



Région  
Hauts-de-France

UNIVERSITÉ DE LILLE

# THÈSE

pour obtenir le grade de :

DOCTEUR DE L'UNIVERSITÉ DE LILLE

dans la spécialité

“INFORMATIQUE”

par

Matthieu DABROWSKI



## Méthodes d'apprentissage profond inspirées par la physique pour l'inférence de la qualité de l'air

Thèse soutenue le 9/12/2024 devant le jury composé de :

JEAN-FRANÇOIS LÉON	Directeur de Recherche CNRS, HDR Laboratoire d'Aérodynamique, Université de Toulouse	(Rapporteur)
LAURENT PERRINET	Directeur de Recherche, HDR INT (CNRS), Université d'Aix-Marseille	(Rapporteur)
RENAUD PÉTERI	Maître de Conférences, HDR Laboratoire MIA, Université de la Rochelle	(Examineur)
VÉRONIQUE RIFFAULT	Professeure, HDR CERI Énergie et Environnement, IMT Nord-Europe	(Présidente du Jury)
RIM SLAMA SALMI	Maître de Conférences LICIT-ECO7, ENTPE	(Examinatrice)
JOSÉ MENNESSON	Maître de Conférences CERI Systèmes Numériques, IMT Nord-Europe	(Co-encadrant)
JÉRÔME RIEDI	Professeur, HDR LOA, Université de Lille	(Co-directeur de Thèse)
CHAABANE DJÉRABA	Professeur, HDR CRISTAL, Université de Lille	(Directeur de Thèse)

**Centre de Recherche en Informatique, Signal et Automatique (CRISTAL)**  
BÂTIMENT ESPRIT AVENUE HENRI POINCARÉ 59655 VILLENEUVE D'ASCQ  
FRANCE

**Laboratoire d'Optique Atmosphérique (LOA)**  
BÂTIMENT P5 AVENUE JEAN PERRIN 59650 VILLENEUVE D'ASCQ  
FRANCE

**École Doctorale Mathématiques, sciences du numérique et de leurs interactions (MADIS)**



## Résumé

---

Dans cette thèse, nous nous intéressons à la prédiction de la concentration en aérosols (plus précisément en Particules de Matière ayant un diamètre inférieur à  $2.5 \mu m$  (PM<sub>2.5</sub>)) au niveau du sol. Cette valeur étant liée à la qualité de l'air, et par extension à des dégradations de la santé humaine, l'objectif de cette tâche dépasse le seul enjeu de l'amélioration des connaissances théoriques.

On trouve dans le domaine de la physique atmosphérique des solutions pour la réalisation de cette tâche. Elles présentent une bonne précision et une excellente explicabilité, cependant elles sont gourmandes en temps de calcul. Il existe donc un besoin de solutions performantes mais aussi rapides pour la réalisation de cette tâche.

Notre objectif dans le cadre de cette thèse est donc de proposer une méthode d'estimation de la concentration en aérosols au niveau du sol, bénéficiant de l'efficacité (en termes de rapport entre précision et temps d'inférence) des Réseaux de Neurones Artificiels (Artificial Neural Network) (ANN). Pour développer ces nouvelles solutions nous nous appuyons sur les connaissances a priori du domaine. Elles se révèlent en effet utiles dans le contexte de l'apprentissage semi-supervisé. Cela peut par ailleurs permettre d'augmenter l'explicabilité des modèles. Cet objectif s'inscrit donc dans un contexte précis, mais qui nous laisse toutefois la liberté d'explorer successivement trois directions de recherche.

Dans ce contexte, notre accès à des données annotées est limité. Certaines de nos annotations sont notamment dites "éparses", car elles ne concernent que quelques pixels d'une image donnée. Nous proposons ainsi premièrement une méthode permettant à un modèle de Réseaux Génératifs Adverses (Generative Adversarial Networks) (GAN) d'exploiter ces vérités terrain éparses. Cette solution s'inspire du rôle des Conditions Aux Limites (CAL) dans la résolution d'équations différentielles en physique. Nous démontrons expérimentalement la compétitivité de cette approche.

Nous étudions dans un second temps l'impact de l'utilisation d'une sélection de variables météorologiques comme données d'entrée supplémentaires sur nos résultats. La relation entre ces variables et certains choix d'architecture, comme les méthodes de fusion d'information employées, est en particulier questionnée. Cette étude

nous a permis de préconiser des choix de variables et d'architectures permettant d'améliorer les résultats.

Enfin, nous proposons un modèle de Réseaux Informés par la Physique (Physics-Informed Neural Network) (PINN), utilisant une fonction de coût basée sur une équation physique permettant de décrire la dispersion des aérosols. Celle-ci permet de guider l'apprentissage sans données annotées, ce qui rend cette méthode particulièrement adaptée dans notre contexte d'entraînement semi-supervisé. De plus l'architecture particulière de ce modèle lui permet de générer lui-même certaines des données nécessaires à l'utilisation de cette fonction de coût. Nos expérimentations démontrent l'efficacité de ce modèle comparé à l'état de l'art.

**Mots-Clés :** Intelligence Artificielle, Vision par Ordinateur, Apprentissage Semi-Supervisé, Apprentissage Informé par la Physique

## Abstract

---

In this thesis, we are interested in the prediction of aerosols (more precisely Particulate Matter of diameter inferior to  $2.5\mu m$  (PM<sub>2.5</sub>)) at ground level. This value is linked to air quality, and through it to human health degradations. The objective of this thesis goes therefore beyond the improvement of theoretical knowledge alone.

Solutions to realise this task can be found in the domain of atmospheric physics. They present a good precision and an excellent explicability, however they are also computationally demanding. There is therefore a need for performant and fast solutions for the completion of this task.

Our objective in the context of this thesis is therefore to propose a method of estimation of ground-level aerosol concentration, benefiting from the efficiency (in terms of relationship between accuracy and inference time) of Artificial Neural Networks (ANN). To develop these new solutions, we wish to rely on prior domain knowledge. It can indeed prove useful in the context of semi-supervised learning. It can moreover increase the explainability of the models. This objective is therefore relative to a precise context, but it however lets us the freedom of exploring up to three research directions successively.

In this context, our access to annotated data is limited. Some of our labels are said "sparse" as sometimes only a few pixels of a given image are properly annotated. This is why we propose firstly a method allowing a Generative Adversarial Network to exploit these sparse ground truths. This solution is inspired from the role of Boundary Conditions in solving differential equations in physics. We demonstrate experimentally the competitiveness of this approach.

We study in a second time the impact of the use of a selection of meteorological variables as additional inputs on our results. The relationship between these variables and some architecture choices, such as the fusion strategy used, is particularly explored. Crucial conclusions on the behaviour of ANNs in this type of context are drawn from this study.

Finally, we propose a Physics-Informed Neural Network model, using a physics-

based loss function allowing to describe aerosol dispersion. This loss function allows the model to learn without annotated data, which makes this method all the more adapted to our context of semi-supervised learning. Moreover the particular architecture of this model allows it to generate by itself some of the data necessary to the use of this loss function. Our experiments demonstrate the efficiency of this model compared to the state-of-the-art.

**Keywords :** Artificial Intelligence, Computer Vision, Semi-Supervised Learning, Physics-Informed Learning

## Remerciements

---

Durant cette thèse, un grand nombre de personnes m'ont apporté leur soutien, leurs conseils, ou parfois simplement leur présence. Ils ont en cela joué un rôle crucial dans la réalisation de cette thèse, et je souhaite pour cela les remercier.

Mes premiers remerciements vont à ma compagne, Gabriela Wuilliez, pour m'avoir écouté et vu me plaindre, m'énerver, me décourager pendant ces trois années et m'avoir envers et contre tout apporté un soutien et des encouragements continus.

Ma gratitude va également à ma famille et mes amis, pour leur compréhension et leur écoute indéfectibles. Ils ont su m'offrir des moments de paix salvateurs à travers les difficultés traversées.

Je souhaite aussi exprimer ma gratitude envers mes collègues doctorants Sami, Mireille, Nicolas, Gaspard et Théo, pour avoir grâce à leur bonne humeur et leur sympathie fait de notre bureau commun un environnement de travail agréable.

De même, je souhaite remercier l'ensemble du personnel de l'IMT Nord Europe, pour leur excellent accueil lors des missions d'enseignement effectuées dans leurs locaux. Je remercie en particulier Corentin, Justin, Aymeric, Raphaël et Adrien, pour m'avoir plus d'une fois accueilli dans leur bureau.

Je suis reconnaissant envers Chaabane Djeraba, mon directeur de thèse, pour son encadrement et ses bons conseils.

Je remercie aussi chaleureusement Jérôme Riedi, mon co-directeur, pour sa passion et ses nombreux conseils mais aussi pour m'avoir accompagné dans ma découverte de la physique atmosphérique, et pour s'être ouvert au domaine de l'Intelligence Artificielle et avoir ainsi grandement facilité la communication. Je lui suis également reconnaissant de m'avoir accompagné et ainsi servi de point de repère lors des quelques conférences auxquelles j'ai eu la chance de participer.

Je tiens également à exprimer ma plus grande gratitude à José Mennesson, co-encadrant de cette thèse. Sa confiance en moi, sa disponibilité, son honnêteté, sa bonne humeur et surtout son implication sans failles ont été des atouts précieux

tout au long de cette thèse.

Mes quelques animaux de compagnie ont toujours été pour moi une source de paix, de joie et de force. Durant ces trois années, plusieurs d'entre eux m'ont malheureusement successivement quitté, puis d'autres m'ont rejoint. Il m'est impossible d'envisager terminer ces quelques lignes sans adresser mes pensées les plus chaleureuses et ma gratitude la plus vive à Abu, Loulou, Cookie, Bam, Kumo, Tournesol, Miel et Namu.

Je souhaite aussi adresser ma reconnaissance à toutes les personnes qui, durant mes précédentes années de formation, ont su me transmettre leur passion pour le monde de la recherche et pour celui de l'Intelligence Artificielle, me donnant ainsi l'envie de réaliser cette thèse. Cette décision restera l'une des plus cruciales, et des meilleures, de ma vie.

# Table des Matières

<b>1</b>	<b>Introduction et Contexte</b>	<b>16</b>
1.1	Motivation . . . . .	17
1.2	Objectifs . . . . .	18
1.3	Défis . . . . .	19
1.4	Variables d'intérêt . . . . .	21
1.5	Données simulées . . . . .	22
1.6	Caractérisation de la tâche . . . . .	25
1.7	Contributions et Organisation du Manuscrit . . . . .	26
<b>2</b>	<b>État de l'Art</b>	<b>28</b>
2.1	Méthodes utilisées en sciences atmosphériques . . . . .	30
2.1.1	Méthodes basées sur les Modèle de Transport Chimique (Chemical Transport Model) (CTM) . . . . .	31
2.1.2	Méthodes statistiques . . . . .	31
2.1.3	Apprentissage Automatique (Machine Learning) (ML) . . . . .	33
2.1.4	Apprentissage Profond (Deep Learning) (DL) . . . . .	35
2.2	Méthodes de base pour comparaison . . . . .	38
2.2.1	Interpolation polynômiale . . . . .	38
2.2.2	Algorithme Random Forest . . . . .	39
2.2.3	Krigeage . . . . .	40
2.2.4	Encodeurs-Décodeurs avec Connexions Résiduelles . . . . .	41
2.3	Translation image-image semi-supervisée . . . . .	43
2.3.1	Réseaux Génératifs Adverses (GANs) . . . . .	43
2.3.2	Auto-Encodeurs Variationnels (VAEs) . . . . .	46
2.4	Acceptabilité des modèles . . . . .	49
2.4.1	Modèles Explicables . . . . .	49
2.4.2	Apprentissage Informé par la Physique . . . . .	51
2.5	Conclusion . . . . .	54

<b>3</b>	<b>BC-GAN : un modèle utilisant les vérités terrains éparses</b>	<b>57</b>
3.1	Proposition . . . . .	60
3.1.1	Contrainte dure . . . . .	61
3.1.2	Contrainte douce . . . . .	62
3.2	Protocole expérimental . . . . .	64
3.2.1	Données . . . . .	64
3.2.2	Pré-traitement des données . . . . .	65
3.2.3	Description des expériences . . . . .	66
3.2.3.1	Modèles pour entraînement strictement supervisé . . .	66
3.2.3.2	Modèles pour entraînement semi-supervisé . . . . .	67
3.2.3.3	Configuration des CAL . . . . .	69
3.3	Evaluation des modèles . . . . .	69
3.3.1	Disponibilité du code et des données . . . . .	71
3.4	Résultats . . . . .	72
3.4.1	Encodeurs-décodeurs en apprentissage purement supervisé . . .	72
3.4.2	Contrainte dure et contrainte douce . . . . .	73
3.4.3	Comparaison avec les méthodes proposées comme base de com- paraison . . . . .	76
3.5	Conclusion . . . . .	79
<b>4</b>	<b>Étude d'impact des variables d'entrée</b>	<b>80</b>
4.1	Proposition . . . . .	83
4.1.1	Stratégies de fusion d'information . . . . .	84
4.1.2	Fusion de données / Concaténation de Canaux (CC) . . . . .	84
4.1.3	Fusion de Représentations (FR) . . . . .	85
4.1.4	Fusion de Décisions (FD) . . . . .	87
4.1.5	Approches de fusion hybrides . . . . .	89
4.2	Protocole expérimental . . . . .	90
4.2.1	Description des données . . . . .	91
4.2.2	Pré-traitement des données . . . . .	92
4.2.3	Description des expériences . . . . .	94
4.2.4	Métrique supplémentaire : la FSIM . . . . .	94
4.2.5	Scores . . . . .	95
4.2.6	Disponibilité du code et des données . . . . .	96
4.3	Résultats . . . . .	96
4.3.1	Aperçu . . . . .	96

4.3.2	Meilleurs résultats . . . . .	99
4.3.3	Résultats des méthodes hybrides de fusion . . . . .	104
4.3.4	Comparaison avec les méthodes proposées comme base de comparaison . . . . .	106
4.4	Conclusion . . . . .	107
<b>5</b>	<b>Modèle informé par la physique pour la prédiction de la concentration en aérosols au niveau du sol</b>	<b>110</b>
5.1	Proposition . . . . .	113
5.2	Protocole expérimental . . . . .	116
5.2.1	Données . . . . .	116
5.2.2	Pré-traitement des données . . . . .	117
5.2.3	Evaluation des modèles . . . . .	117
5.2.4	Conditions Aux Limites : configurations et méthodes . . . . .	118
5.2.5	Disponibilité du code et des données . . . . .	119
5.3	Résultats . . . . .	119
5.3.1	Comparaison à l'état de l'art . . . . .	119
5.3.2	Impact de la fonction de coût informée par la physique . . . . .	120
5.3.3	Impact des CAL . . . . .	121
5.3.4	Impact de la proportion d'entraînement non supervisé . . . . .	123
5.3.5	Comparaison visuelle . . . . .	125
5.4	Conclusion . . . . .	125
<b>6</b>	<b>Conclusion et Travaux Futurs</b>	<b>127</b>
6.1	Bilan des Contributions . . . . .	128
6.1.1	GAN avec utilisation de vérités terrain éparses sous forme de CAL	128
6.1.2	Variables d'entrée et choix d'architecture . . . . .	129
6.1.3	PINN pour la prédiction des PM <sub>2.5</sub> au niveau du sol . . . . .	130
6.2	Travaux futurs . . . . .	131
6.2.1	Des métriques spécialisées pour la prédiction des termes de sources et de pertes . . . . .	131
6.2.2	La robustesse face aux changements de résolution et de couverture spatiales . . . . .	132
6.2.3	Une meilleure prise en compte de l'aspect temporel . . . . .	132
6.2.4	L'application aux données réelles . . . . .	133
<b>A</b>	<b>Annexes</b>	<b>134</b>

A.1	Procédure d'égalisation des valeurs de concentration . . . . .	135
A.2	Architecture des discriminateurs des GANs pour chaque stratégie de fusion . . . . .	136
A.2.1	Fusion de données / Concaténation de Canaux . . . . .	136
A.2.2	Feature Fusion . . . . .	136
A.2.3	Decision Fusion . . . . .	136
A.3	Utilisation de la FSIM en tant que fonction de coût additionnelle . . . . .	138
A.4	Complexité des modèles en fonction de la stratégie de fusion . . . . .	138
A.5	Performances d'encodeurs-décodeurs convolutifs et linéaires . . . . .	139
A.6	Performance de notre PINN pour la prédiction des termes de Sources et de Pertes . . . . .	140
A.7	Performance de notre PINN sur les valeurs extrêmes . . . . .	141
A.8	Etude d'ablation des différentes couches du UNet . . . . .	142
A.9	Performance de notre PINN avec d'autres horizons de prédiction . . . . .	143
<b>B</b>	<b>Financement</b>	<b>149</b>
<b>C</b>	<b>Bibliographie</b>	<b>151</b>
	Références . . . . .	152
	Publications . . . . .	161

# Liste des Abréviations

- ANN** – Réseaux de Neurones Artificiels (Artificial Neural Network).
- AOD** – Profondeur Optique en Aérosols (Aerosol Optical Depth).
- AOT** – Epaisseur Optique en Aérosols (Aerosol Optical Thickness).
- BC-GAN** – Réseau Génératif Adverse avec utilisation de Conditions Aux Limites (Generative Adversarial Network with use of Boundary Conditions).
- CAL** – Conditions Aux Limites.
- CAL/I** – Conditions Aux Limites et Initiales.
- CAMS** – Service Copernicus de Surveillance de l’Atmosphère (Copernicus Atmosphere Monitoring Service).
- CC** – Concaténation de Canaux.
- CI** – Conditions Initiales.
- CNN** – Réseaux de Neurones Convolutif (Convolutional Neural Network).
- CTM** – Modèle de Transport Chimique (Chemical Transport Model).
- DL** – Apprentissage Profond (Deep Learning).
- ED** – Equation Différentielle.

- FD** – Fusion de Décisions.
- FR** – Fusion de Représentations.
- FSIM** – index de SIMilarité de Caractéristiques ou de Représentations (Feature SIMilarity index).
- GAN** – Réseaux Génératifs Adverses (Generative Adversarial Networks).
- GOESNG** – Satellite Géostationnaire Opérationnel et Environmental - Nouvelle Génération (Geostationary Operational Environmental Satellites - New Generation).
- LSTM** – Longue Mémoire à Court Terme (Long Short-Term Memory).
- MAE** – Erreur Absolue Moyenne (Mean Absolute Error).
- MBE** – Biais Moyen (Mean Bias Error).
- ML** – Apprentissage Automatique (Machine Learning).
- MSE** – Erreur Quadratique Moyenne (Mean Squared Error).
- MSG** – Meteosat Deuxième Génération (Meteosat Second Generation).
- PI-GAN** – Réseaux Génératifs Adverses Informés par la Physique (Physics-Informed Generative Adversarial Network).
- PICNN** – Réseaux de Neurones Convolutifs Informés par la Physique (Physics-Informed Convolutional Neural Networks).
- PINN** – Réseaux Informés par la Physique (Physics-Informed Neural Network).
- PM<sub>2.5</sub>** – Particules de Matière ayant un diamètre inférieur à 2.5  $\mu\text{m}$ .
- QE** – Erreur Quantifiée (Quantized Error).
- rMAE** – Erreur Absolue Moyenne relative (relative Mean Absolute Error).
- rMBE** – Biais Moyen relatif (relative Mean Bias Error).
- RNN** – Réseaux de Neurones Récurrent (Recurrent Neural Network).

**VAE** – Auto-Encodeurs Variationnels (Variational Autoencoders).

**XAI** – Intelligence Artificielle eXplicable (eXplainable Artificial Intelligence).

# 1

## Introduction et Contexte

### Sommaire

---

1.1	Motivation . . . . .	17
1.2	Objectifs . . . . .	18
1.3	Défis . . . . .	19
1.4	Variables d'intérêt . . . . .	21
1.5	Données simulées . . . . .	22
1.6	Caractérisation de la tâche . . . . .	25
1.7	Contributions et Organisation du Manuscrit . . . . .	26

---

## 1.1 Motivation

Dans ce manuscrit, nous nous intéressons à la concentration en aérosol dans les plus basses couches de l’atmosphère, donc au niveau du sol. Plus précisément, nous nous intéressons en particulier à la concentration en PM<sub>2.5</sub> (Particules de Matière ayant un diamètre inférieur à 2.5  $\mu m$ ). A cette taille, ces particules sont à la fois difficiles à détecter et dangereuses pour la santé humaine. En effet le niveau d’exposition aux PM<sub>2.5</sub> est fortement associé à un certain nombre de conséquences néfastes pour la santé [1, 2, 3]. De plus l’étude du Global Burden of Disease (Charge Mondiale de la Maladie) a reconnu la pollution de l’air comme le cinquième plus important facteur de risque de mortalité dans le monde [4]. C’est pour ces raisons que la concentration en PM<sub>2.5</sub> est un important indicateur de la qualité de l’air. L’obtention de ces valeurs de concentration, à échelle régionale ou globale, présente donc un intérêt certain en termes de politique de santé publique, en plus de son aspect scientifique.

Dans le domaine de la physique atmosphérique, des solutions purement mathématiques peuvent être utilisées, souvent basées sur la résolution d’un système d’équations différentielles. Cependant si les méthodes purement physique présentent une bonne précision ainsi qu’une excellente explicabilité, elles sont gourmandes en temps de calcul. Or pour ce type de problématique, un modèle efficace et présentant un temps d’inférence court, est nécessaire.

Les ANN classiques représentent en ce sens une solution intéressante. Cependant du fait de leur manque d’explicabilité, ils souffrent de problèmes d’acceptabilité au sein de la communauté de physique atmosphérique. L’apparition d’algorithmes dit explicables, mais également et en particulier des Réseaux Informés par la Physique (Physics-Informed Neural Network) (PINN), a prouvé que ces deux objectifs d’efficacité et d’interprétabilité, pouvaient être réconciliés. Elle a prouvé par la même occasion l’utilité d’intégrer dans la conception des solutions des informations a priori sur les spécificités du problème à résoudre.

La motivation des travaux présentés ici est basée sur cette observation. Nous cherchons en effet à concevoir des algorithmes d’apprentissage profond inspiré par les connaissances sur la physique du problème traité. Cela permettrait ainsi à ces solutions de gagner en interprétabilité comme en efficacité.

Avant de procéder plus avant dans la présentation de nos travaux, nous attirons l’attention du lecteur sur le fait que les études réalisées et présentées ci-après sont fon-

damentalement pluridisciplinaires. En effet nos travaux traitent à la fois de physique atmosphérique, en raison du domaine d’application spécifique, et d’apprentissage machine et de vision artificielle, au regard des solutions développées pour répondre au problème proposé. Les travaux présentés ici se situent à l’interface de ces deux domaines, et il est nécessaire de maîtriser certains concepts de base pour en saisir pleinement l’approche. En effet, la nature des données utilisées, issues de la composition atmosphérique, influence significativement la direction des recherches et motive de nombreux choix d’architecture et d’implémentation. Pour faciliter la compréhension de ces résultats par des scientifiques des deux disciplines, nous avons jugé essentiel de présenter certaines généralités concernant la vision artificielle et les techniques d’apprentissage automatique, tout en apportant des explications sur les approximations faites quant à l’utilisation des données atmosphériques.

Du fait de cette pluridisciplinarité, il est possible que certaines descriptions puissent paraître superflues pour une communauté alors qu’elles sont essentielles pour l’autre. Des approximations ont été inévitables dans l’utilisation des données atmosphériques, ce qui pourrait aussi apparaître comme de sérieux raccourcis pour des spécialistes du domaine. Néanmoins, il s’agissait d’une nécessité afin de présenter nos résultats de manière compréhensible à un public plus familier avec la manipulation d’images contenant des objets bien délimités ou des informations sémantiques. Nous demandons ainsi la patience du lecteur concernant ces descriptions, qui visent à assurer une compréhension mutuelle. Le but de cette thèse n’était pas d’appliquer des méthodes disponibles en intelligence artificielle pour la prévision de la qualité de l’air, mais d’explorer plus fondamentalement comment certaines caractéristiques des problèmes en sciences atmosphériques peuvent améliorer les architectures et les modes d’apprentissage des réseaux de neurones artificiels, tout en garantissant une certaine interprétabilité des résultats obtenus.

### **1.2 Objectifs**

Notre objectif dans le cadre de cette thèse consiste en la conception de solutions inspirées par la physique pour la prédiction de la qualité de l’air au niveau du sol. L’étude expérimentale de ces méthodes nous permet de mieux comprendre comment les connaissances physiques peuvent servir à améliorer la performance de techniques de Apprentissage Profond (Deep Learning) (DL). Nous nous intéressons à trois idées majeures :

**La conception d'une solution pour l'utilisation de vérités terrain éparses pour l'apprentissage.** Les données de qualité de l'air obtenues via capteurs ont la particularité d'être des valeurs éparses, plutôt que des images ou matrices. L'utilisation de ce type de valeurs pour l'apprentissage d'un modèle de Vision Artificielle représente donc un challenge intrinsèque. Nous souhaitons donc proposer une méthode pour la prise en compte de ce type de données par nos modèles de DL, inspirée de la manière dont ce type d'information est typiquement traité en physique. L'utilité de cette méthode pourrait tout-à-fait dépasser son cadre initial d'utilisation.

