détaillé dans ce manuscrit commence par l'échantillonnage de points de référence tirés aléatoirement dans le milieu d'étude. Un calcul de la grandeur radiative d'intérêt est réalisée en chacun de ces points par méthode de Monte Carlo avec un modèle spectral fin. Les résultats de ces calculs sont ensuite stockés dans une base de données sur laquelle un réseau de neurones est entraîné selon un algorithme de Levenberg-Marquardt. Le paradigme bayésien nous permet de définir la topologie optimale du réseau et d'adapter automatiquement l'ensemble des paramètres inhérents au réseau sans intervention de l'utilisateur. Le réseau entraîné permet de reconstituer les champs de flux ou de divergence de flux au sein du domaine. Après une description théorique des notions mises en oeuvre dans ces travaux, des résultats sur cas académiques sont présentés dans différentes configurations. Nous démontrons alors la capacité de la méthode développée à modéliser le rayonnement en milieu participatif avec une précision très élevée et des coûts de calcul faibles.

**Mots clés :** *Rayonnement, Réseaux de neurones, transfert thermique, méthode de Monte Carlo*

---

## Abstract

---

Aeronautical combustion chambers are constrained environments operating at temperatures and pressures that have been steadily increasing over the last decades. Under these conditions, radiative heat transfer is significant and its influence on flame structure and wall temperatures is noticeable. It is therefore essential to be able to accurately model this transfer mode on very heterogeneous industrial cases and on mesh of several hundred million cells. Common resolution methods do not provide satisfactory accuracy in view of the growing needs. A fine resolution and the use of spectral gas models would be necessary but would imply a prohibitive consumption of computing resources.

To overcome this problem, we propose in this work a new radiation modelling methodology based on Monte Carlo methods and Bayesian neural networks. The principle of the concept detailed in this manuscript starts with the sampling of reference points randomly drawn in the study domain. The radiative quantity of interest is computed at each of these points using a Monte Carlo method and a fine spectral model. The results of these calculations are then stored in a database on which a neural network is trained using a Levenberg-Marquardt algorithm. The Bayesian paradigm allows us to define the optimal topology of the network and to automatically adapt the set of parameters inherent to the network without user intervention. The trained network allows to reconstitute the flux or flux divergence fields within the domain. After a theoretical description of the notions used in this work, results on academic cases are presented in different configurations. We then demonstrate the ability of the developed method to model radiation in a participatory environment with very high accuracy and low computational costs.

**Keywords :** *Radiative heat transfer, artificial neural network, Monte Carlo method*