**La conception et l'étude de modèles d'apprentissage multi-variés pour la physique atmosphérique.** Certaines des méthodes typiquement utiles en sciences atmosphériques pour la prédiction de la qualité de l'air ont la particularité d'utiliser un grand nombre de variables. Celles-ci sont choisies pour leur pouvoir explicatif vis-à-vis des valeurs à prédire. L'un de nos objectifs est donc d'étudier si certaines de ces variables peuvent également se révéler utiles pour un modèle de DL. Nous voulons également à travers cette même étude vérifier si la nature des variables d'entrée utilisées peut avoir une influence sur les choix d'architecture du réseau de neurones utilisé.

**La conception d'un modèle de Réseaux de Neurones Informé par la Physique.** Les différentes variables disponibles sont liées entre elles et à la qualité de l'air par le biais d'équations physiques. Ces équations se montrent utiles pour résoudre le problème de la prédiction de la qualité de l'air par des méthodes purement basées sur la physique. Un objectif de cette thèse est d'employer l'une de ces équations pour faciliter l'apprentissage de nos modèles. Nous visons ainsi l'augmentation à la fois de leur interprétabilité et de leur efficacité.

### 1.3 Défis

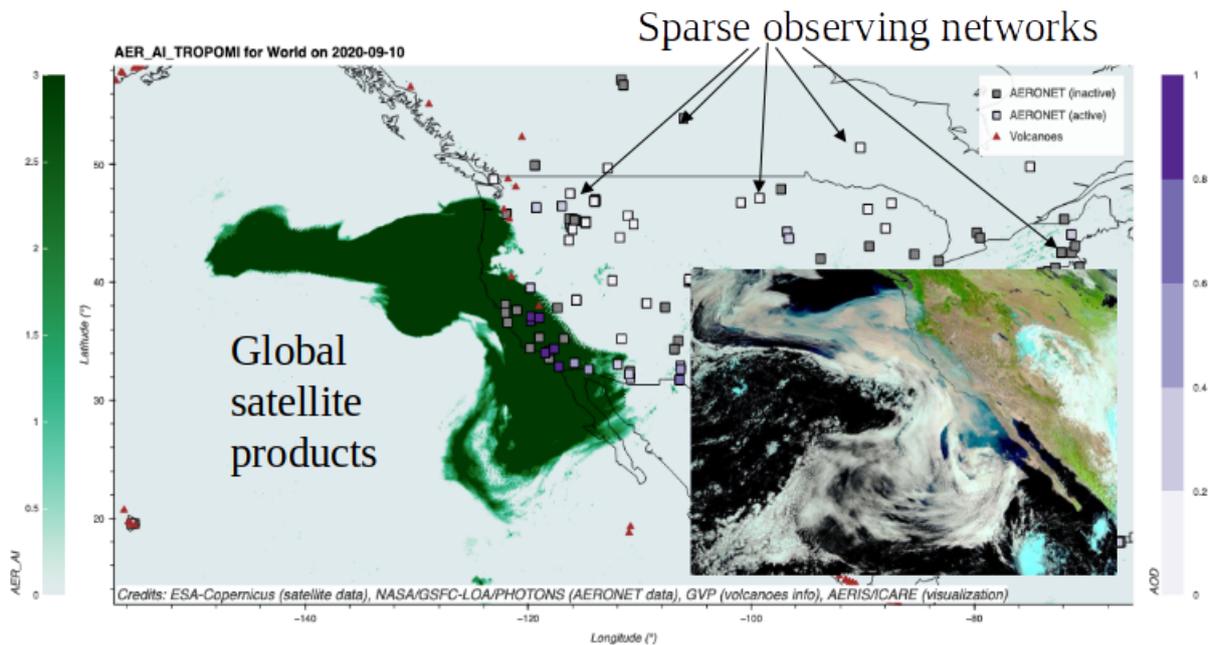
L'objectif de cette thèse est l'estimation de la qualité de l'air à partir de mesures de télédétection spatiale de l'Épaisseur Optique en Aérosols (Aerosol Optical Thickness) (AOT), aussi appelée parfois Profondeur Optique en Aérosols (Aerosol Optical Depth) (AOD).

L'AOD est une quantité optique intégrée sur toute la colonne atmosphérique. Cette mesure, en plus d'être obtenue par des capteurs au sol (mesure locale), peut aussi être obtenue par des instruments de mesure optique satellite (mesure plus globale).

La qualité de l'air est ici définie par la concentration en fines particules d'aérosols.

Nous nous intéressons en particulier à la concentration en  $PM_{2.5}$  au niveau du sol, du fait de ses effets néfastes sur la santé humaine, comme évoqué en section 1.1.

Contrairement à l'AOD, la mesure des  $PM_{2.5}$  ne peut être réalisée que localement par des capteurs au sol (voir figure 1.1). En effet, du fait de leur taille, ces aérosols sont difficiles à observer de manière isolée depuis l'espace avec les instruments actuellement disponibles.



**Figure 1.1:** Réseau de capteurs de concentration (au sol) en aérosols. L'image principale est en fausse couleur et représente une mesure liée à la concentration en aérosols. La petite image dans le coin inférieur droit est une visualisation d'aérosols dans l'atmosphère, en bleu clair, superposée à une image en vraies couleurs.

Des solutions complexes, basées sur le couplage entre modèles chimiques et de transport [5], permettent la modélisation et la prédiction de la concentration en  $PM_{2.5}$ . Cependant ces modèles sont aussi très gourmands en puissance de calcul et ne permettent pas tous l'obtention de valeurs de  $PM_{2.5}$  au niveau du sol en temps réel ni quasi-réel.

Il y a donc un important besoin d'une méthode fiable et rapide d'estimation (voire de prédiction) de la concentration en  $PM_{2.5}$  à partir d'autres variables et observations, et c'est là le défi majeur de cette thèse.

## 1.4 Variables d'intérêt

Comme exposé en section 1.3, il existe un besoin d'une solution d'estimation de la concentration en aérosols. Pour cela, nous devons déterminer quelles données peuvent être exploitées pour l'estimation de cette concentration.

L'AOD est un excellent candidat car c'est une donnée disponible en quantité et qui peut être obtenue en temps quasi-réel, par exemple grâce à l'utilisation d'instruments optiques opérant depuis cinq satellites géostationnaires [6, 7]. Parmi ces cinq satellites, on trouve deux satellites Meteosat Deuxième Génération (Meteosat Second Generation) (MSG) couvrant l'Europe et le Proche-Orient. On trouve aussi le satellite Himawari, qui couvre l'Asie de l'Est. Enfin, on trouve deux Satellite Géostationnaire Opérationnel et Environmental - Nouvelle Génération (Geostationary Operational Environmental Satellites - New Generation) (GOESNG) qui couvrent les Amériques. Grâce à la couverture cumulée de ces satellites, il est possible d'obtenir des valeurs de l'AOD sous forme de matrices couvrant la planète entière. Ces valeurs ne sont donc pas éparses.

Cependant, la concentration en PM<sub>2.5</sub> n'est pas fortement corrélée avec l'AOD obtenue par satellite. En effet, l'AOD ne dépend pas seulement de la concentration en aérosols, mais également de la distribution de tailles de ceux-ci, de la forme de leurs particules, de leur composition chimique, ainsi que de l'humidité atmosphérique (à cause de la croissance hygroscopique des particules). Boucher [8] décrit à travers plusieurs équations la relation entre concentration en aérosols dans l'atmosphère et valeurs de l'AOD. Cette relation est hautement multi-variée et complexe. Pour cette raison, de précédentes tentatives de dériver la concentration en PM<sub>2.5</sub> au niveau du sol à partir des seules observations spatiales de l'AOD ont montré de claires limitations, malgré des résultats intéressants [9]. L'utilisation conjointe de données satellites et de réseaux dédiés à l'observation des PM<sub>2.5</sub> en surface fait partie intégrante des stratégies développées au cours des dernières années pour la surveillance de la qualité de l'air à l'échelle globale [10].

Comme dit précédemment, des variables météorologiques telles que l'humidité de l'air peuvent avoir un impact sur la concentration en aérosols, ses variations et sa relation avec l'AOD. C'est aussi le cas de la température et de la pression atmosphérique. Différents modèles de transports des aérosols présentent aussi le vent comme ayant un fort impact sur leurs déplacements [8]. Ces quatre variables météorologiques (cinq si l'on sépare vitesse et direction du vent) sont donc également candidates en tant que données d'entrée des solutions envisagées. Elles peuvent être

fournies par la plupart des modèles météorologiques, et sont disponibles en temps réel et pour différents horizons de prédiction.

De plus la taille des particules des aérosols a de même un impact sur leur transport, et sur leur relation avec l'AOD. Une variable optique, nommée le coefficient d'Ångström, permet de fournir une information sur la taille des particules. Ce coefficient est défini et peut être aisément calculé à partir de deux mesures de l'AOD à deux longueurs d'onde différentes, comme illustré par l'équation 1.1. Dans cette équation,  $\lambda_1$  et  $\lambda_2$  représentent deux longueurs d'onde différentes, tandis que  $AOD(\lambda_1)$  et  $AOD(\lambda_2)$  représentent les valeurs de l'AOD pour chacune de ces deux longueurs d'onde. "AE" représente l'"Ångström Exponent" ou Coefficient d'Ångström.

$$AE(\lambda_1, \lambda_2) = -\frac{\ln\left(\frac{AOD(\lambda_2)}{AOD(\lambda_1)}\right)}{\ln\left(\frac{\lambda_2}{\lambda_1}\right)} \quad (1.1)$$

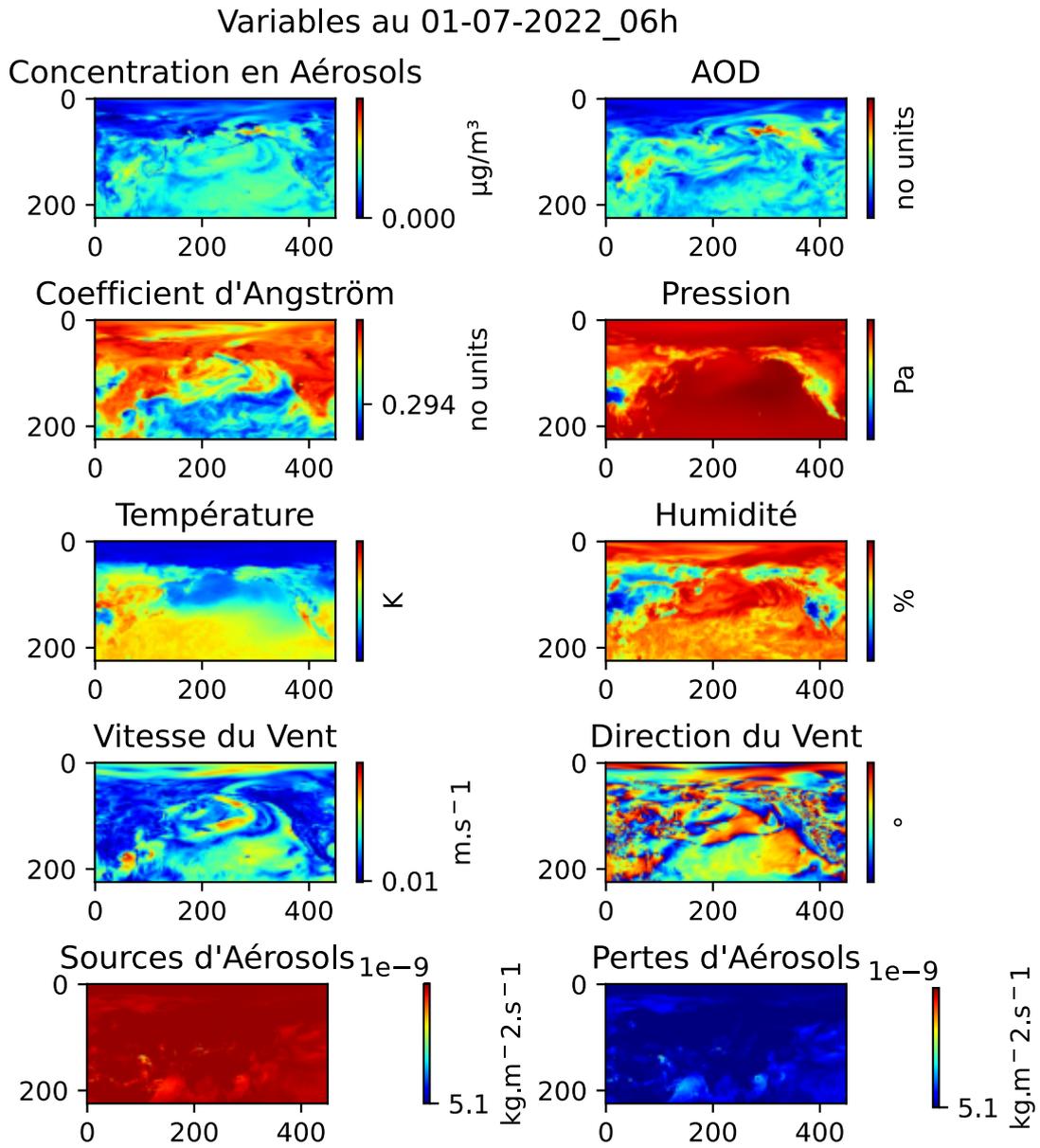
Le coefficient d'Ångström est en général un paramètre tout aussi disponible que l'AOD, en particulier lorsque des capteurs modernes et/ou dédiés à l'observation des aérosols sont considérés. Cela en fait un candidat supplémentaire en tant que donnée d'entrée pour les solutions envisagées.

Enfin, un dernier type de variable est utile pour l'estimation de la concentration en PM<sub>2.5</sub> : les termes de sources et de pertes d'aérosols. Ces termes interviennent dans des modèles de transport d'aérosols [8], et représentent des valeurs qui ne peuvent pas être assimilées aux processus de diffusion ou de convection. Les feux de forêts ou éruptions volcaniques sont des évènements qui peuvent être sources d'aérosols. Les processus de sédimentation ou de déposition sont, quant à eux, représentés dans les termes de pertes. Cependant, ces valeurs, bien que très utiles, ne sont pas aussi aisées à obtenir que les variables météorologiques par exemple, et ne sont pas toujours disponibles en temps réel.

La figure 1.2 donne un aperçu de chacune de ces variables.

## 1.5 Données simulées

Dans cette thèse, nous abordons l'estimation de la concentration en aérosols via des méthodes d'apprentissage profond (Deep Learning) semi-supervisées. Ces approches nécessitent d'avoir une vérité terrain, soit la concentration en aérosols en tout point



**Figure 1.2:** Exemple de chacune des variables considérées pour l'estimation de la concentration en aérosols.

du globe, pour réaliser l'apprentissage et pour obtenir des scores de précision. Or, cette information ne peut être collectée que localement. C'est pourquoi, au cours des différentes expériences réalisées durant cette thèse, des données simulées sont utilisées à la place de données réelles. Ces données sont fournies par deux modèles : le Service Copernicus de Surveillance de l'Atmosphère (Copernicus Atmosphere Monitoring Service) (CAMS) [6], et le modèle ALADIN <sup>1</sup>. Cela permet un accès rapide et assuré à toutes les variables d'intérêt. Elles prennent en effet ainsi toutes les formes de matrices (ou images) de la même taille. Cela rend également plus aisé la manipulation des jeux de données, afin notamment de concevoir différents scénarios durant les expérimentations. Enfin, cela est nécessaire pour l'évaluation des solutions, puisque les données réelles ne contiennent que des valeurs éparses de la concentration en aérosols, or des valeurs complètes sont requises pour l'évaluation.

Ces données sont tout de même utilisées dans le but de reproduire un scénario réaliste. Ainsi, le scénario typique considéré lors des expériences sépare les données disponibles pour l'entraînement en deux périodes. Pendant la première période, les données générées par le modèle de simulation sont considérées comme étant disponibles. En effet, ces données ne sont pas disponibles en temps réel, et sont souvent obtenues de manière plus fiable après un processus de réanalyse (ce qui sous-entend qu'il est réalisé *a posteriori*, après les événements météorologiques considérés). Elles ne sont donc disponibles que pendant une période. Durant cette première période, des données complètes de concentration en PM<sub>2.5</sub> sont disponibles, tout comme les termes de source et de perte d'aérosols.

Pendant la deuxième période, seules les données réelles sont considérées comme disponibles. Les termes de source et de perte des aérosols sont donc indisponibles, et seules des valeurs éparses de la concentration en aérosols sont disponibles. Pour reproduire cette situation, des matrices éparses représentant la concentration en PM<sub>2.5</sub> sont calculées à partir des matrices complètes fournies par les modèles. Cela permet d'avoir un certain contrôle sur le caractère éparses de ces matrices : il est par exemple possible de contrôler le nombre de pixels qu'elles doivent contenir, ou encore la méthode utilisée pour choisir ces pixels.

Les modèles CAMS et ALADIN présentent plusieurs différences. La taille des images fournies diffère en fonction du modèle, mais pas uniquement. En effet le modèle CAMS fournit des images de tailles différentes selon certains paramètres. Les matrices

---

<sup>1</sup><https://www.aeris-data.fr/en/landing-page/?uuid=c5e98410-1f44-4d65-9fbb-97ee350ad7e6>

fournies par le modèle CAMS couvrent toujours la planète entière. Celles fournies par le modèle ALADIN couvrent l'Europe et une partie du Proche-Orient et de l'Afrique du Nord. Du fait de cette couverture géographique réduite, leur résolution (au sens de l'inverse de la surface en  $km^2$  par pixel) est plus importante.

L'objectif des solutions proposées n'est pas de réaliser de meilleures estimations (en termes de précision ou de ressemblance à la réalité) que les modèles CAMS ou ALADIN. En effet, ceux-ci agissent davantage en tant que guides qu'en tant que membre de la compétition. Ce type de modèle fournit les seules données complètes de concentration en aérosols disponibles, donc le seul point de comparaison disponible pour évaluer les solutions proposées. L'objectif est de fournir des estimations aussi proches que possible de celles fournies par ces deux modèles, tout en étant plus rapide.

Les données du modèle CAMS sont disponibles via la banque de données atmosphériques Copernicus <sup>2</sup>. Les données du modèle ALADIN utilisées durant ces travaux de recherche sont disponibles sur la plateforme AERIS <sup>3</sup>. Les ensembles de données utilisés dans le cadre de cette thèse, provenant des modèles CAMS et ALADIN, sont tous accessibles via le centre de données et de service AERIS/ICARE <sup>4</sup>.

## 1.6 Caractérisation de la tâche

La donnée d'entrée principale utilisée lors des expériences effectuées durant cette thèse est l'AOD. Cette donnée prend la forme d'une matrice ou image. Comme évoqué en section 1.4, d'autres variables d'entrée peuvent être utilisées en plus de l'AOD, mais elles prennent également cette forme. L'objectif est d'obtenir une matrice (ou image) complète de la concentration en aérosols. Ainsi, une image (ou éventuellement plusieurs) est utilisée comme entrée, et une autre image est obtenue en sortie. La tâche à effectuer correspond donc en ce sens à de la translation image-vers-image.

Cependant, cette tâche pourrait aussi être envisagée d'un autre point de vue. En effet, des valeurs éparses de concentration en PM<sub>2.5</sub> sont disponibles, et on cherche à obtenir une matrice complète de ces mêmes valeurs. Puisqu'une partie de la sortie désirée est disponible, cette tâche peut être assimilée à de l'in-painting ou out-painting. Ces processus permettent l'obtention d'une image complète à partir d'une portion de celle-ci [11].

<sup>2</sup><https://ads.atmosphere.copernicus.eu/#!/home>

<sup>3</sup><https://www.aeris-data.fr/en/landing-page/?uid=c5e98410-1f44-4d65-9fbb-97ee350ad7e6>

<sup>4</sup><https://www.icare.univ-lille.fr/>

La complexité du problème, la multiplicité des données d’entrées utilisables, et la faible quantité de valeurs de concentration en aérosols disponibles, amènent à préférer l’approche de la translation image-vers-image.

## 1.7 Contributions et Organisation du Manuscrit

Plusieurs contributions scientifiques ont été réalisées durant cette thèse, en accord avec les objectifs définis dans la section 1.2. Dans cette section, nous présentons ces contributions et leur organisation au sein de ce manuscrit.

**Chapitre 2 - État de l’Art.** Il est important, avant toute description de nos contributions, de proposer une description des domaines de recherches connexes à ces travaux. Nous y décrivons notamment les méthodes typiquement utilisées en physique atmosphérique pour répondre à notre problématique, mais également des concepts clés liés à l’apprentissage semi-supervisé et à la notion d’acceptabilité des modèles. Cette description est primordiale pour la bonne compréhension de cette thèse, ainsi que des défis scientifiques auxquels elle tente de répondre. Les méthodes utilisées comme références pour l’évaluation de nos modèles sont également présentées dans ce chapitre.

**Chapitre 3 - BC-GAN : un modèle utilisant les vérités terrain éparses.** Ce chapitre se concentre sur le premier objectif défini dans la section 1.2. Nous y proposons de considérer les vérités terrain éparses comme des Conditions Aux Limites (CAL). La littérature contient différentes solutions pour la gestion de ces CAL, habituellement appliquées à des Réseaux Informés par la Physique (Physics-Informed Neural Network) (PINN). Nous proposons de les appliquer à un Réseaux Génératifs Adverses (Generative Adversarial Networks) (GAN), d’où le terme de Réseau Génératif Adverse avec utilisation de Conditions Aux Limites (Generative Adversarial Network with use of Boundary Conditions) (BC-GAN). Des expériences sont réalisées pour étudier l’impact de ces CAL sur la performance de nos modèles, et également pour identifier la méthode la plus efficace dans ce contexte.

**Chapitre 4 - Étude d’impact des variables d’entrée.** Nous abordons dans ce chapitre le deuxième objectif de la section 1.2. Une sélection de différentes variables météorologiques est proposée, ainsi que différents modèles et architectures. Une étude extensive est proposée, afin d’évaluer l’impact individuel et conjoint de ces éléments sur les résultats. Une réflexion sur la nature des variables utilisées en lien avec les choix d’architecture y est également présentée.

**Chapitre 5 - Modèle informé par la physique pour la prédiction de la concentration en aérosols au niveau du sol.** Le troisième et dernier objectif présenté en section 1.2 est atteint dans ce chapitre. Un modèle de Réseaux Informés par la Physique (Physics-Informed Neural Network) (PINN) y est proposé. Son architecture atypique lui permet d'utiliser une fonction de coût basée sur une équation physique. Celle-ci permet à son tour de guider efficacement l'apprentissage du modèle sans besoin de données annotées. Cela rend ce modèle particulièrement adapté à un régime d'apprentissage semi-supervisé. Des expériences sont réalisées pour évaluer l'efficacité du modèle, et ce avec différentes quantités de données annotées. Une étude d'ablation est proposée pour évaluer l'influence de différentes caractéristiques du modèle (notamment son architecture et sa fonction de coût) sur ses performances. Enfin les méthodes de gestion des CAL proposées dans le chapitre 3 sont appliquées à ce PINN. Des expériences sont alors proposées pour mieux comprendre ces méthodes et leur influence sur les résultats en fonction du contexte et du modèle auxquelles elles sont appliquées.

**Chapitre 6 - Conclusion et Travaux Futurs.** Ce dernier chapitre présente un récapitulatif de nos contributions et de leurs résultats, et résume les conclusions finales de nos travaux. Des pistes de recherche pour d'éventuels travaux futurs sur cette thématique y sont également proposées.

# 2

## État de l'Art

### Sommaire

---

2.1	Méthodes utilisées en sciences atmosphériques . . . . .	30
2.1.1	Méthodes basées sur les Modèle de Transport Chimique (Chemical Transport Model) (CTM) . . . . .	31
2.1.2	Méthodes statistiques . . . . .	31
2.1.3	Apprentissage Automatique (Machine Learning) (ML) . . . . .	33
2.1.4	Apprentissage Profond (Deep Learning) (DL) . . . . .	35
2.2	Méthodes de base pour comparaison . . . . .	38
2.2.1	Interpolation polynômiale . . . . .	38
2.2.2	Algorithme Random Forest . . . . .	39
2.2.3	Krigeage . . . . .	40
2.2.4	Encodeurs-Décodeurs avec Connexions Résiduelles . . . . .	41
2.3	Translation image-image semi-supervisée . . . . .	43
2.3.1	Réseaux Génératifs Adverses (GANs) . . . . .	43
2.3.2	Auto-Encodeurs Variationnels (VAEs) . . . . .	46
2.4	Acceptabilité des modèles . . . . .	49
2.4.1	Modèles Explicables . . . . .	49

2.4.2	Apprentissage Informé par la Physique . . . . .	51
2.5	Conclusion . . . . .	54

---

Ce chapitre propose une revue de la littérature, dans le but de situer les travaux décrits dans cette thèse par rapport aux dernières avancées dans le domaine de la vision artificielle. L’accent est mis sur les réseaux de neurones, et sur les méthodes pertinentes dans le contexte décrit dans le chapitre 1. Cet état de l’art est divisé en cinq sections principales.

La section 2.1 s’intéresse aux méthodes typiquement utilisées dans la communauté des sciences atmosphériques pour répondre à la problématique proposée.

Dans la section 2.2 différentes méthodes et solutions sont présentées. Celles-ci serviront de base de comparaison pour évaluer la performance relative des modèles et contributions proposés dans cette thèse.

Le chapitre 1 spécifie que les vérités terrain ne sont disponibles que sur une partie du jeu de données. Il y a par conséquent un besoin d’une solution permettant l’utilisation de données à la fois avec et sans vérité terrain. Cela amène naturellement vers le domaine de l’apprentissage semi-supervisé, sujet de la section 2.3.

Ces travaux sont amenés à intéresser non seulement la communauté de la vision artificielle, mais aussi celle de la physique atmosphérique. Pour cette raison, l’acceptabilité du réseau est un paramètre important dans la conception des solutions. Le principal obstacle des ANN à leur acceptabilité est leur aspect boîte noire. La section 2.4 évoque les pistes de solution à ce problème.

Enfin la section 2.5 résume le positionnement de ces travaux de recherche vis-à-vis de l’état de l’art.

### **2.1 Méthodes utilisées en sciences atmosphériques**

Au sein de la littérature propre à la communauté de physique atmosphérique, et parmi les très nombreuses études ayant pour but de générer les concentrations en PM<sub>2.5</sub> à partir d’images satellite, plusieurs catégories de méthodes peuvent généralement être identifiées.

### 2.1.1 Méthodes basées sur les Modèles de Transport Chimique (Chemical Transport Model) (CTM)

Les CTM sont des modèles numériques simulant typiquement la chimie atmosphérique, et qui peuvent être utilisés pour fournir des prédictions de pollution de l'air. Ces modèles ne se focalisent pas sur tous les processus atmosphériques mais s'intéressent plutôt à un nombre restreint d'espèces chimiques. Ils s'appuient sur des équations de transport et, pour chaque espèce, représentent son cycle complet. Ainsi les phénomènes de production et pertes chimiques, les dépôts et les flux, doivent être pris en compte. Ils sont parfois inclus dans des modèles de plus grande ampleur, comme des Modèles de Circulation Générale (GCM). Le modèle CHIMERE [12] est un exemple de CTM.

Des méthodes basées sur ces CTM permettent l'inférence de valeurs de concentration en PM<sub>2.5</sub> au niveau du sol. Elles établissent un facteur d'échelle entre les valeurs simulées d'AOD et de PM<sub>2.5</sub> [13, 14], qui peut ensuite être transféré pour estimer la concentration en PM<sub>2.5</sub> au niveau du sol à partir de l'AOD obtenue par satellite [15, 16]. La précision de ce type de méthode dépend fortement de la variabilité spatio-temporelle du facteur d'échelle. Il souffre donc de limites claires dans le cas où cette variabilité n'est pas convenablement prise en compte ou représentée par le modèle de dimensionnement.

### 2.1.2 Méthodes statistiques

Le second type de méthodes est axé sur les données. Il a pour but d'établir des relations univariées ou multivariées entre la concentration en PM<sub>2.5</sub> observée au niveau du sol et divers facteurs (dont par exemple l'AOD) influant ces valeurs.

Différentes méthodes de régression peuvent être utilisées dans ce cadre [17, 18], de la simple régression linéaire (éventuellement multiple) à la régression géographiquement pondérée, en passant par des modèles à effets mixtes.

**La régression linéaire multiple** est une méthode cherchant à exprimer une variable  $y$  en fonction d'une série de  $p$  variables explicatives  $x_1, \dots, x_p$ . Si  $n$  échantillons  $(y_i, x_{1i}, \dots, x_{pi})$  avec  $i = 1, \dots, n$  sont disponibles, l'équation de régression peut alors s'écrire comme l'équation 2.1. Dans cette équation, les termes  $a_0, \dots, a_p$  sont des paramètres à estimer en exploitant les observations, et  $\epsilon_i$  représente l'erreur du modèle, qui doit être minimisée. Le principe de la régression linéaire simple est le même,

mais avec une seule variable  $x$ .

$$y_i = a_0 + a_1x_{1i} + \dots + a_px_{pi} + \epsilon_i \quad (2.1)$$

**Les modèles à effets mixtes (MEM)** sont des méthodes de régression cherchant à atteindre un équilibre entre les notions de précision et de parcimonie [18]. Pour chaque paramètre à estimer, on introduit une partie dite fixe, et une partie aléatoire. Le modèle est donc dit à effets mixtes dans le sens où il représente des effets fixes et aléatoires. Ces derniers représentent en général le caractère non linéaire des relations entre certaines des variables utilisées, ou de la relation de ces variables au temps par exemple. L'idée générale des MEM est représentée par l'équation 2.2. Cette équation garde les mêmes notations que l'équation 2.1, à ceci près que les termes  $a_0, \dots, a_p$  sont des paramètres fixes tandis que les termes  $b_0, \dots, b_p$  sont des termes aléatoires.

$$y_i = (a_0 + b_0) + (a_1 + b_1)x_{1i} + \dots + (a_p + b_p)x_{pi} + \epsilon_i \quad (2.2)$$

**La régression géographiquement pondérée** est une méthode adaptée au domaine de l'analyse spatiale. Les coordonnées géographiques des points considérés sont donc prises en compte. Cette méthode est adaptée aux situations dans lesquelles les relations entre les différentes variables dépendent de la localisation [18]. L'équation 2.3 décrit la structure fondamentale de cette méthode.  $\mu_i$  et  $\nu_i$  représentent les coordonnées du point auquel l'échantillon  $(y_i, x_{1i}, \dots, x_{pi})$  correspond. Les fonctions  $\beta_0, \dots, \beta_p$  définissent la relation spatiale entre les variables.

$$y_i = \beta_0(\mu_i, \nu_i) + \beta_1(\mu_i, \nu_i)x_{1i} + \dots + \beta_p(\mu_i, \nu_i)x_{pi} + \epsilon_i \quad (2.3)$$

Hu et al. [19] proposent l'utilisation d'un modèle à deux étages, utilisant un modèle linéaire à effets mixtes linéaire pour le premier étage, et une régression géographiquement pondérée sur le second, ce qui leur permet de n'appliquer la pondération géographique que sur une partie des variables utilisées. Des méthodes de krigeage sont parfois utilisés également [18]. Ce type de méthode est décrit plus amplement en section 2.2.3.

**Les modèles additifs généralisés**[20] ont pour spécificité leur capacité à prendre en compte l'éventuelle non-linéarité de la relation entre les variables explicatives  $x_m$  et la variable à prédire  $y$ . Ils associent en effet à chaque variable explicative  $x_m$  une

fonction  $f_m$ , qui peut être linéaire ou non. De son côté, l’espérance  $E$  de la variable  $y$  est liée aux valeurs des prédicteurs  $f_m(x_m)$  à travers une fonction de lien  $g$  et une constante  $\beta_0$ . L’équation 2.4 décrit la structure générale de ces modèles.

$$g(E(y)) = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_m(x_m) \quad (2.4)$$

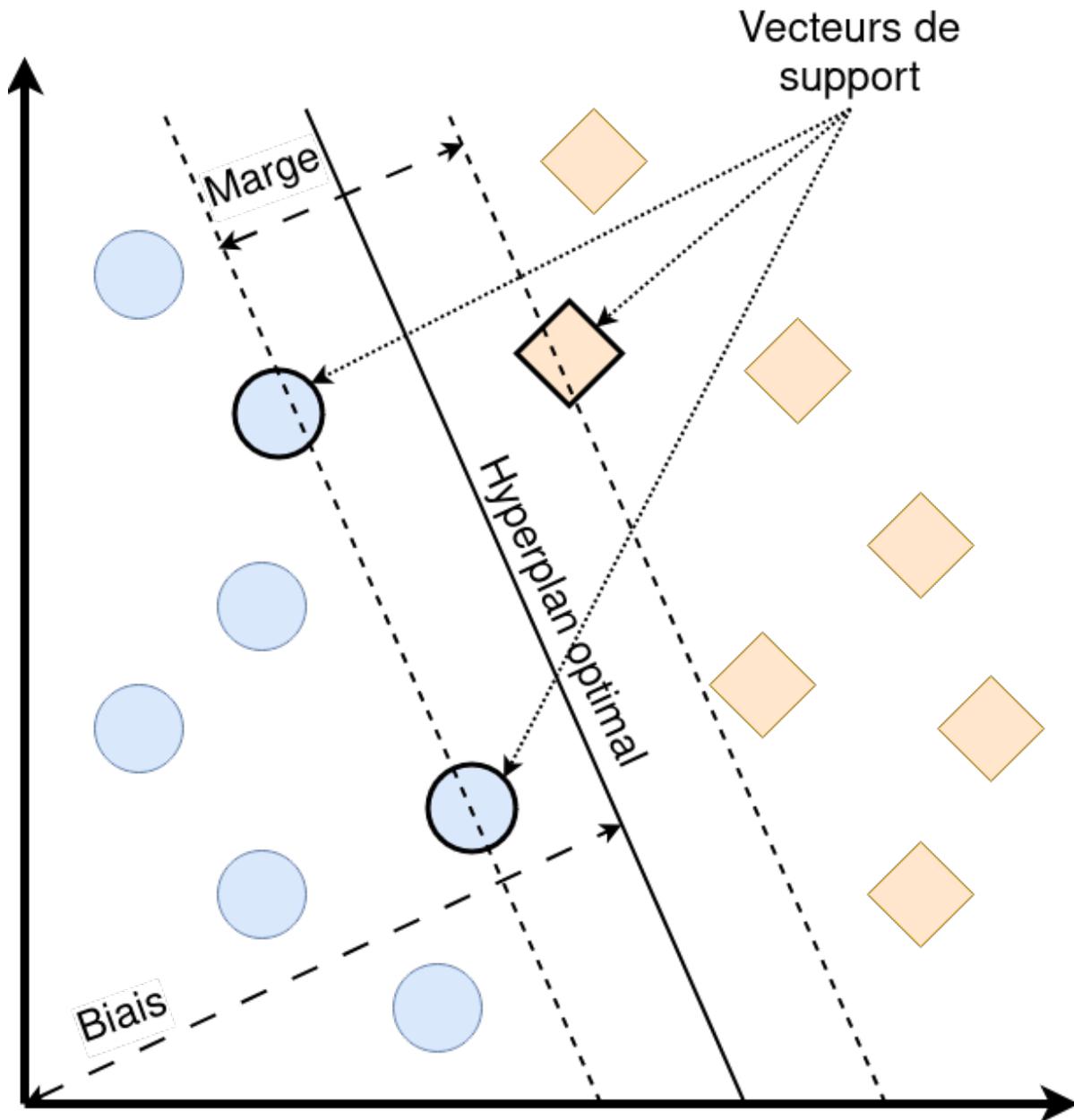
Zou et al. [21] ont proposé un modèle additif généralisé pour la prédiction de la concentration en PM<sub>2.5</sub> au niveau du sol. Ce modèle utilise un grand nombre de variables explicatives, incluant l’AOD et des paramètres météorologiques, mais aussi des variables liées à la population ou au réseau routier. Les relations entre ces variables explicatives et la concentration en PM<sub>2.5</sub>  $y$  en l’occurrence sont représentées par des splines, et la considération d’aspects linéaires et non-linéaires dans ces relations augmente la robustesse du modèle.

### 2.1.3 Apprentissage Automatique (Machine Learning) (ML)

Les techniques d’Apprentissage Automatique (Machine Learning) (ML) ont été rapidement développées [22, 23] et se sont montrées hautement efficaces pour représenter les relations non linéaires entre les PM<sub>2.5</sub> et de multiples variables [24]. Les plus populaires parmi ces méthodes sont les Machines à Vecteurs Support (SVM - Support Vector Machines) mais aussi et surtout les algorithmes basés sur des arbres de décisions, comme Random Forest (Forêt Aléatoire) ou encore l’application de méthodes de Boosting (accélération) à des arbres de décisions.

Les SVM sont des méthodes dont l’objectif général est d’identifier un hyperplan à  $N$  dimensions permettant de séparer les données qui lui sont présentées en groupes. Ces données ont elles aussi  $N$  dimensions. Cet hyperplan est situé au milieu d’une marge, représentée par la figure 2.1 issue de Unik et al. [23]. L’objectif de cette méthode est de trouver l’hyperplan optimal permettant de maximiser la taille de cette marge. Les points de données les plus proches de cet hyperplan sont appelés vecteurs supports, d’où le nom de la méthode.

L’algorithme Random Forest utilise les concepts d’Arbre de Décision et d’Ensemble Learning. Il est décrit plus en détail dans la section 2.2.2. La méthode d’ensemble learning appliquée dans ce cas appartient à la catégorie du bagging. Cela signifie que plusieurs classificateurs sont entraînés en parallèle, et un modèle final est obtenu à partir de cet ensemble.



**Figure 2.1:** Illustration du concept de la méthode des SVM, et de l'hyperplan optimal que la méthode cherche à obtenir. Source : modifié à partir de [23].

Jin et al. [25] proposent l’utilisation d’un algorithme Random Forest non pas pour estimer directement la concentration en PM<sub>2.5</sub> au niveau du sol, mais pour l’estimation d’un autre paramètre physique complexe. Celui-ci s’avère nécessaire pour l’utilisation d’une méthode basée sur des modèles et équations physiques. Cette seconde méthode permet à son tour l’estimation de la concentration en PM<sub>2.5</sub>. Il s’agit donc là d’une méthode hybride entre ML et physique. Il est possible d’y voir là une variante des méthodes d’apprentissage informé par la physique décrite en section 2.4.2, à ceci près qu’aucun réseau de neurones n’est utilisé par les auteurs de [25].

**Le Boosting** est également une méthode d’ensemble learning, permettant donc la construction d’un classifieur final à partir d’un ensemble de modèles individuellement moins performants. Elle est souvent appliquée à des arbres de décision. Ici, les modèles sont typiquement entraînés en série, les uns après les autres. La méthode la plus typique de Boosting se nomme AdaBoost. L’idée est d’entraîner un premier classifieur, et après analyse de ses résultats, des poids sont attribués aux données : les échantillons mal classifiés se voient attribuer un poids plus important, et inversement. Un nouveau modèle est entraîné sur ces données pondérées. À partir de l’analyse des résultats de ce deuxième classifieur, une nouvelle pondération peut être obtenue en appliquant le même principe. Un troisième modèle peut alors être entraîné, et ainsi de suite.

Les performances des méthodes de ML restent finalement affectée par la distribution et la densité des stations au sol utilisées pour nourrir ces dernières [9, 26].

Gupta et Christopher [9] ont introduit l’utilisation de réseaux de neurones pour estimer la distribution spatiale de particules fines. Cela ouvrit la voie à l’utilisation des techniques de Deep Learning pour répondre à ce type de problématique.

#### 2.1.4 Apprentissage Profond (Deep Learning) (DL)

D’après Zhou et al. [27], les architectures d’ANN typiquement utilisées pour l’estimation de la concentration en PM<sub>2.5</sub> sont réparties en plusieurs catégories : les Réseaux de Croyances Profondes (DBN - Deep Belief Network), les Réseaux Convolutifs (CNN - Convolutional Neural Networks), les Réseaux Récurents (RNN - Recurrent Neural Networks) ainsi que les Transformers.

**Les Deep Belief Networks (Réseau de Croyances Profondes) (DBN)** sont des algorithmes de DL typiquement utilisés pour des tâches non supervisées, telles que l’extraction de caractéristiques ou encore la réduction de la dimensionnalité. Ils sont

connus pour leur capacité à apprendre des représentations hiérarchiques complexes des données d'entrée, prenant en compte des caractéristiques de différents niveaux d'abstraction. Après entraînement non supervisé, il est possible d'affiner les DBN pour leur faire réaliser une tâche supervisée comme la classification ou la régression [27].

**Les Réseaux de Neurones Convolutif (Convolutional Neural Network) (CNN)** sont des ANN connus pour leur capacité à apprendre efficacement à partir de données multi-dimensionnelles, comme des images par exemple. Ils utilisent le concept de noyaux de convolution pour prendre en compte le voisinage de chaque point de données (ou pixel dans le cas d'une image) lors de son traitement, représentant ainsi l'aspect spatialisé de l'information. Une architecture typique des CNN pour la régression ou la translation image-image est l'encodeur-décodeur. L'idée est de réduire la dimensionnalité de la donnée dans l'encodeur puis de la réaugmenter, souvent de manière symétrique, dans le décodeur. Li et al. [28] ont proposé un modèle d'encodeur-décodeur avec connexions résiduelles pour l'estimation de la concentration en PM<sub>2.5</sub>, ce qui permet d'exploiter des représentations de plusieurs niveaux d'abstraction. Les réseaux convolutifs à graphe (GCNN - Graph Convolutional Neural Networks) sont une variante des CNN permettant la représentation de corrélations spatio-temporelles entre les données via des graphes, et sont eux aussi utilisés pour l'estimation de la concentration en PM<sub>2.5</sub> [27].

**Les Réseaux de Neurones Récurrent (Recurrent Neural Network) (RNN)** sont des réseaux particulièrement efficaces pour apprendre les dépendances temporelles au sein des données. Ils fonctionnent grâce à une boucle de retour, leur permettant de conserver une mémoire du contexte historique des entrées qui leur ont été présentées, et de le prendre en compte lors du traitement de nouvelles données.

Les réseaux dit à Longue Mémoire à Court Terme (Long Short-Term Memory) (LSTM) sont un type de RNN particulier, utilisant un système de portes leur permettant d'oublier ou retenir différentes informations obtenues durant les étapes précédentes, et ainsi d'apprendre des dépendances temporelles à plus long terme.

Ces caractéristiques se révèlent utiles pour l'estimation de valeurs telles que la concentration en PM<sub>2.5</sub>, pour lesquelles l'aspect temporel est important. Certains modèles utilisent des couches de convolutions pour permettre l'apprentissage des dépendances spatiales au sein des données d'entrée, suivi de couches récurrentes comme les LSTM qui prendront alors en entrée une séquence de caractéristiques et permettront l'apprentissage des dépendances temporelles au sein de celles-ci [27].

**Les Transformers** sont un type de réseau introduit par Vaswani et al. [29], dont la spécificité réside dans son mécanisme d’auto-attention. Celui-ci lui permet d’apprendre des relations entre différentes valeurs d’une séquence d’entrée, peu importe la distance entre ces valeurs, et sans recourir à des couches récurrentes ou de convolution. Cela le rend particulièrement apte à la prise en compte de dépendances spatio-temporelles à grandes échelles. Ce modèle a été originalement proposé pour des tâches de traitement automatique du langage naturel (comme la traduction ou la génération de texte), cependant de nombreuses variantes ont été proposées, notamment pour l’adapter au domaine de la Vision Artificielle, et ce avec succès comme démontré par Yuan et al. [30].

En ce qui concerne l’estimation de la concentration en PM<sub>2.5</sub>, des variantes de Transformers optimisés pour cette tâche ont été proposés, et ont démontré leur performance [31].

La performance des Réseaux de Neurones Artificiels (Artificial Neural Network) (ANN) est en augmentation permanente, en particulier les modèles optimisés pour une région particulière [32]. Cependant ils n’aident pas nécessairement à comprendre l’importance relative des différents paramètres d’entrée sur la décision finale. Une critique persistante des ANN parmi la communauté des physiciens est en effet que leur efficacité a souvent pour prix leur tendance à obstruer la compréhension physique, en particulier parce que ces réseaux ont tendance à se reposer sur des architectures de plus en plus complexes.

De manière peu surprenante, l’intérêt général croissant pour le domaine de l’Intelligence Artificielle eXplicable (eXplainable Artificial Intelligence) (XAI) est également retrouvé dans le domaine scientifique [33], y compris en sciences atmosphériques. En effet l’utilisation de techniques d’apprentissage profond amène des changements de paradigme en modélisation de l’atmosphère. L’étude de Park et al. [34] propose une réponse à ce phénomène. Elle propose une approche intéressante pour l’évaluation de la sensibilité des modèles à différentes variables d’entrée, en utilisant la propagation de pertinence couche-par-couche [35]. Mais cette étude reste une exception dans l’océan de modèles de la concentration en PM<sub>2.5</sub>.

Combiner les avantages en termes d’interprétabilité de modèles physiques avec les forces des méthodes de ML, est une voie intéressante pour l’obtention de valeurs précises de concentration en PM<sub>2.5</sub> à partir d’observations satellites, comme illustré par Jin et al. [25].

A partir de cette analyse, un ensemble restreint de méthodes peut être sélectionné pour servir de référence lors de l’évaluation de nos propres modèles.

## 2.2 Méthodes de base pour comparaison

Les jeux de données utilisés dans le cadre de cette thèse ne sont pas aussi populaires au sein de la communauté de la Vision Artificielle que peuvent l’être MNIST [36], ImageNet [37] ou encore COCO [38]. Par conséquent, il est difficile de trouver dans la littérature des référentiels auxquels se comparer sur ce problème et avec ces jeux de données.

Il est nécessaire de décider d’un petit nombre de méthodes à implémenter pour servir de solution de départ. Cela permettra ainsi d’y comparer les méthodes proposées durant cette thèse. Trois méthodes sont retenues pour constituer cette base de comparaison: l’interpolation polynômiale (de degré 3), l’algorithme Random Forest (dans sa version applicable pour la régression), la méthode du krigeage, et enfin des modèles d’encodeur-décodeur avec connexions résiduelles. Ces méthodes ont été choisies après une analyse minutieuse des solutions typiquement utilisées dans la communauté de physique atmosphérique, proposée en section 2.1. Elles sont choisies entre autres pour leur adaptabilité au problème. Elles sont par ailleurs représentatives de la diversité de ces solutions. Nous allons dans cette section décrire les quatre méthodes retenues.

### 2.2.1 Interpolation polynômiale

L’idée derrière cette méthode est d’approximer une fonction par un polynôme. En l’occurrence, il s’agit d’approximer la fonction prenant des valeurs d’AOD en entrée et produisant des valeurs de concentration en aérosols en sortie.

Plus précisément, à partir des données disponibles, des points de coordonnées  $(a_i, c_i)$  sont placés, avec  $a_i$  une valeur d’AOD pour un pixel donné et  $c_i$  la valeur de concentration en aérosols pour ce même pixel. L’existence d’une fonction  $F$  telle que pour tout  $i$  dans le domaine considéré,  $F(a_i) = c_i$  est supposée. Pour trouver une bonne estimation polynômiale de  $F$ , il faut alors chercher le polynôme  $P$  passant par un nombre maximal de points  $(a_i, c_i)$ . Ce polynôme  $P$  est ensuite utilisé pour produire des valeurs de concentration  $c_i$  à partir des valeurs d’AOD  $a_i$ .

Il est important de noter que cette méthode est appliquée pixel par pixel, et ne prend

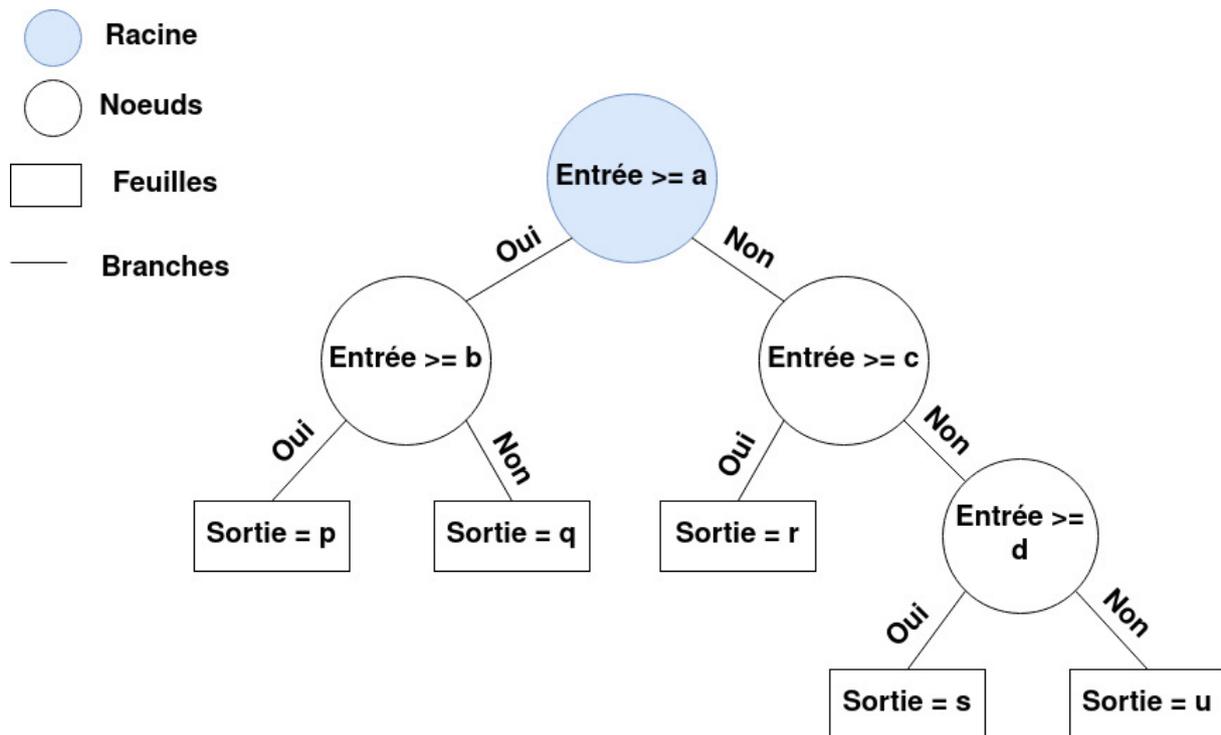


Figure 2.2: Illustration du concept d'arbre de décisions.

donc pas en compte le voisinage local des valeurs.

### 2.2.2 Algorithme Random Forest

Les algorithmes Random Forest, proposés par Ho [39] sont basés sur le principe des Arbres de Décisions et celui de l'Ensemble Learning. Le concept de l'Arbre de Décision est illustré par la figure 2.2. À chaque inférence, la variable d'entrée passe par un premier test logique, celui de la racine. En fonction du résultat obtenu, le chemin de gauche ou de droite est alors emprunté. La branche empruntée aboutit à un noeud, où un nouveau test logique est présenté. De la même manière, selon le résultat obtenu, le chemin de gauche ou de droite est suivi. Ce processus est répété jusqu'à arriver à une feuille, qui correspond à la sortie du modèle. La feuille représente la "décision" prise par cet arbre. Ce concept peut être appliqué pour réaliser de la régression comme de la classification. L'apprentissage de ce type de modèle passe par un calcul de gain d'information pour chaque test réalisé, afin de déterminer quels tests sont les plus pertinents dans ce processus de prise de décision.

Le principe de l'Ensemble Learning consiste à entraîner plusieurs modèles, et à utiliser comme résultat final la moyenne des résultats proposés par ceux-ci. Cela

repose sur l’hypothèse que les erreurs des modèles sont indépendantes et différentes les unes des autres. Dans le cas de l’algorithme Random Forest, les modèles utilisés sont des arbres (d’où le terme de Forêt, constituée de plusieurs arbres). Chacun de ces arbres est entraîné sur un sous-ensemble différent du jeu de données. Pour chaque inférence, les réponses de tous les arbres sont moyennées afin de fournir un résultat final. Cela permet la construction d’un modèle robuste à partir de plusieurs modèles qui, individuellement, le sont moins.

Comme l’interpolation polynômiale, cet algorithme est appliqué pixel par pixel. Le voisinage local des valeurs n’est donc pas pris en compte.

### 2.2.3 Krigeage

Le krigeage est une méthode d’interpolation et extrapolation spatiale formalisée par Matheron [40]. Elle est décrite comme la méthode d’estimation optimale (au sens statistique du terme) par Gratton [41]. L’équation 2.5 décrit l’idée derrière cette méthode.

$$F(x_p) = \sum_{i=1}^m W_i \cdot F(x_i) \quad (2.5)$$

L’objectif de cette méthode est d’estimer les valeurs de la fonction  $F$  au point  $x_p$ , en utilisant les valeurs connues de  $F$  en un nombre  $m$  de points  $x_i$  environnants. Atteindre cet objectif nécessite de déterminer les poids  $W_i$  de chacun de ces points. Cela est réalisé par l’estimation de ce qui est appelé le variogramme. Celui-ci est lié aux valeurs de la variance entre deux points divisée par la distance entre ces mêmes points.

D’après Baillargeon [42], la méthode du krigeage atteint de meilleures performances avec d’importants volumes de données, et quand les valeurs à estimer suivent une distribution normale.

Comparé à d’autres méthodes discutées dans ce papier, le krigeage souffre de longs temps d’inférence, pouvant aller d’une quinzaine de secondes à deux minutes selon les paramètres utilisés. Ces temps sont longs en comparaison des quelques fractions de seconde pour les autres méthodes mises en oeuvre pour notre travail. Cela est dû au fait qu’un nouveau modèle de krigeage est construit pour chaque inférence, car de nouveaux points  $x_i$  sont utilisés, avec des valeurs de  $F$  associées.

### 2.2.4 Encodeurs-Décodeurs avec Connexions Résiduelles

Le terme d’encodeur-décodeur définit un type d’architecture de réseaux de neurones. Elle contient un encodeur, conçu de manière à réduire la dimensionnalité de la donnée d’entrée au fur et à mesure du passage de celle-ci dans ses couches successives. Ce processus aboutit à l’obtention d’une représentation latente, ou encodage, de la donnée (d’où le terme d’encodeur). Cette représentation est la donnée d’entrée du décodeur, la deuxième partie de cette architecture. Le décodeur va quant à lui réaugmenter la dimensionnalité de la représentation latente à travers le passage dans ses couches. On obtient ainsi une architecture en forme de sablier.

L’ajout de connexions résiduelles à ce type d’architecture permet de créer des connexions entre deux couches non successives. Par exemple, entre la première et la dernière couche. Cela permet notamment de préserver les informations les plus superficielles, typiquement retrouvées dans les premières couches du réseau. Ronneberger, Fischer et Brox [43] ont montré l’efficacité de modèles d’encodeur-décodeur convolutifs avec connexions résiduelles (alors appelés UNet) dans le domaine de l’analyse d’images médicales. Li et al. [28], quant à eux, proposent l’application de ce concept à des modèles convolutifs et purement linéaires (ou dits "Fully Connected") pour la prédiction de la concentration en PM<sub>2.5</sub>. Ces modèles proposés sont illustrés par la figure 2.3, issue de Li et al. [28]. Ces mêmes auteurs démontrent la performance de modèles d’encodeurs-décodeurs purement linéaires avec connexions résiduelles pour cette tâche, mais n’évaluent pas celle des modèles convolutifs qu’ils proposent.

Le modèle d’encodeur-décodeur avec connexions résiduelles utilisé dans cette thèse pour servir de référence pour l’évaluation de nos propres propositions est un modèle contenant à la fois des couches convolutives et linéaires. Il est appelé UNet en référence au travail de Ronneberger, Fischer et Brox [43] et pour des raisons de praticité. Il est illustré par la figure 3.5. Le chapitre 3 aborde la question de son architecture et démontre par l’expérience à la fois l’intérêt de l’ajout de couches linéaires et de connexions résiduelles. L’intérêt des connexions résiduelles pour la prédiction des PM<sub>2.5</sub> est également démontré par Li et al. [28]. La section A.5, quant à elle, démontre la supériorité dans notre contexte de ce modèle comparé au modèle purement linéaire proposé par Li et al. [28].

Ces méthodes vont donc servir de référence pour y comparer nos propres modèles. Le chapitre 1 ayant mis en lumière le besoin de méthodes d’apprentissage semi-supervisé, nous allons à présent nous intéresser à l’état de l’art de ce type de méthodes.

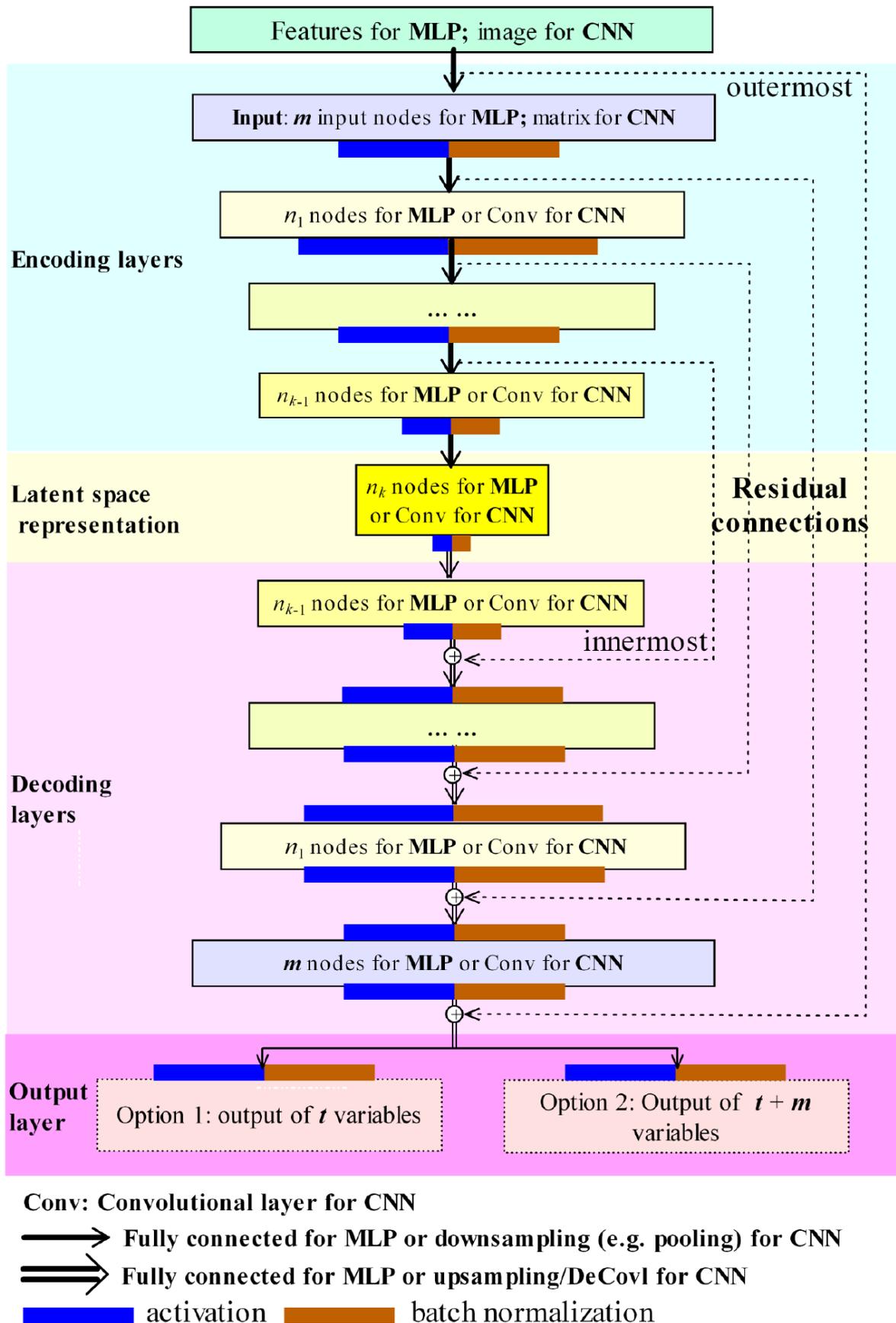


Figure 2.3: Schéma d’un encodeur-décodeur avec connexions résiduelles tel que proposé par les auteurs de [28]. Source : [28].

## 2.3 Translation image-image semi-supervisée

Comme mis en avant dans le chapitre 1, dans une situation réelle, l’ensemble d’apprentissage disponible contiendrait à la fois des données labellisées et non labellisées. Or l’idée derrière l’apprentissage semi-supervisé est justement d’utiliser des données labellisées et non labellisées durant la phase d’entraînement d’un modèle, que cela soit dans le but d’améliorer la performance, ou à cause d’un manque de données étiquetées.

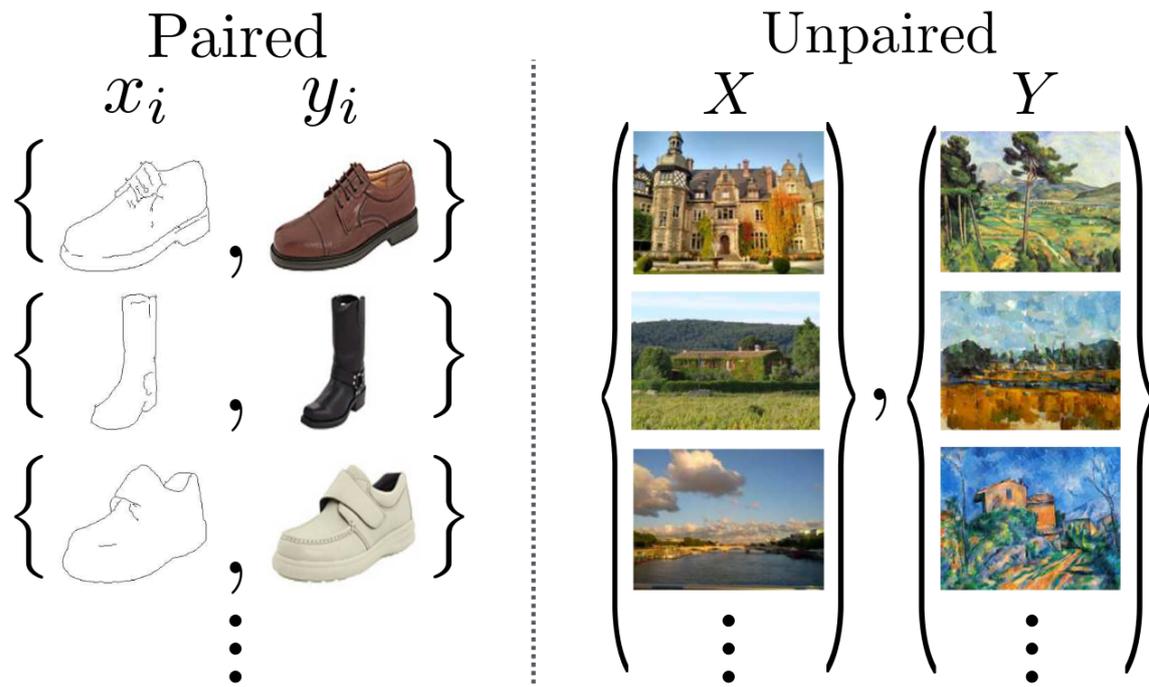
Une part importante des travaux sur l’apprentissage semi-supervisé se concentre principalement sur la classification semi-supervisée, selon Engelen et Hoos [44]. La tâche considérée dans ces travaux de thèse peut être vue comme de la translation image-vers-image. Les travaux les plus récents sur la réalisation de ce type de tâche dans un contexte semi-supervisé reposent souvent sur les Réseaux Génératifs Adverses (Generative Adversarial Networks) (GAN) ou sur les Auto-Encodeurs Variationnels (Variational Autoencoders) (VAE) [45, 46, 44, 47].

D’autres méthodes encore reposent sur des informations a priori fournies au modèle, comme c’est le cas pour les Réseaux Informés par la Physique (Physics-Informed Neural Network) (PINN) [48, 49]. Ce type d’ANN est décrit plus précisément dans la section 2.4.2.

Dans la littérature, on distingue la translation d’image appairée et non appairée. Le terme de translation appairée signifie que le jeu de données contient pour chaque image d’entrée un exemple de son équivalent après translation, autrement dit son équivalent dans le domaine cible [50]. Il contient donc des couples image source / image cible. Le terme de translation non appairée en revanche signifie que le jeu de données contient des échantillons d’images à la fois dans les domaines source et cible, mais ces images ne sont pas nécessairement équivalentes l’une à l’autre [51]. Le jeu de données ne contient pas forcément de couples image source / image cible. La tâche considérée dans ce papier est précisément une tâche de translation appairée. La figure 2.4 illustre la différence entre le contexte d’une translation appairée et non appairée.

### 2.3.1 Réseaux Génératifs Adverses (GANs)

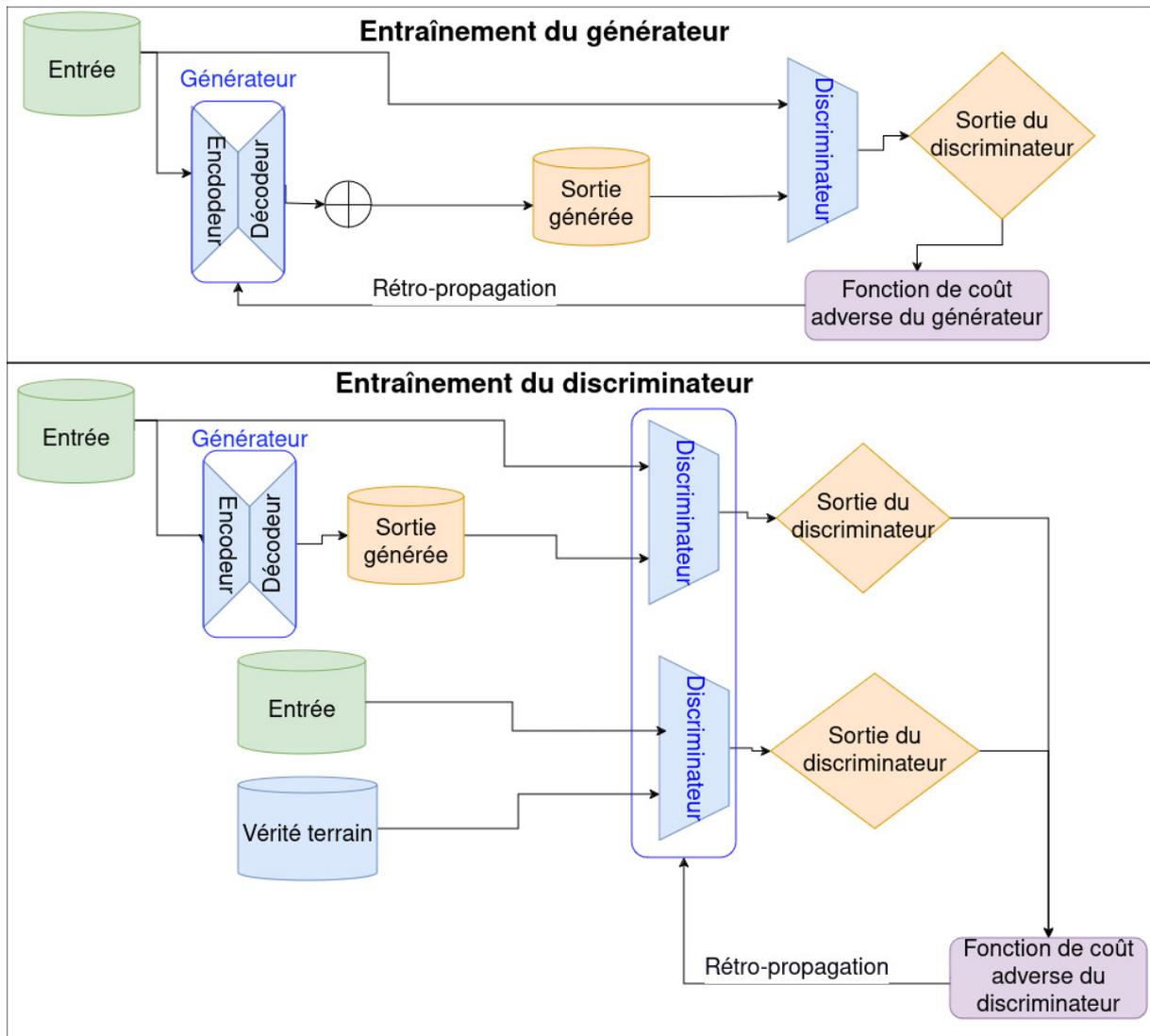
Les GANs n’ont pas cessé de gagner en popularité depuis leur introduction par Goodfellow et al. [52]. L’idée derrière ce type de modèle est d’entraîner simultanément un réseau générateur et un réseau discriminateur.



**Figure 2.4:** Représentation d'ensembles de données pour une tâche de translation appairée (à gauche) et non appairée (à droite). Dans le cas de la translation appairée, chaque échantillon de l'ensemble contient deux images. L'une est l'image source ( $x_i$ ) et l'autre l'image cible ( $y_i$ ) qui lui correspond directement. Dans le cas de la translation non appairée, deux ensembles distincts sont disponibles : l'ensemble source ( $X$ ) et l'ensemble cible ( $Y$ ). Ces ensembles ne sont pas alignés : il n'y a pas de correspondance entre les éléments de  $X$  et ceux de  $Y$ . Source de l'image : [51].

Le but du discriminateur est d'être capable de déterminer, lorsque des données lui sont présentées, si elles sont réelles (en d'autres termes, extraites de la distribution de données originale), ou s'il s'agit d'une sortie produite par le générateur. De son côté, le but du générateur est de produire des données capables de tromper le discriminateur. Autrement dit, de faire en sorte que celui-ci se trompe et classe les sorties du générateur comme étant réelles. Ainsi, plus la fonction de coût du générateur est basse, plus celle du discriminateur est élevée, et inversement, d'où la nature adverse de ce type de réseaux. La figure 2.5 décrit un processus d'entraînement d'un GAN, pour une tâche de translation image-image appairée semi-supervisée. Ce processus est similaire à celui proposé par Isola et al. [50], qui utilisent un GAN dit "conditionnel" pour réaliser ce type de tâche.

Le GAN original souffre de problèmes de stabilité. Un modèle plus récent, appelé GAN Wasserstein avec Pénalités de Gradient (ou WGAN-GP) a été proposé par



**Figure 2.5:** Schéma d'un entraînement d'un GAN pour une tâche de translation image-image appairée semi-supervisée. Les vérités terrain ne sont nécessaires que pour l'entraînement du discriminateur. Celui-ci reçoit tour-à-tour des exemples réels (entrée-vérité terrain) et non réels (entrée-sortie générée) qu'il doit être capable de différencier. Le générateur est un encodeur-décodeur.

Arjovsky, Chintala et Bottou [53] et Gulrajani et al. [54] pour résoudre ce problème.

La popularité des GANs dans le domaine de la Vision Artificielle est dûe en particulier à leur capacité à générer des images réalistes. Ils ont aussi été prouvés efficaces pour la réalisation d’autres tâches, telles que la translation image-vers-image [55, 56, 57].

Les GANs classiques souffrent du même problème qu’un grand nombre de modèles d’apprentissage profond. En effet, ce sont ce qu’on appelle des boîtes noires. Il est difficile de comprendre et d’expliquer sur quoi le modèle base ses décisions. Un domaine de l’IA tente de s’éloigner de cet aspect boîte noire : il s’agit du domaine de l’Intelligence Artificielle eXplicable.

### 2.3.2 Auto-Encodeurs Variationnels (VAEs)

Les Auto-Encodeurs Variationnels (Variational Autoencoders) (VAE) sont des modèles introduits par Kingma et Welling [58], initialement conçus pour réaliser de l’apprentissage non supervisé.

Nous évoquons ici sa version utilisable en pratique, qui contient une astuce de reparamétrisation. La figure 2.6 illustre le principe de fonctionnement d’un VAE. Ainsi ce type de modèle contient un encodeur, qui prend une image  $x$  en entrée et produit deux vecteurs, appelés  $\mu$  et  $\sigma$ . Ces deux vecteurs ainsi qu’un vecteur aléatoire  $\epsilon$  sont utilisés pour construire le vecteur  $z$  en suivant la formule décrite par l’équation 2.6. L’opérateur  $\cdot$  représente le produit de Hadamard (produit de deux vecteurs élément-par-élément).

$$z = \mu + \sigma \cdot \epsilon \quad (2.6)$$

Par la suite ce vecteur  $z$  est utilisé comme entrée par un décodeur, qui produit en sortie une image  $x'$ . Kingma et Welling [58] proposent alors l’utilisation d’une fonction de coût comme l’erreur quadratique moyenne ou l’entropie croisée pour la reconstruction, autrement dit pour s’assurer que  $x'$  est aussi proche que possible de  $x$ . Une fonction de coût dite de régularisation utilisant les vecteurs  $\mu$  et  $\sigma$  est également utilisée, il s’agit de la divergence de Kullback Leibler.

Dans le cadre de la translation d’images semi-supervisée, ce type de modèle est souvent utilisé en couplage avec un modèle servant de discriminateur [47, 59], le VAE remplaçant alors le générateur d’un GAN. L’idée derrière ce type de travaux est alors

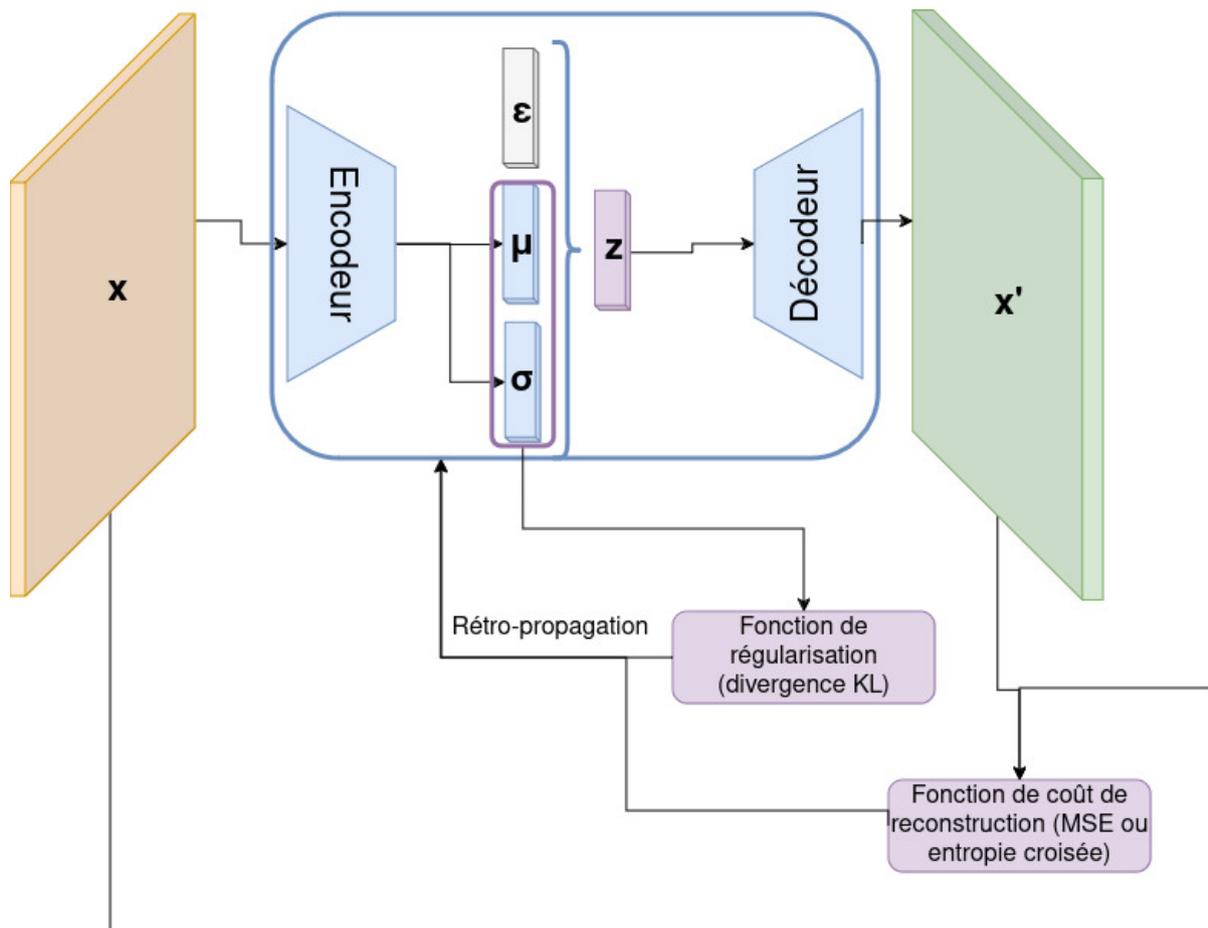


Figure 2.6: Principe de fonctionnement d'un VAE, tel que décrit dans [58]

d'entraîner deux VAEs, un pour chaque domaine (source et cible), en prenant soin d'obtenir des représentations latentes issues du même espace pour ces deux modèles. Une fois cette première phase terminée, il est possible d'utiliser l'encodeur du premier modèle avec le décodeur du deuxième ou inversement, ce qui permet alors de passer d'un domaine à l'autre dans le sens voulu.

Le discriminateur et sa fonction de coût adverse ont pour rôle principal d'assurer que les images générées par le VAE sont aussi proches que possible des images réelles. Le discriminateur permet donc de continuer l'apprentissage du VAE dans un contexte semi-supervisé.

La figure 2.7 représente le processus d'entraînement d'un VAE utilisant un discriminateur.

L'un des intérêts d'un modèle tel que le VAE-GAN pour la translation image-vers-

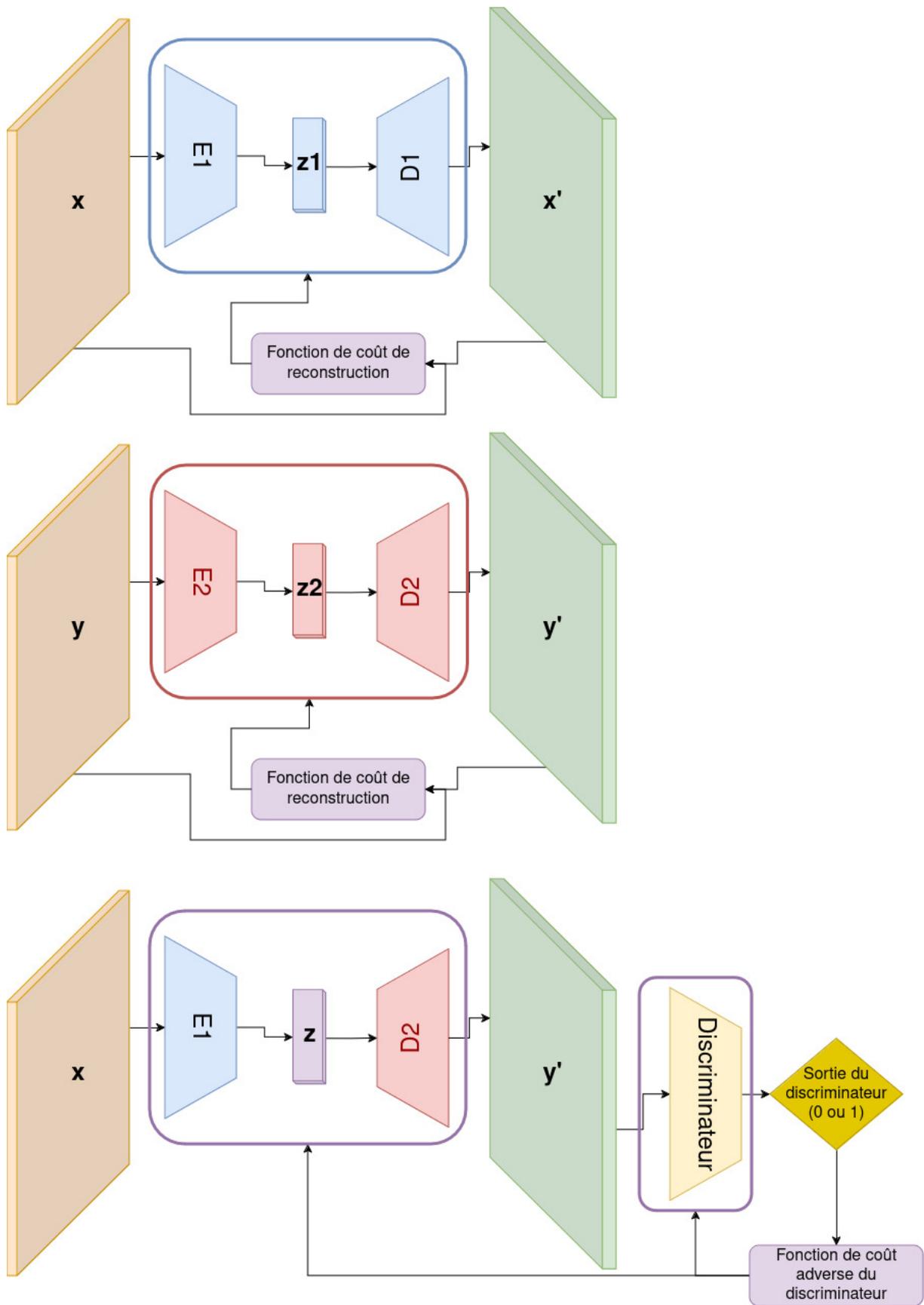


Figure 2.7: Principe d'apprentissage d'un modèle utilisant deux VAEs et un discriminateur.

image semi-supervisée est donc qu’il permet de réaliser cette translation dans les deux sens : de l’entrée vers la vérité terrain, et inversement. Cependant cette caractéristique n’est pas recherchée dans notre contexte, et elle induit une augmentation du nombre de paramètres du modèle global.

De plus cette augmentation de la complexité du modèle accentuerait son aspect boîte noire, et diminuerait donc son acceptabilité. Or nos modèles sont destinés à être utilisés par la communauté de sciences atmosphériques. Il nous faut donc accorder une importance particulière aux méthodes plus susceptibles d’être acceptables par cette dernière.

## 2.4 Acceptabilité des modèles

Dans le domaine de l’optique atmosphérique, comme dans tout domaine scientifique, l’interprétabilité d’une solution proposée revêt une importance capitale. Elle permet en effet de tirer des conclusions à partir des résultats, et par là même d’améliorer la connaissance du domaine. Cependant l’aspect boîte noire d’un grand nombre d’ANN diminue justement leur explicabilité. Cette section s’intéresse donc aux modèles de DL s’éloignant de cet aspect boîte noire, ce qui les rend plus acceptables, et attractifs, aux yeux de la communauté de physique atmosphérique.

### 2.4.1 Modèles Explicables

Un domaine de recherche en IA a spécifiquement pour but la conception de modèles plus interprétables, et plus facilement acceptables par leurs utilisateurs. Il s’agit de l’Intelligence Artificielle eXplicable (eXplainable Artificial Intelligence) (XAI).

L’idée générale derrière ce domaine est de concevoir des architectures dont la conception est explicable, ou dont les résultats et leur qualité sont interprétables par les utilisateurs. Ce domaine existe depuis plusieurs décennies [60], cependant la récente croissance de sa popularité peut être vue comme une réponse à des préoccupations légitimes sur l’aspect boîte noire d’un nombre importants de réseaux de neurones.

Ce gain de popularité est particulièrement remarquable dans des domaines d’applications tels que la finance, la médecine, la loi et même la production scientifique [33, 61, 62, 63]. Dans ces domaines, la capacité à expliquer un modèle et ses résultats peut représenter sa capacité à assurer une certaine sécurité, équité ou rigueur scientifique. De manière plus générale, cela permet à l’utilisateur de faire plus facilement confiance au modèle.

Belle et Papantonis [62] présentent clairement l’intérêt que peut représenter ce type de démarche pour l’utilisateur final.

Roscher et al. [63], dans le contexte de l’XAI, il y a trois éléments importants à considérer lors de l’évaluation de l’explicabilité (au sens général) d’un modèle.

1. **Transparence** : un modèle est transparent si les processus permettant l’extraction des paramètres du modèle à partir des données d’entraînement et ceux permettant la production de sorties à partir des données de test peuvent être décrits et motivés par le concepteur de ce modèle. Il existe différents niveaux de transparence selon les algorithmes, et il est irréaliste d’espérer concevoir une méthode d’apprentissage transparente sous tous aspects. Par exemple, bien qu’un réseau de neurones donné puisse être entièrement décrit mathématiquement, le choix de son architecture et de ses hyper paramètres n’est que rarement motivé par la connaissance du domaine d’application.
2. **Interprétabilité** : il s’agit de la capacité à comprendre de manière générale ce sur quoi le modèle base ses décisions. Certaines approches pour la conception de modèles interprétables sont basées sur des arbres de décisions, car leur structure permet d’illustrer de manière intuitive le processus de prise de décision d’un modèle. Parmi les méthodes utilisées pour augmenter l’interprétabilité d’un ANN, il existe notamment des méthodes dites *post-hoc*, qui consistent à inspecter la manière dont la sortie d’un modèle répond à la modification de certaines de ses entrées. Des scores de pertinence, de sensibilité ou d’importance de différentes caractéristiques peuvent alors être calculés et visualisés par des cartes thermiques, pour permettre l’identification de motifs.
3. **Explicabilité** : une explication est la collection de caractéristiques du domaine interprétable qui ont contribué, pour un exemple donné, à produire une décision. Pour qu’un modèle soit explicable, il doit en général être possible de comprendre pourquoi la décision du modèle dans le cas A est différente de celle dans le cas B. L’un des principaux obstacles à l’application de ce concept à un ANN est que ce type d’algorithme ne contient, de manière inhérente, aucune représentation explicite de concepts symboliques comme des objets ou événements. C’est pour cette raison que différents auteurs, comme ceux de [64], proposent d’investiguer des modèles hybrides entre l’IA symbolique et connectionniste (i.e. les ANN), afin d’exploiter les forces de ces deux domaines.

Selon les mêmes auteurs, la connaissance du domaine d’application, et l’intégration de cette connaissance dans le modèle, peut permettre d’augmenter son explicabilité générale. Les modèles utilisant ce type de connaissance sont appelés modèles informés.

De plus, si le domaine d’application est un domaine scientifique, alors il est possible d’augmenter la cohérence du modèle avec les connaissances scientifiques, ce qui produit le même type d’effets sur son explicabilité générale. Cette dernière idée semble particulièrement adaptée à notre contexte d’utilisation.

### 2.4.2 Apprentissage Informé par la Physique

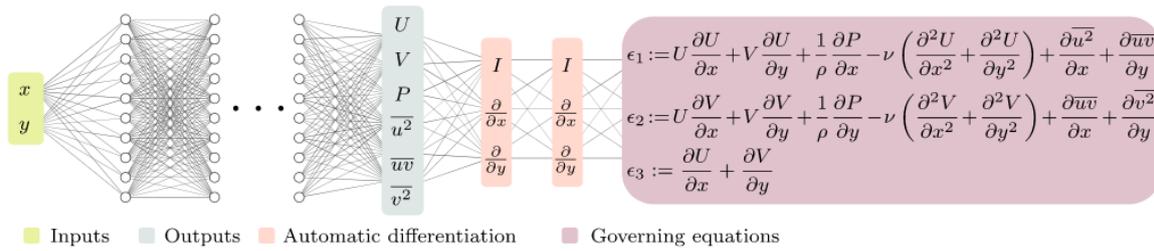
Il est intéressant de noter que la connaissance du domaine d’application peut être utilisée pour augmenter l’explicabilité d’un modèle [33]. Dans ce sens, les Réseaux Informés par la Physique (Physics-Informed Neural Network) (PINN) peuvent être vus comme une porte d’entrée vers l’augmentation de l’acceptabilité d’un modèle, dans la mesure où ils exploitent la connaissance de la physique et de la géométrie du problème considéré.

L’introduction de ce type d’ANN par Raissi, Perdikaris et Karniadakis [48] a donné naissance à un nouveau domaine de recherche en lui-même. L’idée principale derrière ce type de réseaux est d’utiliser des informations a priori sur la physique et la géométrie du problème dans le but de guider l’apprentissage.

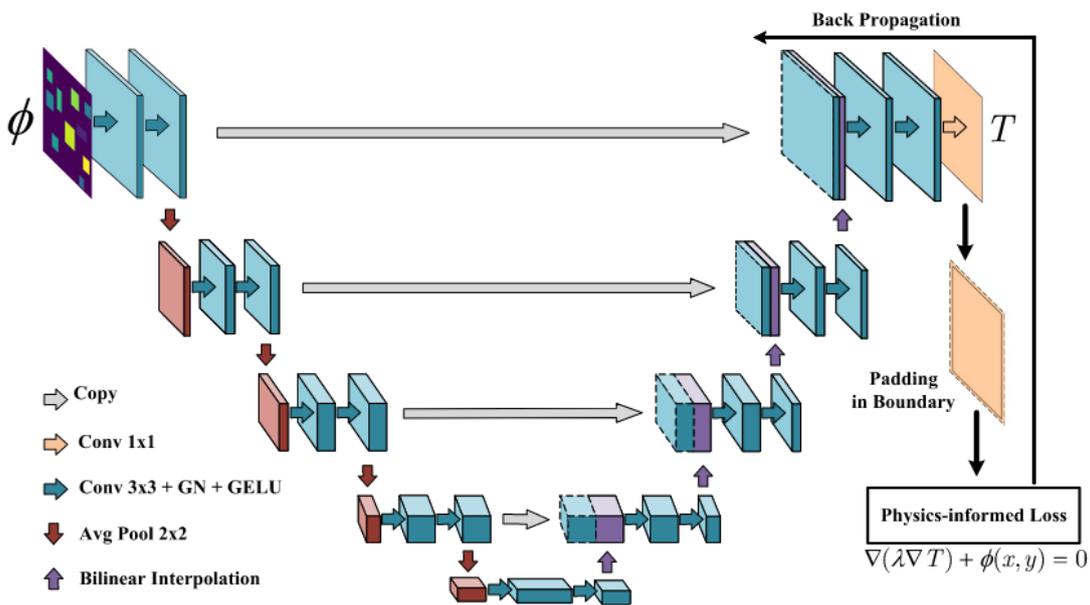
Cela peut typiquement être fait à travers la conception d’une fonction de coût qui quantifie essentiellement la cohérence (au regard de la physique) des sorties du modèle. En général, cette fonction de coût correspond à une équation (souvent différentielle) décrivant une relation entre les entrées et sorties du modèle, et qui doit être respectée par ces dernières. La figure 2.8, issue de Eivazi et al. [65], représente un PINN conçu justement de cette manière. Il s’agit là d’un exemple de PINN utilisant un Perceptron Multi-Couches.

D’autres travaux [66, 67, 68] proposent d’exploiter à la fois les capacités des PINNs et celles des Réseaux de Neurones Convolutif (Convolutional Neural Network) (CNN), afin de concevoir des PI-CNNs, plus adaptés à l’utilisation de données sous forme d’images, ou plus généralement de matrices. La figure 2.9, issue de Zhao et al. [68], représente un modèle de ce type dont l’architecture est basée sur le principe du UNet [43].

Une bibliothèque entière a été proposée par Lu et al. [69], avec pour but de pro-



**Figure 2.8:** Schéma du fonctionnement d’un PINN utilisant les équations de Navier-Stokes (moyennées par le Reynolds) pour un flux incompressible turbulent. On retrouve trois composantes de ces équations à droite du schéma : elles sont utilisées en tant que fonctions de coût à minimiser. Le modèle est un Perceptron Multi-Couches. Source : [65].



**Figure 2.9:** Schéma d’un PI-CNN, avec une architecture basée sur le UNet. Ce modèle est conçu pour la prédiction champs de température  $T$ , à partir de  $\phi$  une fonction de distribution d’intensité corrélée avec la disposition de sources de chaleur. Source : [68].

poser des solutions facilement implémentables basées sur les PINNs pour résoudre différentes sortes d’équations différentielles (telles que des équations différentielles partielles, des équations intégro-différentielles...).

Les PINNs sont en général particulièrement utiles dans le contexte de l’apprentissage semi-supervisé. Ils permettent de guider l’apprentissage d’une manière moins centrée sur les données, et de réduire ainsi le besoin en données labellisées.

La littérature sur les PINNs propose deux méthodes principales pour assurer le respect des Conditions Aux Limites et Initiales (CAL/I), et exploiter les informa-

tions qu’elles contiennent durant l’entraînement. La première s’appelle la contrainte douce, ou méthode douce, et consiste en la conception d’une fonction de coût afin d’apprendre au modèle à respecter ces CAL/I. La seconde s’appelle la contrainte (ou méthode) dure, et consiste à appliquer une transformation directement sur les sorties du modèle pour le forcer à respecter ces CAL/I. Sun et al. [49] montrent dans leurs travaux la supériorité de la méthode dure dans le cas des PINNs : elle facilite en effet davantage la convergence.

Certains modèles exploitant les forces à la fois de l’apprentissage adverse et informé par la physique ont été proposés [70, 71]. Les plus notables sont des modèles appelés Réseaux Génératifs Adverses Informés par la Physique (Physics-Informed Generative Adversarial Network) (PI-GAN) [72, 73].

La figure 2.10, tirée de Yang, Zhang et Karniadakis [73], illustre un exemple de PI-GAN. Dans cet exemple, l’équation physique utilisée pour concevoir ce PI-GAN est l’Équation Différentielle Stochastique (EDS) (voir Equation 2.7). Cette équation est très générale et peut être utilisée pour décrire différents phénomènes. Dans cette équation,  $x$  et  $\omega$  représentent des coordonnées qui sont utilisées comme donnée d’entrée par le modèle.

$$\begin{aligned} N_x[u(x;\omega);k(x;\omega)] &= f(x;\omega) & x \in D, \omega \in \Omega \\ B_x[u(x;\omega)] &= b(x;\omega) & x \in \Gamma \end{aligned} \quad (2.7)$$

**Equation 2.7:** EDS décrite dans [73].  $x$  est une coordonnée spatiale à  $d$  dimensions, appartenant au domaine spatial  $D$  inclus dans  $\mathbb{R}^d$ .  $\omega$  est un évènement aléatoire appartenant à l’espace des probabilités  $\Omega$ .  $N_x$  représente un opérateur différentiel général, et  $B_x$  l’opérateur de Conditions Aux Limites (CAL), agissant sur l’espace  $\Gamma$  représentant les limites du domaine  $D$ . La fonction  $b(x;\omega)$  participe également à la définition des CAL.  $k(x;\omega)$  est un coefficient et  $f(x;\omega)$  un terme de forçage, tous deux modélisés comme des processus aléatoires. Enfin,  $u(x;\omega)$  est la solution recherchée, qui dépend donc de  $k(x;\omega)$  et  $f(x;\omega)$ .

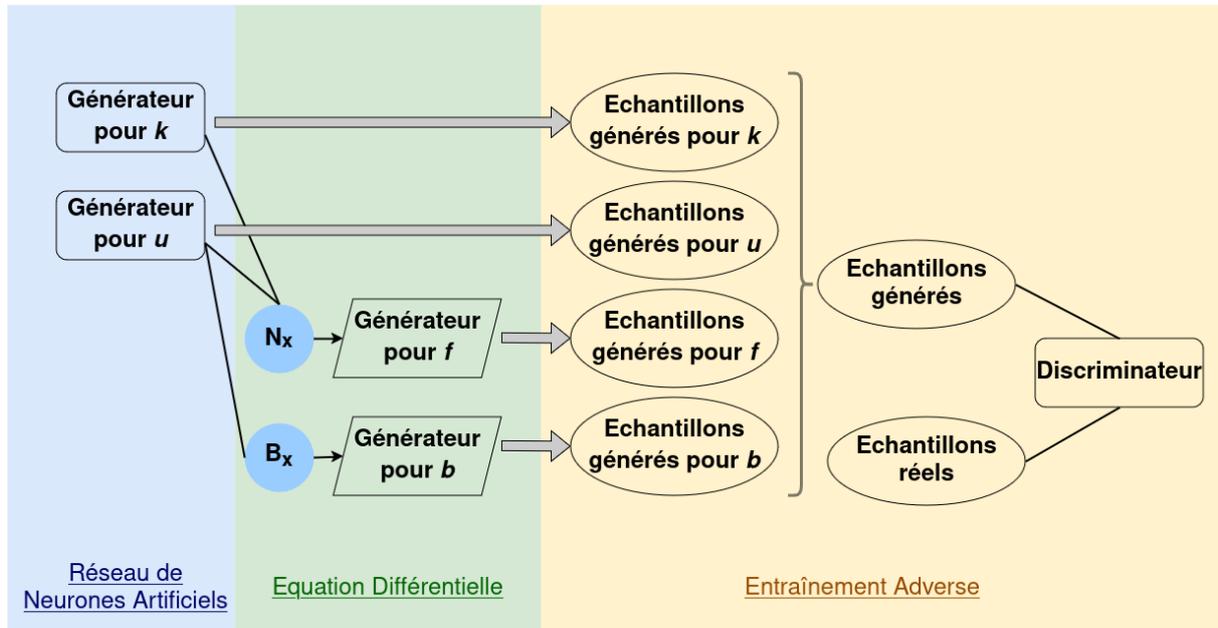
L’équation 2.8 propose un exemple plus spécifique du concept représenté par l’équation 2.7, et conserve les mêmes notations.

Dans le modèle proposé par Yang, Zhang et Karniadakis [73], l’EDS n’est pas utilisée dans la fonction de coût, mais dans l’architecture du modèle, comme le montre la figure 2.10. Par construction, le modèle produit nécessairement des prédictions pour  $k(x;\omega)$ ,  $u(x;\omega)$ ,  $f(x;\omega)$  et  $b(x;\omega)$ . Le discriminateur détermine ensuite si les données qui lui sont présentées, sont réelles ou générées. Le reste du fonctionnement de ce modèle suit alors le fonctionnement typique d’un GAN. Les générateurs ont ainsi pour

$$-\frac{1}{10} \frac{d}{dx} \left[ k(x; \omega) \frac{d}{dx} u(x; \omega) \right] = f(x; \omega) \quad x \in [-1, 1], \omega \in \Omega \quad (2.8)$$

$$u(-1) = u(1) = 0$$

**Equation 2.8:** EDS décrite dans Yang, Zhang et Karniadakis [73], représentant un exemple spécifique de l’équation 2.7, plus générale.  $k(x; \omega)$  et  $f(x; \omega)$  sont des processus stochastiques indépendants, et  $k(x; \omega)$  est strictement positif.



**Figure 2.10:** Schéma d’un PI-GAN conçu pour la résolution d’EDS. Les générateurs pour  $k$  et  $u$  sont des ANN à couches linéaires. L’utilisation des opérateurs  $N_x$  et  $B_x$  aboutit à la production de prédictions pour  $f$  et  $b$  également. Les prédictions obtenues sont concaténées avant d’être présentées au discriminateur. Source : modifié à partir de [73].

objectif de tromper le discriminateur en produisant des sorties suffisamment réalistes pour que celui-ci les détermine comme réelles (et non générées).

## 2.5 Conclusion

Nous avons proposé dans ce chapitre 2 un aperçu des travaux de l’état de l’art pertinents vis-à-vis de notre problématique de recherche. Nous avons commencé par un inventaire des méthodes typiquement utilisées comme réponses à notre problématique au sein de la communauté de physique atmosphérique, proposé en section 2.1. Les méthodes de Deep Learning y apparaissent comme prometteuses, malgré un certain nombre de limitations dues notamment à leur manque d’interprétabilité. À la lumière de ce premier aperçu de l’état de l’art, nous avons décrit en section 2.2

trois méthodes adaptables à notre contexte et pouvant être utilisées comme baseline de comparaison, afin de répondre à l’obstacle du manque de référentiel sur cette axe de recherche. Le chapitre 1 a mis en avant un défi important relatif à notre problématique : le manque de données annotées. La section 2.3 a proposé des méthodes d’apprentissage semi-supervisées pouvant répondre à ce défi, et met notamment en avant les VAE et les GAN comme pistes prometteuses. Enfin la section 2.4 aborde le sujet de l’acceptabilité des modèles proposés par la communauté de physique atmosphérique, pour laquelle l’interprétabilité des solutions est un avantage indéniable. Nous y avons donné plusieurs définitions importantes vis-à-vis des notions d’interprétabilité et d’explicabilité, avant de décrire des pistes de réponse à ce besoin, basées sur l’apprentissage informé par la physique. À la suite de ces réflexions, nous avons pu identifier trois pistes de recherche majeures, qui sont investiguées en détail dans les chapitres suivants.

La première voie de recherche que nous avons empruntée est née d’observations sur différents modèles. Ainsi, la méthode du krigeage, présentée en section 2.2.3, utilise les vérités terrain éparses (évoquée dans le chapitre 1) comme donnée d’entrée. De plus, la littérature sur l’apprentissage informé par la physique propose des méthodes permettant d’exploiter l’information contenue dans les Conditions Aux Limites (CAL), comme décrit en section 2.4.2. Celles-ci représentent des valeurs connues (en certains points) de la quantité à estimer : des vérités terrain éparses peuvent donc en ce sens être perçues comme des CAL et utilisées par ce type de méthode.

Il est donc certain que des vérités terrain éparses représentent une information utile, et qui peut être prise en compte pour améliorer la performance d’un modèle. La première piste de recherche, explorée dans le chapitre 3, est donc la conception d’un modèle d’apprentissage semi-supervisé capable d’exploiter l’information contenue dans des vérités terrain éparses (lors de la partie non supervisée de son apprentissage).

La plupart des méthodes présentées en section 2.1 sont capables de prendre plusieurs variables en entrée. D’après Ma et al. [17], Unik et al. [23] et Xu, Zhang et Liang [18], la question des variables à utiliser pour expliquer, estimer ou prédire la concentration en PM<sub>2.5</sub> continue d’être le sujets de travaux de recherche.

De cette observation est née notre deuxième piste de recherche. Dans le chapitre 4, nous proposons une étude sur une sélection réduite de variables d’entrée, utilisées pour la prédiction de la concentration en PM<sub>2.5</sub> par différents modèles. Cette étude s’intéresse également à la question de la stratégie de fusion d’informations à employer

pour pouvoir traiter efficacement ces multiples données d’entrée.

Enfin, la fin de la section 2.1 met en avant le manque d’interprétabilité de la majorité des modèles de DL d’estimation de la concentration en PM<sub>2.5</sub>. Nous y avons également souligné l’interprétabilité importante des modèles physiques. De plus dans la section 2.4.2 nous précisons que les PINN sont adaptés à l’apprentissage semi-supervisé.

La conception d’un PINN pour la prédiction de la concentration en PM<sub>2.5</sub> à partir de multiples variables, et capable d’exploiter les vérités terrain éparses en tant que CAL, est donc l’évolution naturelle des travaux de recherche précédents. C’est donc également l’axe de recherche exploré dans le chapitre 5.

# 3

## BC-GAN : un modèle utilisant les vérités terrains éparses

### Sommaire

---

3.1	Proposition . . . . .	60
3.1.1	Contrainte dure . . . . .	61
3.1.2	Contrainte douce . . . . .	62
3.2	Protocole expérimental . . . . .	64
3.2.1	Données . . . . .	64
3.2.2	Pré-traitement des données . . . . .	65
3.2.3	Description des expériences . . . . .	66
3.3	Evaluation des modèles . . . . .	69
3.3.1	Disponibilité du code et des données . . . . .	71
3.4	Résultats . . . . .	72
3.4.1	Encodeurs-décodeurs en apprentissage purement supervisé . . . . .	72
3.4.2	Contrainte dure et contrainte douce . . . . .	73
3.4.3	Comparaison avec les méthodes proposées comme base de comparaison . . . . .	76

3.5 Conclusion . . . . . 79

---

Comme exposé dans la section 1.2, l'objectif de cette thèse est de prédire par l'apprentissage machine la concentration en particules fines proches de la surface, à partir de valeurs d'AOD intégrées sur la colonne atmosphérique. Or, dans un cas réel, seules des valeurs éparses de la concentration en PM<sub>2.5</sub> sont disponibles (voir section 1.3). Dans notre contexte, cela correspond alors à ce qu'on appelle des vérités terrains éparses.

La littérature sur les PINNs propose des méthodes pour utiliser l'information contenue dans les Conditions Aux Limites (CAL) afin de guider l'apprentissage (voir section 2.4.2). Les CAL sont généralement définies comme des restrictions limitant les solutions possibles d'une équation différentielle. Elles peuvent aussi représenter des valeurs connues de la solution ou sortie désirée. C'est justement ce que représentent les vérités terrains éparses : en ce sens, elles peuvent donc être vues comme des CAL.

La connaissance des CAL limite ainsi les potentielles sorties du modèle qui pourraient correspondre à la vérité terrain. Le terme de CAL est utilisé tout au long de ce chapitre en référence aux vérités terrain éparses, et ce bien qu'aucun système d'équations différentielles ne soit utilisé.

Les problèmes contenant ce type de CAL en tant qu'information disponible a priori ne peuvent pas toujours être résolus avec des méthodes purement basées sur la physique. Cela peut être dû au coût important de calcul des méthodes basées sur la physique, ou à l'absence d'une équation physique pertinente décrivant le problème concerné.

La prédiction de l'exposition aux champs électro-magnétiques en milieu urbain est un exemple de ce type de problème [74]. Dans ce contexte, différents capteurs et transmetteurs peuvent fournir des valeurs éparses, représentées donc par une matrice éparse. Les indices d'une telle matrice représenteraient alors la géométrie du problème, et ses valeurs des CAL.

Il y a par conséquent un réel besoin dans le domaine de la Vision Artificielle d'une méthode permettant l'utilisation de CAL durant l'entraînement en tant qu'information disponible a priori, sans utiliser d'équation physique.

Dans ce chapitre, une méthode est proposée pour permettre à des modèles d'apprentissage profond d'exploiter les CAL en tant qu'information a priori durant l'apprentissage. Ce chapitre démontre aussi dans le cas d'utilisation décrit en section 1 l'intérêt de ce type d'approche pour l'amélioration de la performance des modèles.

Les CAL peuvent être considérées avec différentes quantités et distributions de pixels dans les matrices éparses qui les représentent. Par conséquent différentes configurations de CAL sont implémentées et leur impact sur les résultats est discuté. Au vu des caractéristiques du jeu de données utilisé, l'apprentissage semi-supervisé semble être le régime d'entraînement le plus pertinent pour l'étude de telles méthodes.

Les principales contributions présentées dans ce chapitre peuvent être résumées ainsi :

- une nouvelle méthode pour l'exploitation des CAL avec des GANs semi-supervisés et non informés par la physique est proposée
- l'intérêt en termes de performance de l'exploitation des CAL pour l'entraînement d'un GAN semi-supervisé est empiriquement prouvé
- l'impact de la quantité et de la distribution des CAL sur la performance est étudié
- la méthode proposée est appliquée avec succès au cas d'utilisation décrit dans le chapitre 1

La section 3.1 détaille les contributions proposées, puis la section 3.2 décrit leur implémentation et les expériences réalisées. Une analyse des résultats obtenus est proposée en section 3.4, et enfin la section 3.5 résume les avancées réalisées grâce aux travaux réalisés dans ce chapitre.

### 3.1 Proposition

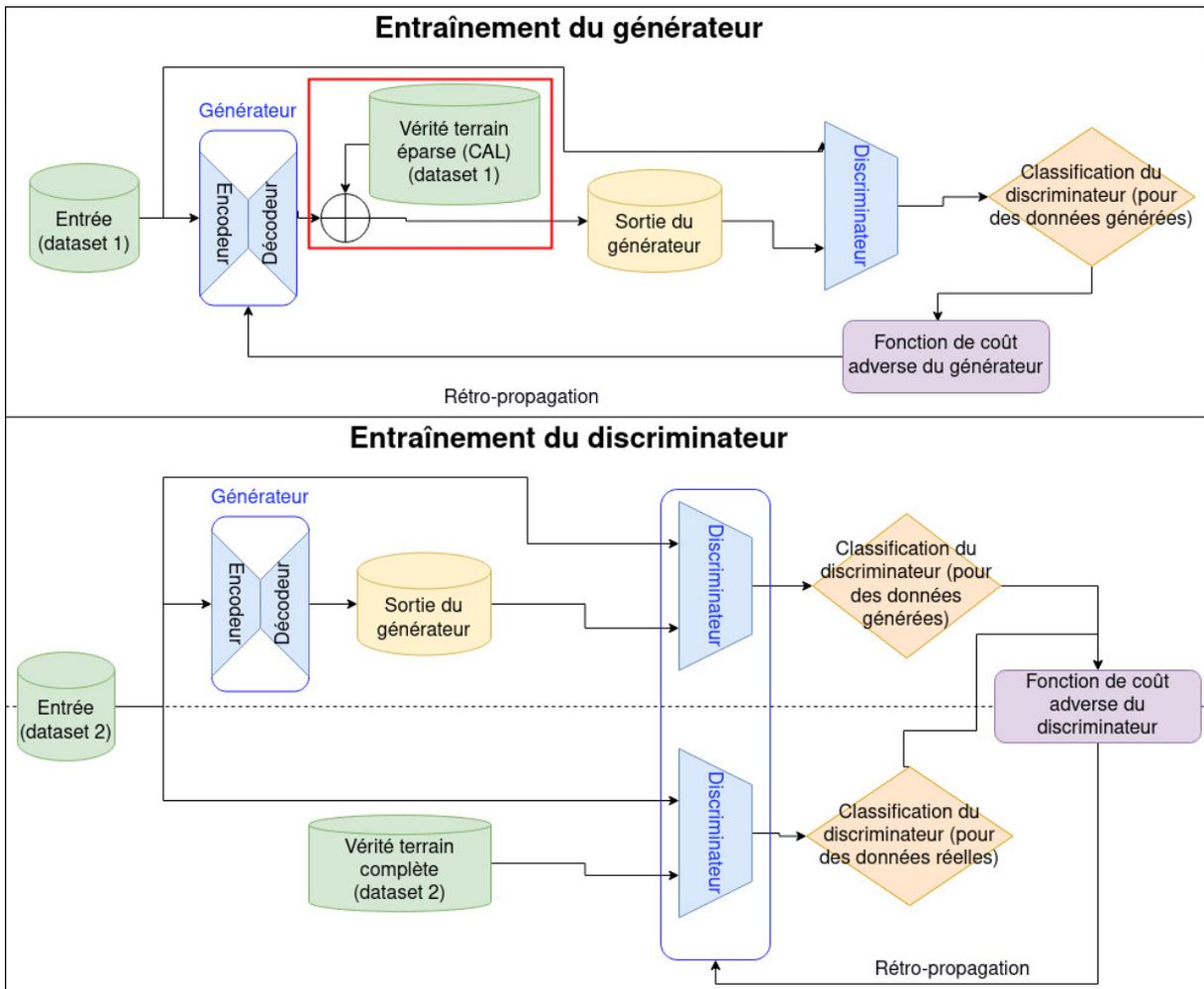
La première contribution de cette thèse concerne l'utilisation des informations contenues dans les CAL pour l'entraînement d'un GAN. Plus précisément, il s'agit de démontrer que ce type d'information, et les méthodes permettant de les exploiter, ne sont pas utiles que pour les PINNs mais également pour d'autres modèles.

Deux types de contraintes peuvent être appliquées à un PINN pour exploiter les CAL : la contrainte dure et la contrainte douce. Dans cette section, ces deux méthodes sont appliquées à un GAN, afin de comparer leur efficacité dans le contexte défini au chapitre 1. Sun et al. [49] ont montré, dans le cas des PINNs, la supériorité de la contrainte dure sur la contrainte douce. L'un des objectifs de l'étude proposée ici est également de déterminer si cela est également vrai pour les GANs. Différentes configurations de CAL sont également proposées, afin de discuter leur impact sur les

résultats. Ces configurations sont décrites en section 3.2.

### 3.1.1 Contrainte dure

L'utilisation de la contrainte dure consiste à appliquer une transformation sur les sorties du modèle afin de les forcer de manière directe à respecter les CAL. Cette méthode est illustrée par la Figure 3.1.



**Figure 3.1:** Schéma représentant l'entraînement d'un GAN, avec utilisation de la contrainte dure sur les CAL. Dans ce contexte, les CAL sont des valeurs éparsees de la concentration en aérosols au niveau du sol.

L'équation 3.1 décrit la manière dont les sorties du modèle sont forcées à respecter les CAL. Dans cette équation,  $y$  est la vérité terrain, tandis que  $\hat{y}$  et  $\hat{y}_{cal}$  représentent la sortie du modèle respectivement avant et après transformation.  $O$  est une matrice contenant exclusivement des valeurs de 1, et  $CAL$  est une matrice représen-

tant la position des valeurs éparses de concentration représentant nos CAL. Elle contient 1 aux positions pour laquelle la valeur de concentration est connue, et 0 aux autres endroits. L'opérateur  $\cdot$  représente le produit de Hadamard (produit de deux matrices élément-par-élément).

$$\hat{y}_{cal} = (O - CAL) \cdot \hat{y} + CAL \cdot y \quad (3.1)$$

Le GAN représenté dans la Figure 3.1 est entraîné de manière semi-supervisée. Le discriminateur utilise pour son entraînement des entrées et vérités terrains complètes, alors que le générateur utilise des entrées complètes et des vérités terrains éparses. C'est pour cette raison qu'il est spécifié sur la Figure 3.1 qu'ils sont entraînés sur deux datasets différents.

L'avantage principal de cette méthode est qu'elle force le modèle à respecter les CAL, au lieu de lui apprendre à les respecter. Par conséquent, le respect des CAL est assuré.

Cependant cela implique que la méthode se repose sur la fonction de coût utilisée par le modèle et la rétro-propagation pour apprendre au modèle à produire des sorties avec une répartition de valeurs cohérentes autour des CAL. Autrement dit, on espère éviter toute discontinuité autour des CAL.

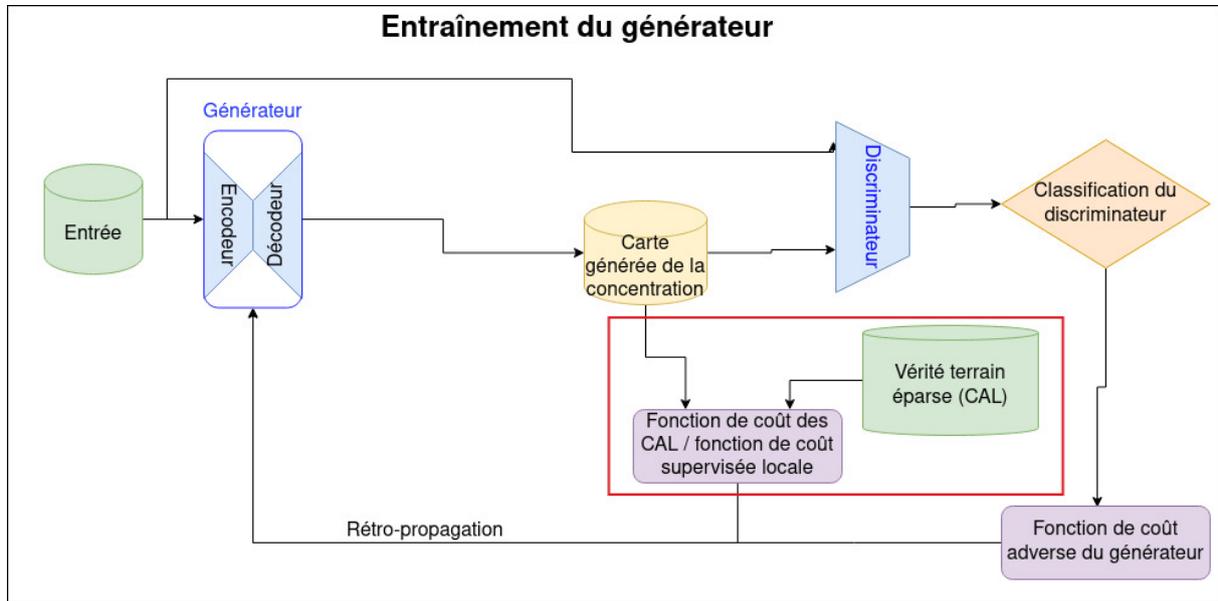
L'hypothèse faite durant ces travaux est que, pour que cette méthode puisse fonctionner, la fonction de coût doit contenir des informations sur les relations locales entre les valeurs des sorties, et sur leur répartition globale. En effet, il est supposé que ce type d'information est nécessaire pour pouvoir préserver la continuité des valeurs.

Les fonctions de coût des PINNs sont conçues à partir d'une équation physique pertinente dans le contexte dans lequel ils sont utilisés. Cette équation contient souvent des informations sur la distribution de valeurs de la sortie recherchée et les relations locales entre celles-ci. Cela rend l'utilisation de la contrainte dure particulièrement adaptée pour les PINNs [49]. A l'inverse, la fonction de coût adverse, conçue pour l'apprentissage des GANs, ne contient pas ce type d'informations.

### 3.1.2 Contrainte douce

L'idée derrière la méthode de la contrainte douce est de concevoir une fonction de coût additionnelle, spécialement pour apprendre au modèle à respecter les CAL. Au lieu de se reposer sur la fonction de coût existante, cette méthode en introduit une nouvelle. Le

modèle n'est pas forcé à respecter les CAL, comme c'est le cas avec la contrainte dure. A la place, il est explicitement entraîné à les respecter. La Figure 3.2 illustre cette idée.



**Figure 3.2:** Schéma représentant l'entraînement du générateur d'un GAN, avec utilisation de la contrainte douce sur les CAL. Dans ce contexte, les CAL sont des valeurs éparses de la concentration en aérosols au niveau du sol.

L'entraînement du discriminateur n'est pas représenté sur ce second schéma, car il est entraîné de la même façon qu'illustré par la Figure 3.1. Le générateur et le discriminateur utilisent les mêmes jeux de données avec les contraintes douces et dures.

La fonction de coût additionnelle est essentiellement une fonction de coût supervisée, mais localisée, appliquée à une matrice éparsée. En effet, elle ne se concentre que sur les pixels de l'image dont la localisation correspond aux CAL. L'Erreur Quadratique Moyenne (Mean Squared Error) (MSE) est appliquée sur ces matrices éparsées. L'équation 3.2 représente la manière dont cette nouvelle fonction de coût, appelée  $loss_{CAL}$ , est calculée. La même notation que pour l'équation 3.1 est utilisée.

$$loss_{CAL}(CAL, \hat{y}, y) = MSE(CAL \cdot \hat{y}, CAL \cdot y) \quad (3.2)$$

Comme le modèle n'est cette fois pas forcé à respecter les CAL, il est possible qu'il produise des sorties avec des valeurs des CAL moins proches de la réalité, comme précisé par Sun et al. [49]. Cependant l'introduction d'une nouvelle fonction de coût permet d'entraîner le modèle à respecter les CAL de manière plus explicite. On ne se repose ainsi plus sur la fonction de coût existante du modèle pour produire des

sorties sans discontinuité autour des CAL.

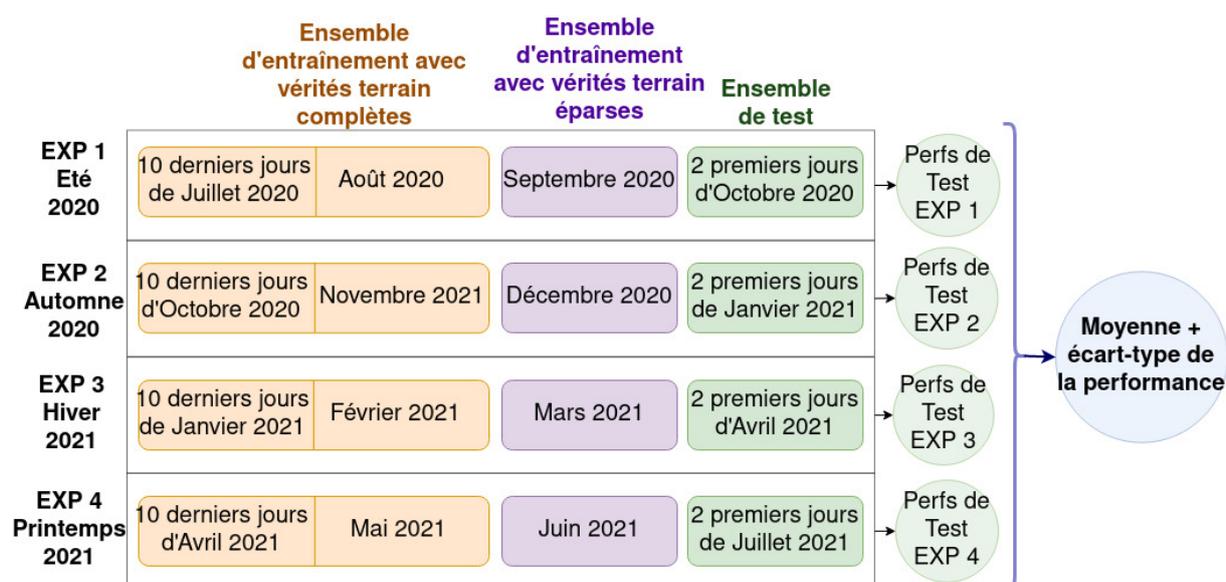
## 3.2 Protocole expérimental

### 3.2.1 Données

Plusieurs expériences sont réalisées pour étudier les méthodes proposées en section 3.1. Les modèles utilisés lors de ces expériences utilisent comme donnée d'entrée des valeurs de l'AOD, et comme vérité terrain des valeurs de concentration en aérosols au niveau du sol. Ces valeurs prennent la forme de matrices géo-référencées de taille 241x480 couvrant la planète entière, et sont issues du modèle CAMS. Avec cette couverture géographique, cette résolution est relativement grossière. Cela doit être pris en compte dans l'analyse des résultats.

La nature météorologique de ces données peut induire un biais dans les résultats. En effet, un certain nombre de phénomènes et processus météorologiques induisent une variabilité significative dans les valeurs de concentration en PM<sub>2.5</sub> en fonction des saisons. Or si les ensembles de tests et d'apprentissage présentent des valeurs moyennes de concentration en PM<sub>2.5</sub> significativement différentes, cela augmente le risque de surapprentissage. Les résultats et performances obtenus peuvent donc aussi varier en fonction de la période temporelle à laquelle le jeu de données correspond : c'est ce que l'on appelle le biais saisonnier. Pour éviter ce problème et obtenir des résultats représentatifs de la performance générale des modèles, chaque expérience est réalisée quatre fois, en utilisant quatre ensembles de données différents : un pour chaque saison. La moyenne et l'écart-type des résultats sont présentés, ce qui met en évidence la résistance (ou la sensibilité) des différentes méthodes au biais saisonnier. La figure 3.3 illustre la manière dont ces expériences sont réalisées, et présente les jeux de données utilisés pour ces dernières. Pour chaque période et chaque expérience, deux jeux de données d'entraînement sont considérés. Le premier peut être appelé ensemble d'entraînement complet. Il contient 336 échantillons, eux-même composés de matrices complètes de l'AOD et de vérités terrains complètes (autrement dit, de matrices complètes de concentration en aérosol). Cet ensemble est utilisé pour réaliser de l'entraînement supervisé et semi-supervisé.

Le deuxième peut être appelé ensemble d'entraînement épars. Il contient 240 échantillons, composés également de matrices complètes de l'AOD, servant de données d'entrée. Cependant sa spécificité est qu'il ne contient que des vérités terrains



**Figure 3.3:** Illustration des quatre périodes temporelles sur lesquelles chaque expérience est réalisée. Cela permet de prendre en compte le biais saisonnier dans les résultats.

éparses. Cette particularité permet de se rapprocher du cas réel puisque, comme décrit en section 1.3, les réseaux de capteurs de concentration en aérosols ne fournissent que des valeurs éparses. Ces vérités terrains éparses sont en réalité générées en sélectionnant un certain nombre de pixels des vérités terrains complètes correspondantes. Elles correspondent également aux CAL utilisées durant les expériences en entraînement semi-supervisé.

Le nombre exact de valeurs (ou pixels) contenus dans ces CAL, ainsi que leur répartition, est décrit par ce qui sera appelé ici la "configuration des CAL". Cette configuration est un paramètre important dans les expériences en apprentissage semi-supervisé réalisées, et plusieurs configurations possibles sont décrites en section 3.2.3.3.

L'ensemble de test contient le même type de valeurs que l'ensemble d'entraînement complet. Il contient 16 échantillons.

### 3.2.2 Pré-traitement des données

Les modèles météorologiques présentés en section 1.5 peuvent fournir des valeurs d'AOD et de concentration en PM<sub>2.5</sub>. Cependant ces valeurs doivent subir un pré-traitement afin de faciliter leur utilisation ou d'augmenter les performances de nos modèles.

Toutes les valeurs d’AOD inférieures à 0.005 sont ramenées à 0, car elles sont assimilables à du bruit.

Une analyse rapide de la distribution des valeurs d’AOD et de concentration montre que celle-ci est loin d’être uniforme. Elle a davantage la forme d’une exponentielle, plus précisément de la fonction  $\exp(\frac{-1}{x})$ . Cela implique donc une forte quantité de valeurs proches de 0, et une très faible quantité de valeurs plus élevées. Ce déséquilibre peut rendre plus difficile la convergence des modèles.

De plus comme évoqué en section 1.1, ces valeurs élevées de concentration en aérosols sont associées à des effets néfastes sur la santé. Cela peut donc nous amener à nous questionner sur la performance de certains de nos modèles en prédiction de valeurs extrêmes. Quelques pistes de réponses à ce questionnement sont proposées en section A.7.

Pour obtenir une distribution plus uniforme, la fonction  $\ln(1+x)$  est appliquée sur les entrées et la vérité terrain. Cette fonction est préférable à la fonction  $\ln(x)$  car elle permet d’éviter l’obtention de valeurs indéfinies comme  $\ln(0)$ . La figure 3.4 montre la distribution de valeurs de concentration, avec et sans l’application d’un logarithme. L’application de ce pré-traitement permet d’obtenir de meilleures performances, comme illustré dans la section A.1.

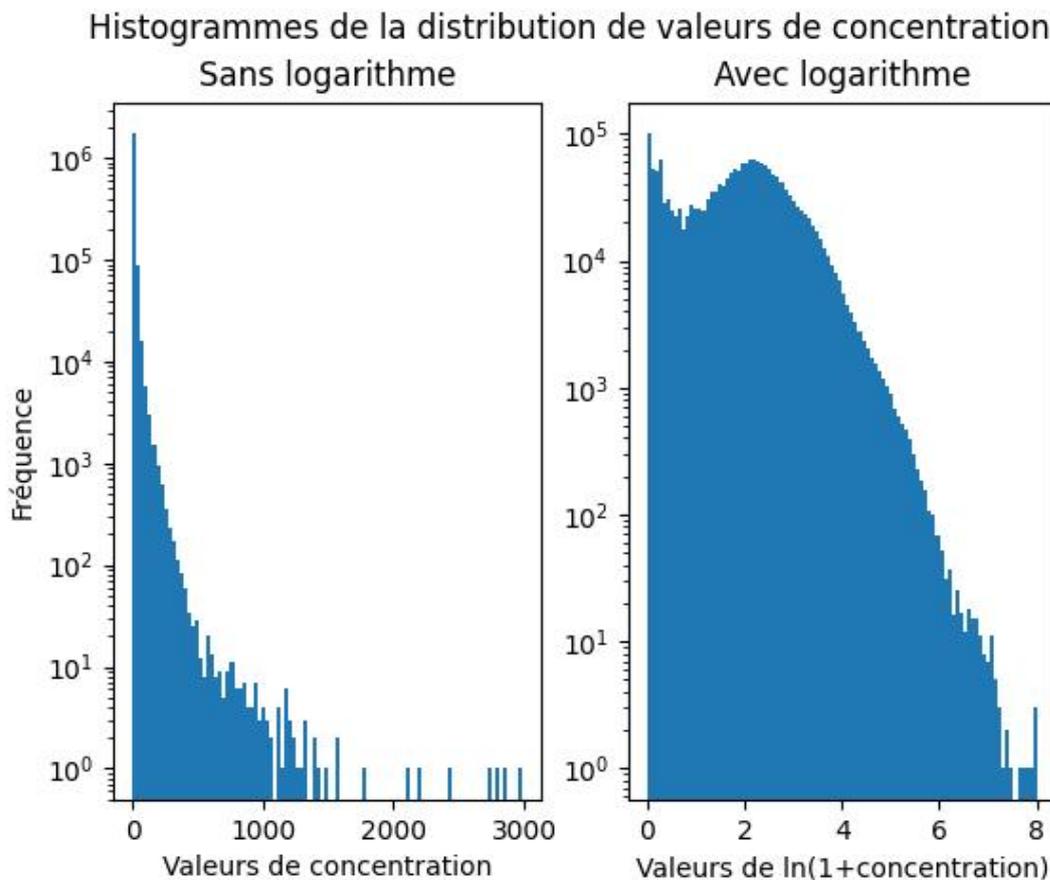
Lors de la phase de test, et après chaque inférence, la fonction  $\exp(x) - 1$  est appliquée aux sorties du modèles et aux vérités terrains, afin d’obtenir des valeurs de concentration en aérosols exprimées en  $\mu\text{g}/\text{m}^3$ . Cela permet d’obtenir des métriques exprimées en  $\mu\text{g}/\text{m}^3$  également, ce qui facilite l’interprétation des résultats.

De même, durant la phase de test et après chaque inférence, les valeurs de concentration en aérosols inférieures à  $1 \mu\text{g}/\text{m}^3$  sont ramenées à  $0 \mu\text{g}/\text{m}^3$ . Cela correspond à environ 19% de la quantité totale de valeurs, en moyenne. Ce traitement permet de mieux évaluer la performance des modèles dans les zones polluées, qui sont les zones d’intérêt principales. Il est appliqué aux vérités terrains et aux prédictions, avant le calcul des métriques.

### 3.2.3 Description des expériences

#### 3.2.3.1 Modèles pour entraînement strictement supervisé

Les premières expériences sont réalisées en apprentissage strictement supervisé. Plusieurs variations de modèles d’encodeur-décodeur sont conçues, implémentées et



**Figure 3.4:** Histogrammes représentant la distribution des valeurs de concentration en aérosols, avec et sans l’application d’un logarithme.

testées sur les mêmes données. On conçoit ainsi des variantes avec et sans couches linéaires (CL), et avec et sans connexions résiduelles (CR). En cas de présence de connexions résiduelles, ces modèles d’encodeur-décodeur sont appelés des UNets, en référence au modèle proposé par Ronneberger, Fischer et Brox [43]. Ces modèles ont en commun leur nombre de couches de convolution ainsi que leurs tailles de noyaux de convolution. La figure 3.5 représente la plus complexe de ces variantes, avec couches linéaires et connexions résiduelles.

### 3.2.3.2 Modèles pour entraînement semi-supervisé

Dans un second temps, des expériences en apprentissage semi-supervisé sont réalisées. Celles-ci utilisent un modèle GAN, composé d’un générateur et d’un discriminateur. Au vu des résultats obtenus en strictement supervisé, il est décidé d’utiliser un modèle de UNet avec couches linéaires comme générateur. Le discriminateur est un encodeur

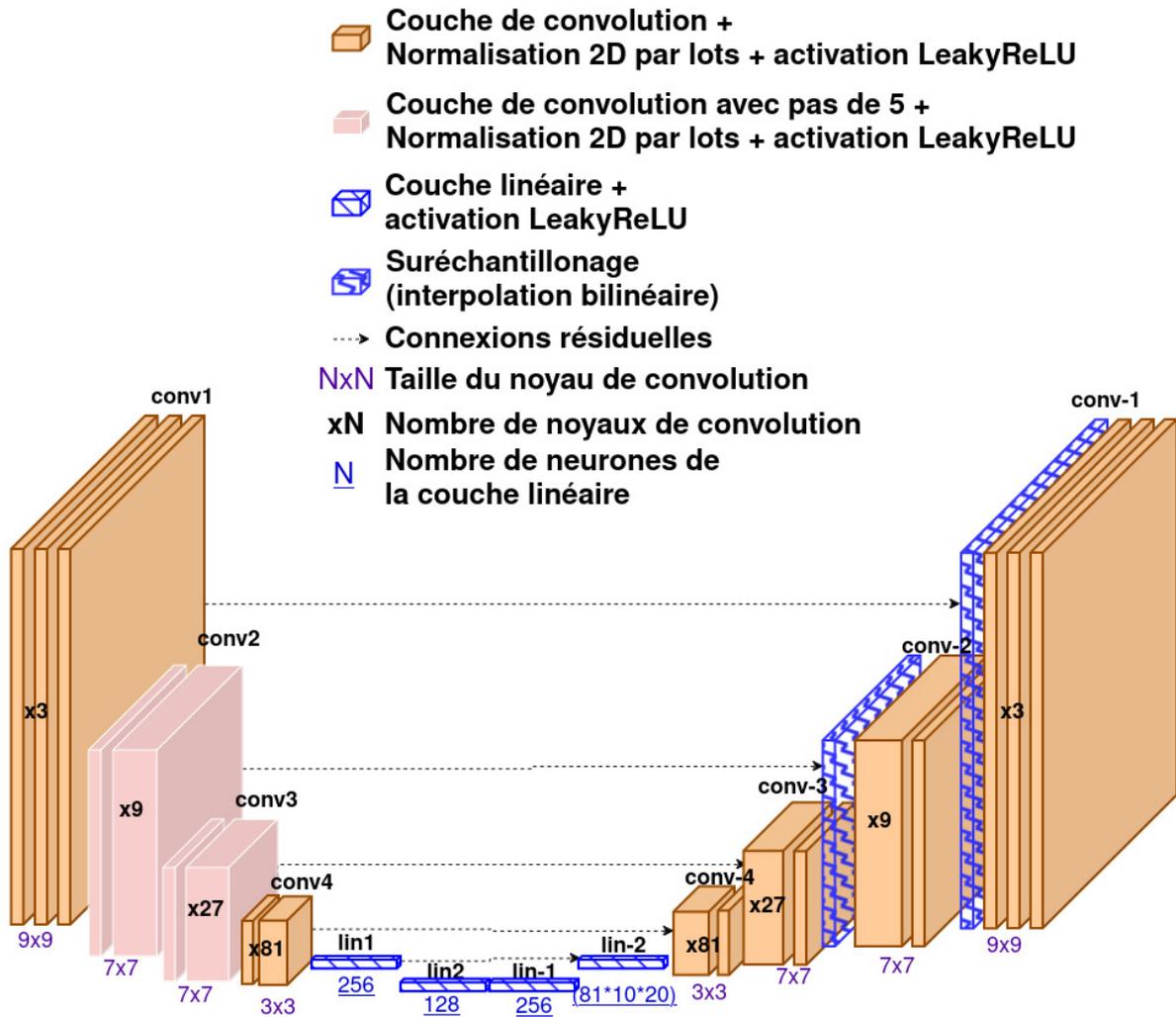


Figure 3.5: Schéma d'un UNet avec couches linéaires.

convolutif suivi de deux couches linéaires. Il peut recevoir deux images de taille 241x480 en entrée, et fournit un unique scalaire en sortie.

L'ensemble d'entraînement complet (tel qu'il est nommé dans la section 3.2.1) est utilisé pour l'apprentissage du discriminateur, tandis que l'ensemble d'entraînement épars est utilisé pour l'apprentissage du générateur.

Deux versions de ce GAN sont conçues, implémentées et testées : une avec la contrainte dure pour le respect des CAL, et une avec la contrainte douce, telles que ces méthodes sont décrites dans la section 3.1. Chaque version est testée avec chacune des configurations de CAL proposées dans la section 3.2.3.3.

### 3.2.3.3 Configuration des CAL

L'un des objectifs de ces expériences en apprentissage semi-supervisé est d'étudier l'impact que les CAL, ainsi que leur configuration, peuvent avoir sur les résultats. Cinq configurations de CAL sont proposées, implémentées et utilisées dans le cadre de ces expériences.

**Sans** : aucune CAL (cas habituel de l'entraînement semi-supervisé)

**Aléatoire 1%** : 1% des pixels de la vérité terrain sont utilisés en tant que CAL, et ils sont localisés à des positions aléatoires

**Aléatoire 5%** : même idée que pour **Aléatoire 1%** mais 5% de pixels sont utilisés au lieu de 1%

**Aléatoire Terre 1%** : même idée que pour **Aléatoire 1%**, mais les pixels choisis en tant que CAL sont toujours géographiquement positionnés sur terre (donc pas dans la mer ou l'océan). L'idée est de simuler les positions géographiques que pourraient avoir des capteurs de concentration en aérosols, or ceux-ci ne peuvent être situés que sur terre.

**Aléatoire Terre 5%** : même idée que pour **Aléatoire Terre 1%** mais avec 5% de pixels au lieu de 1%

Lors de la génération de ces CAL, le but est de simuler le cas réel, où les données sont fournies par un réseau de capteurs au sol. Chaque pixel des CAL représente essentiellement un capteur. A cause de la résolution relativement grossière des images utilisées ici, la couverture de ce type de réseau est surestimée par la représentation proposée ici. Il est intéressant de noter que l'utilisation de 1% des pixels en tant que CAL aboutit à un nombre de pixels plus proche du réel nombre de capteurs sur la surface de la Terre qu'en utilisant 5% de pixels.

## 3.3 Evaluation des modèles

L'évaluation des différents modèles durant le test se fait par l'utilisation de différentes métriques. La valeur de chacune de ces métriques est calculée par comparaison entre la matrice obtenue en sortie et la vérité terrain.

Trois métriques sont utilisées constamment tout au long des expériences menées

durant cette thèse. Il s'agit de l'Erreur Absolue Moyenne (Mean Absolute Error) (MAE), du Biais Moyen (Mean Bias Error) (MBE) et de l'Erreur Quantifiée (Quantized Error) (QE). La MAE et la MBE sont exprimées en  $\mu\text{g}/\text{m}^3$  (l'unité de concentration en aérosols dans l'atmosphère). Les versions relatives de la MAE et MBE (exprimées en pourcentages) peuvent aussi être utilisées, elles sont appelées l'Erreur Absolue Moyenne relative (relative Mean Absolute Error) (rMAE) et le Biais Moyen relatif (relative Mean Bias Error) (rMBE), respectivement. Elles correspondent à la MAE ou MBE divisée par la valeur moyenne de la concentration en PM<sub>2.5</sub> à estimer. Une division par la valeur maximale amènerait à des valeurs trop proches de 0%, et ne permettrait pas une interprétation aisée. Exprimer la même métrique selon différentes unités permet de faciliter l'interprétation des résultats et leur communication à destination de chercheurs de différentes communautés. Cela peut également servir au calcul d'un score global, prenant en compte plusieurs métriques.

Une quatrième métrique est introduite durant ces travaux de recherche, et répond à un besoin de mieux évaluer la performance des modèles proposés. Elle est décrite en section 4.2.4.

- **Erreur Absolue Moyenne (Mean Absolute Error) (MAE).** De toutes les métriques utilisées dans le cadre de cette thèse, celle-ci est la plus populaire et la plus utilisée. Elle sert de représentation de l'erreur globale des modèles et méthodes évalués [75]. Lorsqu'elle est exprimée en pourcentages, on l'appelle la rMAE.
- **Biais Moyen (Mean Bias Error) (MBE).** Cette métrique représente la tendance d'un modèle ou d'une méthode à la sur ou sous-estimation. Un modèle qui ne présenterait pas de telle tendance claire peut tout de même présenter de faibles performances, c'est pourquoi cette métrique seule ne suffit pas pour l'analyse des résultats. Cependant elle représente une information utile qui peut éventuellement être utilisée pour proposer une correction des résultats du modèle. Son mode de calcul revient à retirer la valeur absolue dans le calcul de la MAE. Lorsqu'elle est exprimée en pourcentages, on l'appelle la rMBE.
- **Erreur Quantifiée (Quantized Error) (QE).** Cette métrique a été proposée dans le cadre de cette thèse et spécifiquement conçue pour les besoins de celle-ci. L'idée est de quantifier la prédiction et la vérité terrain avant de les comparer. Cette quantification est réalisée en utilisant les quartiles des valeurs de la vérité terrain, ce qui nous permet d'obtenir quatre classes. De plus, en ce qui concerne la vérité terrain le nombre de pixels appartenant à chacune de ces classes est

quasi-équivalent.

L'équation 3.3 illustre le processus de quantification. La matrice initiale  $y$  est notée  $M$  et ses éléments  $M_{i,j}$ .  $C$  représente la matrice quantifiée et  $C_{i,j}$  ses éléments.

$$C_{i,j} = \begin{cases} 0 & \text{si } M_{i,j} \leq q1 \\ 1 & \text{si } M_{i,j} \in ]q1, q2] \\ 2 & \text{si } M_{i,j} \in ]q2, q3] \\ 3 & \text{si } M_{i,j} > q3 \end{cases} \quad (3.3)$$

Après application de ce processus, une vérité terrain et une prédiction quantifiées sont obtenues. Celles-ci sont respectivement notées  $C_{vt}$  et  $C_{pred}$  dans l'équation 3.4. Cette dernière illustre le mode de calcul de a QE à partir de ces matrices quantifiées.  $N$  y représente le nombre d'éléments (de pixels) dans chacune de ces matrices.

$$QE = \frac{|C_{vt} - C_{pred}|}{N} \quad (3.4)$$

Cette métrique serait traditionnellement utilisée pour des tâches de classification ou segmentation, plutôt qu'une tâche de translation image-vers-image comme c'est le cas ici. L'un des intérêts de cette métrique est sa ressemblance avec une représentation typique de la qualité de l'air. En effet la qualité de l'air est souvent illustrée par des seuils, qui correspondent en général à des niveaux d'alerte ou à des risques de santé publique. Cette métrique nous permet ainsi d'établir l'erreur du modèle lors d'une utilisation par classe plutôt que par valeur continue.

### 3.3.1 Disponibilité du code et des données

Le code utilisé pour ces expériences est inclus dans celui proposé en section 4.2.6.<sup>1</sup> La spécificité des expériences présentées dans ce chapitre vis-à-vis de ce code réside dans l'utilisation d'une seule variable d'entrée (l'AOD), ce qui entraîne l'absence d'un besoin de stratégie de fusion de données.

Les données données (issues du modèle CAMS) utilisées pour ces expériences sont rendues disponibles via une archive Zenodo.<sup>2</sup>

<sup>1</sup><https://doi.org/10.5281/zenodo.13947256>

<sup>2</sup><https://doi.org/10.5281/zenodo.14355407>

## 3.4 Résultats

### 3.4.1 Encodeurs-décodeurs en apprentissage purement supervisé

La table 3.1 présente les résultats des expériences décrites en section 3.2.3.1. On y trouve donc les résultats obtenus avec chacune des variantes d’encodeur-décodeurs proposées.

**Table 3.1:** Résultats obtenus avec les modèles d’encodeurs-décodeurs supervisés. La colonne U définit si l’encodeur-décodeur contient des connexions résiduelles (auquel cas il s’agit d’un UNet), et la colonne L s’il contient des couches Linéaires.

Modèle		MAE	rMAE	MBE	rMBE	QE (sans unités)
U	L	( $\mu\text{g}/\text{m}^3$ )	(%)	( $\mu\text{g}/\text{m}^3$ )	(%)	
Non	Non	8.69 $\pm 1.68$	65.79% $\pm 6.82\%$	-1.03 $\pm 1.18$	-7.98% $\pm 9.12\%$	0.52 $\pm 0.04$
Non	Oui	9.49 $\pm 1.56$	71.99% $\pm 4.89\%$	-3.16 $\pm 0.61$	-23.74% $\pm 2.21\%$	0.69 $\pm 0.07$
Oui	Non	7.94 $\pm 2.12$	59.74% $\pm 10.36\%$	-0.79 $\pm 3.48$	-7.89% $\pm 27.56\%$	0.43 $\pm 0.02$
<b>Oui</b>	<b>Oui</b>	<b>6.69</b> <b><math>\pm 1.03</math></b>	<b>50.71%</b> <b><math>\pm 2.97\%</math></b>	-1.36 $\pm 0.7$	-10.21% $\pm 4.79\%$	<b>0.39</b> <b><math>\pm 0.02</math></b>

Le modèle le plus performant semble être le UNet avec couches linéaires. En ce qui concerne les encodeurs-décodeurs sans connexions résiduelles, ajouter des couches linéaires semble amener à une baisse des performances. Mais pour les UNet, ajouter ce type de couche a l’effet inverse.

L’interprétation proposée pour ce phénomène est qu’en l’absence de connexions résiduelles, ajouter des couches linéaires entre l’encodeur et le décodeur amène à une perte d’information. En effet, dans ce cas de figure la représentation latente fournie au décodeur prend la forme d’un vecteur, au lieu d’une matrice. Il semblerait que présenter l’information sous forme d’une matrice présente un avantage pour le modèle, possiblement parce que l’information est davantage localisée.

En présence de connexions résiduelles, cette perte d’informations n’a pas lieu. En effet, la première couche convolutive du décodeur reçoit quoi qu’il arrive une matrice, issue de la dernière couche de convolution de l’encodeur. Par conséquent, la présence de couches linéaires peut amener à l’ajout d’une information additionnelle sous la forme d’un vecteur, sans perte d’information. Cette information supplémentaire est potentiellement utile pour l’apprentissage, d’où le gain en performance.

La réflexion proposée ici peut amener à des questionnements sur l’impact de chaque

couche du UNet sur la performance, et donc par extension de l'impact de chaque échelle de représentation de nos données. La section A.8 tente de répondre à ces questionnements à travers une étude d'ablation.

### 3.4.2 Contrainte dure et contrainte douce

Les tables 3.2 et 3.3 montrent les résultats des expériences décrites en section 3.2.3.2, utilisant respectivement les contraintes dure et douce.

**Table 3.2:** Résultats des expériences sur les GANs en apprentissage semi-supervisé, avec la contrainte dure.

Type de CAL (contrainte dure)	MAE ( $\mu\text{g}/\text{m}^3$ )	rMAE (%)	MBE ( $\mu\text{g}/\text{m}^3$ )	rMBE (%)	QE (sans unités)
Sans	13.14 $\pm 1.02$	100.47% $\pm 7.77\%$	-12.09 $\pm 1.81$	-91.7% $\pm 5.81\%$	1.41 $\pm 0.07$
Aléatoire 1%	13.45 $\pm 1.2$	102.58% $\pm 2.09\%$	-11.71 $\pm 1.51$	-89.01% $\pm 3.74\%$	1.49 $\pm 0.02$
<b>Aléatoire 5%</b>	<b>12.84</b> $\pm 1.4$	<b>97.78%</b> $\pm 1.2\%$	-11.34 $\pm 1.23$	-86.32% $\pm 2.2\%$	<b>1.41</b> $\pm 0.06$
Aléatoire Terre 1%	13.65 $\pm 0.32$	104.94% $\pm 12.01\%$	-11.44 $\pm 2.16$	-86.54% $\pm 8.7\%$	1.45 $\pm 0.18$
Aléatoire Terre 5%	15.01 $\pm 4.8$	116.06% $\pm 45.75\%$	<b>-9</b> $\pm 6.06$	<b>-66.96%</b> $\pm 42.91\%$	1.42 $\pm 0.09$

**Table 3.3:** Résultats des expériences sur les GANs en apprentissage semi-supervisé, avec la contrainte douce.

Type de CAL (contrainte douce)	MAE ( $\mu\text{g}/\text{m}^3$ )	rMAE (%)	MBE ( $\mu\text{g}/\text{m}^3$ )	rMBE (%)	QE (sans unités)
Sans	13.14 $\pm 1.02$	100.47% $\pm 7.77\%$	-12.09 $\pm 1.81$	-91.7% $\pm 5.81\%$	1.41 $\pm 0.07$
Aléatoire 1%	7.5 $\pm 0.99$	57% $\pm 3.15\%$	<b>-0.85</b> $\pm 2.45$	<b>-5.54%</b> $\pm 17.46\%$	0.42 $\pm 0.04$
<b>Aléatoire 5%</b>	<b>6.51</b> $\pm 1.31$	<b>49.25%</b> $\pm 6.07\%$	-1.33 $\pm 2.61$	-10.85% $\pm 20.03\%$	<b>0.35</b> $\pm 0.03$
Aléatoire Terre 1%	11.73 $\pm 1.45$	89.71% $\pm 9.31\%$	-0.97 $\pm 1.25$	-6.7% $\pm 9.06\%$	0.79 $\pm 0.03$
Aléatoire Terre 5%	8.73 $\pm 0.67$	66.65% $\pm 4.53\%$	-4.11 $\pm 1.99$	-30.43% $\pm 12.96\%$	0.68 $\pm 0.07$

Dans le cas de ces GANs, la contrainte douce pour l'utilisation des CAL semble plus efficace que la contrainte dure, contrairement à ce que la littérature en dit dans le cas

des PINNs. En effet, Sun et al. [49] affirment que l'utilisation de la contrainte dure pour les CAL amène à une meilleure performances des modèles. De plus on remarque sur la table 3.2 que lorsque la contrainte dure est utilisée, l'utilisation de CAL comme information a priori n'amène pas toujours à une meilleure performance du modèle.

Inversement, lorsque la contrainte douce est utilisée, la configuration de CAL amenant à la pire performance est celle sans CAL, comme montré par la table 3.3. Celle-ci montre également qu'ajouter des CAL (en termes de quantité de pixels labellisés constituant les CAL) amène à une augmentation significative de la performance. Observer ce phénomène uniquement avec la contrainte douce et pas avec la contrainte dure paraît surprenant, car l'information fournie au modèle est la même dans les deux cas.

Pour proposer une interprétation à cette observation, il est utile de s'intéresser au fonctionnement des PINNs. Ces modèles utilisent une fonction de coût basée sur une Equation Différentielle (ED). La solution d'une ED dépend des Conditions Aux Limites (CAL) et Conditions Initiales (CI). Par conséquent, une erreur dans les CAL/I peut induire une erreur dans toutes les valeurs de la variable prédite.

C'est pour cela que d'après Sun et al. [49], l'inconvénient majeur de la méthode douce est qu'elle n'assure pas la précision de ces CAL/I, au contraire de la méthode dure. Ces mêmes auteurs expliquent que des "effets de compétition" entre les différentes fonction de coût peuvent avoir lieu lors de l'utilisation de la méthode douce, et ces effets peuvent rendre plus difficile la convergence. Il peut également s'avérer difficile d'attribuer un poids à chaque fonction de coût.

A la lumière de ces explications, une interprétation sur la bonne performance des PINNs avec la contrainte dure peut être proposée. Il semblerait ainsi que les fonctions de coût informées par la physique contiennent des informations sur la dépendance entre valeurs, sur la manière dont les valeurs à prédire s'influencent mutuellement en fonction de la distance qui les séparent. Par conséquent, les PINNs reçoivent des informations sur la manière dont un changement dans la valeur d'une CAL devrait influencer d'autres valeurs.

Une interprétation similaire concernant la mauvaise performance des GANs avec la contrainte dure peut également être proposée. En effet, la fonction de coût adverse des GANs est un simple scalaire, et ne contient pas le même type d'information qu'une fonction de coût informée par la physique. Il est raisonnable de supposer qu'elle ne contient pas d'informations sur la dépendance entre les valeurs des CAL et les autres

valeurs. Le GAN n’ayant donc pas accès à cette information, cela pourrait expliquer sa faible performance avec la contrainte dure sur les CAL (qui force le modèle à respecter les CAL plutôt que le lui apprendre).

Il apparaît par conséquent que les CAL sont effectivement une information a priori utile pour les modèles GANs, mais qu’elle doit être utilisée différemment que pour les PINNs. Il semble en effet que permettre à un GAN d’apprendre à respecter les CAL l’aide à inférer correctement d’autres valeurs.

Un autre argument peut être proposé pour expliquer les phénomènes observés, et leur différence avec ce qui serait attendu dans le cas d’un PINN. En effet, la nature même des CAL utilisées est particulière, dans le sens où elle est différente des CAL habituellement utilisées pour l’entraînement d’un PINN [48, 49, 65, 76]. En effet celles-ci concernent classiquement des valeurs connues pour certaines variables aux limites spatiales (ou temporelles dans le cas de Conditions Initiales) du domaine, plutôt que pour certains points épars et aléatoirement répartis. Les CAL utilisées ici sont moins structurées. Il serait utile de comparer les résultats d’un PINN avec ces mêmes CAL et les deux contraintes proposées. Cela permettrait d’évaluer si la nature et la forme des CAL a un lien avec l’utilité de chacune de ces contraintes. Une telle comparaison est proposée en section 5.3.3.

Enfin, la fonction de coût dédiée aux CAL est localisée, dans le sens où ses valeurs sont réparties (de manière épars) sur une matrice. Il est possible que l’ajout de cette fonction de coût aide le GAN à localiser les erreurs qu’il commet. En effet, la fonction de coût originale du GAN, la fonction de coût adverse, est plus globale : elle ne contient qu’un scalaire par échantillon.

En résumé, l’utilisation de la contrainte douce pour l’utilisation de CAL/I est recommandée lors de l’entraînement d’un GAN.

- **Autres observations**

Il est intéressant de noter que la configuration de CAL **Aléatoire 5%** semble amener à une plus grande sensibilité au biais saisonnier que la configuration **Aléatoire Terre 5%**. Il est possible que l’utilisation de CAL au-dessus de mers ou d’océans amène à une plus grande variabilité parmi les valeurs des CAL que lorsqu’elles sont exclusivement situées sur la terre.

De plus, si l’on compare les résultats obtenus avec les configurations **Aléatoire 5%** et

1%, il apparaît que parfois, utiliser davantage de CAL (en termes de nombre de pixels) amène à plus grande sensibilité au biais saisonnier. Cela peut aussi amener à une MBE légèrement plus éloignée de zéro. Cependant cela peut être dû à la nature aléatoire des CAL. Par exemple, ajouter des CAL à un endroit où les valeurs de concentration sont légèrement plus haute pourrait amener le modèle à légèrement surestimer la concentration en aérosols sur le domaine complet. Ceci devrait être pris en compte lors du développement d'une application réelle à partir d'un réseau de capteurs existant.

### 3.4.3 Comparaison avec les méthodes proposées comme base de comparaison

Dans cette section, nous réalisons une comparaison entre nos propositions et les méthodes classiques de l'état de l'art.

**Table 3.4:** Comparaison des résultats des modèles proposés et des algorithmes et méthodes présentés en section 2.2. Dans la colonne type d'apprentissage, "S" signifie "Supervisé", "Semi-S" "Semi-Supervisé" et "NA" "Non Applicable" (pour les méthodes où aucun apprentissage n'a lieu).

Type	Modèle	MAE ( $\mu\text{g}/\text{m}^3$ )	rMAE (%)	MBE ( $\mu\text{g}/\text{m}^3$ )	rMBE (%)	SE (no units)
S	Interpolation Polynômiale de Degré 3	7.14 $\pm 1.04$	54.12% $\pm 2.69\%$	-2.76 $\pm 1.51$	-21.19% $\pm 11.62\%$	0.41 $\pm 0.03$
S	Random Forest	7.17 $\pm 1.07$	54.34% $\pm 2.89\%$	-2.7 $\pm 1.43$	-20.85% $\pm 11.27\%$	0.41 $\pm 0.03$
S	UNet + Couches Linéaires (proposition)	6.69 $\pm 1.03$	50.71% $\pm 2.97\%$	-1.36 $\pm 0.7$	<b>-10.21%</b> $\pm 4.79\%$	0.39 $\pm 0.02$
NA	Krigeage Aléatoire 5%	<b>5.81</b> $\pm 0.77$	<b>44.07%</b> $\pm 1.98\%$	-3.36 $\pm 0.7$	-25.36% $\pm 3.1\%$	<b>0.32</b> $\pm 0.02$
Semi-S	GAN contrainte douce Aléatoire 5% (proposition)	6.51 $\pm 1.31$	49.25% $\pm 6.07\%$	<b>-1.33</b> $\pm 2.61$	-10.85% $\pm 20.03\%$	0.35 $\pm 0.03$

La table 3.4 montre que le UNet avec couches linéaires présente de meilleures performances que les méthodes supervisées présentées en section 2.2 (i.e. l'Interpolation Polynômiale de degré 3 et l'algorithme Random Forest). Ces deux méthodes sont appliquées pixel-par-pixel et ne tiennent donc pas compte de l'aspect spatial du problème. Cela montre donc l'importance de ce type d'information dans ce contexte.

Cependant, l'écart de performance entre le UNet et ces méthodes est relativement petit. Comme dit précédemment, il est suggéré que cet écart de performance est principalement dû au fait que ces méthodes sont appliquées pixel-par-pixel tandis que

le UNet est un encodeur-décodeur convolutif. Il est donc raisonnable de suggérer que l'utilisation d'images de plus haute résolution amènerait à un écart de performance plus important en faveur du UNet.

Un GAN en apprentissage semi-supervisé, avec la contrainte douce pour l'utilisation des CAL, et la configuration de CAL **Aléatoire 5%**, amène à de meilleurs résultats que ceux obtenus avec le modèle UNet complètement supervisé, et ce sur toutes les métriques. Cela démontre à la fois l'intérêt de l'apprentissage semi-supervisé (par rapport à purement supervisé) et de l'utilisation des CAL durant l'entraînement.

Les GANs semblent également moins robustes au biais saisonnier que les UNets, comme illustré par la table 3.4. Cette plus grande variabilité des résultats des GANs peut être due à l'utilisation de la fonction de coût adverse. Elle permet l'apprentissage semi-supervisé, et l'utilisation de datasets avec des vérités terrains éparses voir absentes, au prix d'une erreur moins précisément définie, avec un caractère plus imprévisible que la MSE utilisée pour entraîner le UNet.

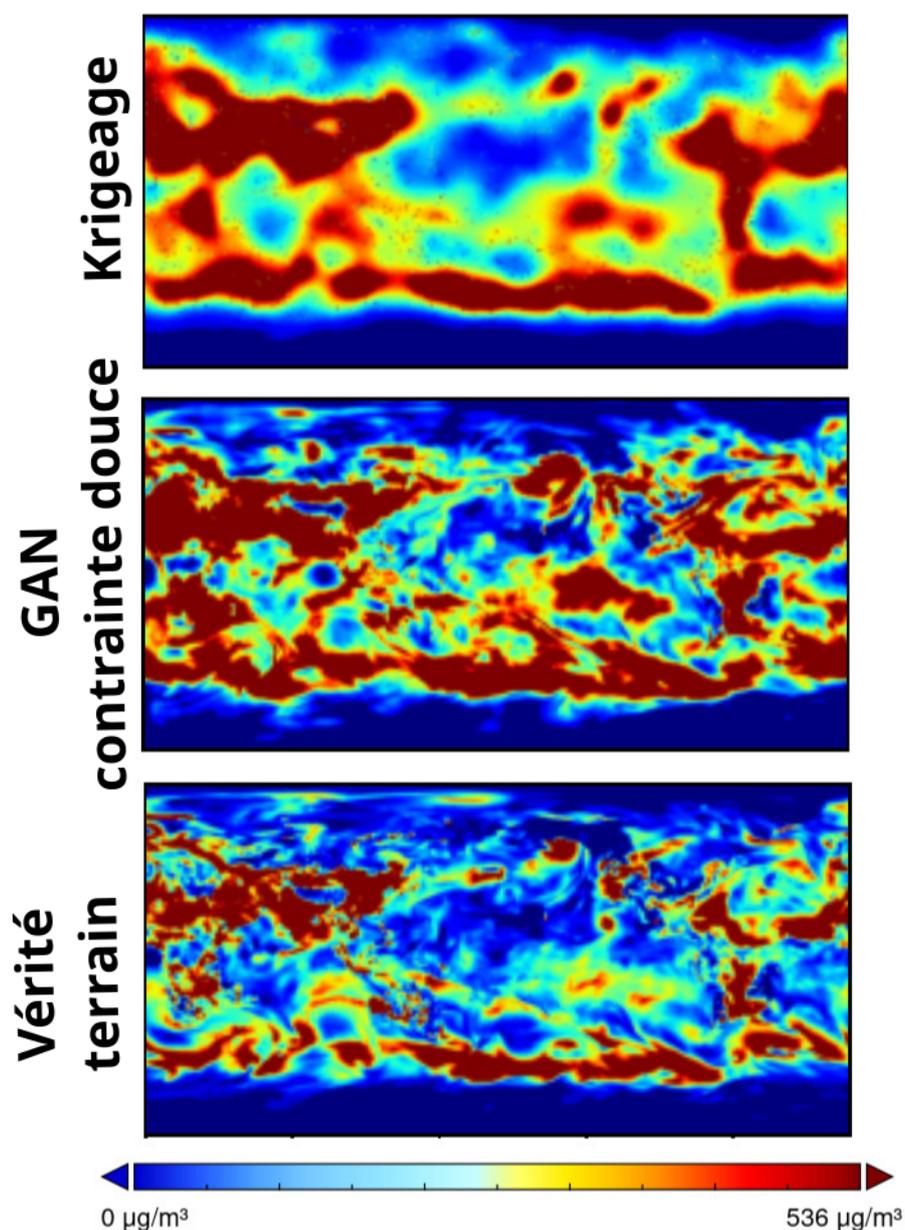
De manière surprenante, la table 3.4 semble indiquer que la méthode fournissant les meilleures performances est le krigeage. C'est aussi la méthode la plus robuste au biais saisonnier. Cela est probablement dû au fait qu'un nouveau modèle de krigeage, basé sur les CAL, est construit pour chaque échantillon (et donc chaque inférence). C'est également la raison pour le grand temps d'inférence de cette méthode, comme montré par la table 3.5. Cela rend le krigeage inutilisable en tant que solution au problème exposé dans le chapitre 1.

**Table 3.5:** Comparaison des temps d'inférence (en secondes) des modèles proposés et des algorithmes et méthodes présentés en section 2.2. "IP deg 3" signifie "Interpolation Polyômiale de degré 3", et "t" représente le temps d'inférence.

Modèle	IP deg 3	Random Forest	UNet	GANs	Krigeage	
					1% CAL	5% CAL
t (s)	0.0006 ± 0.0001	0.978 ± 0.017	0.015 ± 0.001	0.012 ± 0.001	15.774 ± 1.254	129.416 ± 51.205

De plus, comme illustré par la figure 3.6, le krigeage produit des sorties très lisses. Celles-ci sont peut-être précises en moyenne, mais elles manquent de détails. En effet la préservation des formes et structures plus fines est importante pour l'interprétation des sorties du modèle. Les GANs à l'inverse ne souffrent pas de ce problème : leurs sorties préservent les structures fines. Celles-ci sont représentatives de l'influence de la dynamique atmosphérique sur les valeurs de concentration en aérosols. Le krigeage

reste cependant une méthode pertinente à laquelle comparer les modèles proposés.



**Figure 3.6:** Sorties fournies par des modèles GAN et de krigeage avec une configuration de CAL Aléatoire 5%. On trouve la vérité terrain en bas. Les petits points visibles sur la sortie du krigeage représentent les valeurs de CAL.

Sur la figure 3.6, des petits points sont visibles sur les sorties du krigeage. Ils représentent les valeurs de CAL exactes. Celles-ci apparaissent comme des artéfacts ou anomalies car elles diffèrent parfois des valeurs estimées par le modèle dans la même zone.

Enfin, l'argument proposé en section 3.4.3 à propos de la résolution des images peut

être fait ici aussi. En effet, les GANs comme les encodeurs-décodeurs sont supposés bénéficier de l'utilisation d'images de plus haute résolution. Autrement dit leurs performances sont censées s'améliorer dans ce cas. A l'inverse, le krigeage, les méthodes d'interpolation, ainsi que l'algorithme Random Forest sont supposés démontrer de moins bonnes performances lors de l'utilisation d'images de plus haute résolution.

### 3.5 Conclusion

Résoudre un problème physique nécessite parfois d'exploiter les Conditions Aux Limites et l'information qu'elles représentent. Des méthodes ont été proposées pour exploiter cette information au sein de Réseaux de Neurones, mais principalement pour des réseaux Informés par la Physique.

Ces travaux ont fourni une base pour discuter de l'efficacité et de la pertinence de l'application de ces méthodes à un modèle non informé par la physique (plus spécifiquement, à un Réseau Génératif Adverse). L'intérêt en termes de performance de l'exploitation des Conditions Aux Limites durant l'apprentissage d'un Réseau Génératif Adverse a ainsi pu être montré. Ils ont aussi abouti à la recommandation de l'une de ces méthodes d'utilisation des CAL (la contrainte douce).

Pour des raisons pratiques ce type de GAN sera désormais appelé un BC-GAN, pour Réseau Génératif Adverse avec utilisation de Conditions Aux Limites (Generative Adversarial Network with use of Boundary Conditions).

Ces travaux ont mis en évidence l'intérêt d'analyser le comportement de ces méthodes sur des données de plus grande résolution. L'analyse des résultats ainsi que la figure 3.6 met également en avant le besoin d'une métrique additionnelle, capable de représenter la conservation par le modèle des formes et structures fines de l'image.

De plus, si les CAL sont utiles pour un ANN comme pour une solution purement basée sur la physique, alors il est pertinent de se demander si d'autres informations typiquement utiles en physique gagneraient également à être exploitées par un réseau de neurones. En particulier, plusieurs champs météorologiques (humidité, température, pression, vent) sont connus pour leur influence sur la concentration en aérosols dans l'atmosphère et au niveau du sol.

# 4

## Étude d'impact des variables d'entrée

### Sommaire

---

4.1	Proposition . . . . .	83
4.1.1	Stratégies de fusion d'information . . . . .	84
4.1.2	Fusion de données / Concaténation de Canaux (CC) . . . . .	84
4.1.3	Fusion de Représentations (FR) . . . . .	85
4.1.4	Fusion de Décisions (FD) . . . . .	87
4.1.5	Approches de fusion hybrides . . . . .	89
4.2	Protocole expérimental . . . . .	90
4.2.1	Description des données . . . . .	91
4.2.2	Pré-traitement des données . . . . .	92
4.2.3	Description des expériences . . . . .	94
4.2.4	Métrique supplémentaire : la FSIM . . . . .	94
4.2.5	Scores . . . . .	95
4.2.6	Disponibilité du code et des données . . . . .	96
4.3	Résultats . . . . .	96
4.3.1	Aperçu . . . . .	96
4.3.2	Meilleurs résultats . . . . .	99

## Chapitre 4 – Étude d'impact des variables d'entrée

4.3.3	Résultats des méthodes hybrides de fusion . . . . .	104
4.3.4	Comparaison avec les méthodes proposées comme base de comparaison . . . . .	106
4.4	Conclusion . . . . .	<b>107</b>

---

L’observation spatiale des propriétés optiques des aérosols peut fournir des informations à grande échelle, sous la forme de la Profondeur Optique en Aérosols (Aerosol Optical Depth) (AOD) ou du coefficient d’Ångström notamment. Cependant ces variables ne sont pas directement liées à la concentration atmosphérique en PM<sub>2.5</sub> au niveau du sol, et l’obtention de cette concentration à partir de ces variables n’est pas aisée, comme cela a été précisé dans le chapitre 1. D’un autre côté, les réseaux de capteurs au sol de concentration en PM<sub>2.5</sub> ne peuvent fournir que quelques valeurs éparses. Obtenir des cartes complètes de la concentration en PM<sub>2.5</sub> à partir d’observations satellites est donc une tâche importante et intéressante, mais qui reste un challenge conséquent. Cela est dû au fait que la relation liant l’AOD et la concentration en PM<sub>2.5</sub> au niveau du sol est hautement multivariée.

Dans le chapitre 3, seul l’AOD est utilisé pour l’inférence de la concentration, ce qui amène à des résultats prometteurs, surclassant des méthodes telles que l’interpolation polynomiale ou encore l’algorithme Random Forest. Cependant et comme expliqué dans le chapitre 1, d’autres variables, connues pour leur impact sur la concentration atmosphérique en aérosols [23], peuvent être utilisées. Il s’agit des vitesse et direction du vent, des température, pression et humidité atmosphériques, et du coefficient d’Ångström. L’étude proposée dans ce chapitre traite de ces informations additionnelles et évalue l’intérêt de leur utilisation en termes de performance, en fonction de l’architecture du réseau utilisé ainsi que de la stratégie de fusion de l’information utilisée. En effet, l’utilisation d’informations supplémentaires en entrée d’un modèle appelle au choix d’une stratégie de fusion d’information particulière.

Les conclusions apportées par cette étude ont pour but d’aider à la conception d’un modèle efficace et inspiré par la connaissance. En se basant sur une analyse de performances, cette étude propose en effet une sélection d’architectures de réseaux adaptées à l’application. Notons cependant que l’objectif principal des travaux présentés dans ce chapitre n’est pas le développement d’un modèle optimisé pour une utilisation précise, mais plutôt d’investiguer si certaines architectures et stratégies de fusion sont adaptées à l’exploitation de l’information contenue dans des champs atmosphériques bi-dimensionnels à plusieurs composantes pour l’estimation de la concentration en aérosols.

Les principales contributions décrites dans ce chapitre peuvent être résumées ainsi :

- l’étude de l’intérêt de l’utilisation de plusieurs variables (coefficient d’Ångström, vitesse et direction du vent, température, pression, humidité) conjointement à

l’AOD pour la prédiction de la concentration en PM<sub>2.5</sub> au niveau de la surface

- une étude de différentes méthodes de fusion pour la prédiction de la concentration en PM<sub>2.5</sub> au niveau du sol, en fonction des variables utilisées en entrée et du type de modèle utilisé
- la proposition d’une architecture d’ANN ainsi que d’une sélection de variables à utiliser, basée sur la connaissance obtenue à la suite de ces études

Ces contributions sont décrites en section 4.1, tandis que leur implémentation ainsi que le plan d’expériences de l’étude réalisée sont détaillés en section 4.2. L’analyse des résultats de ces expériences se trouve en section 4.3. Enfin les déductions, nouvelles connaissances et conclusions obtenues à la suite de ces travaux sont résumés en section 4.4.

## 4.1 Proposition

Dans ce chapitre, l’utilisation de nouvelles variables d’entrée est proposée. Ces variables ont été présentées en section 1.4 : il s’agit du coefficient d’Ångström ainsi que des variables météorologiques mesurées au niveau du sol. Ces dernières sont l’humidité, la température, la pression, et enfin la vitesse et la direction du vent. Leur influence sur la concentration en aérosols dans l’atmosphère est reconnue et exprimée par différents modèles physiques [23].

Des modèles sont donc conçus avec pour objectif de produire en sortie des cartes (autrement dit des images géo-référencées) de la concentration en PM<sub>2.5</sub> au niveau du sol, à partir de données d’entrée constituées de cartes de l’AOD ainsi que de cartes d’autres variables atmosphériques. Dans ce cadre, des ANN convolutifs tels que les UNets et GANs sont utilisés, car leur capacité à prendre en compte la variabilité spatiale des données qui leur sont présentées n’est plus à prouver.

Un nombre important d’expériences est réalisé dans le but d’étudier l’impact de ces variables additionnelles sur la performance, et ce pour chaque architecture de réseau proposée. L’idée est notamment d’identifier les variables à utiliser. En effet l’issue attendue de ce grand nombre d’expériences est une architecture de réseau de neurones performante, ainsi que la proposition de lignes directrices pour la prédiction de la concentration en PM<sub>2.5</sub> à partir de ces variables.

En ce qui concerne les architectures, deux réseaux convolutifs sont utilisés : un

UNet et un GAN. L’apprentissage du UNet est purement supervisé, tandis que l’apprentissage du GAN est semi-supervisé, avec utilisation de vérité terrain éparses (aussi appelées CAL) : il s’agit en fait d’un BC-GAN, comme proposé dans le chapitre 3. Le UNet est lui aussi similaire à celui proposé dans le chapitre 3.

#### **4.1.1 Stratégies de fusion d’information**

Les modèles d’ANN utilisés ont en commun l’utilisation en tant que donnée d’entrée d’une ou plusieurs variables afin de produire le même type de sortie. Il est par conséquent nécessaire de fusionner ces variables durant l’inférence. Cela permet d’assurer que la production de la sortie du modèle utilise bien toutes les informations proposées en entrée.

Les stratégies de fusion sont par conséquent un aspect important de la conception, et dans un second temps de la performance de ces modèles. Elles représentent différentes méthodes qui peuvent être appliquées pour exploiter plusieurs sources de données (autrement dit plusieurs entrées) au sein du même réseau. Elles peuvent être appliquées aux UNets comme aux BC-GANs. Selon Mangai et al. [77], il existe trois stratégies de fusion principales : la fusion de données, la fusion de représentations, et la fusion de décisions. Les sections 4.1.2, 4.1.3 et 4.1.4 décrivent respectivement chacune de ces approches et la manière dont elles sont appliquées dans le contexte de ces travaux.

L’un des objectifs des expériences réalisées est également d’identifier la meilleure stratégie de fusion en fonction de l’architecture de réseau et des variables d’entrées utilisées.

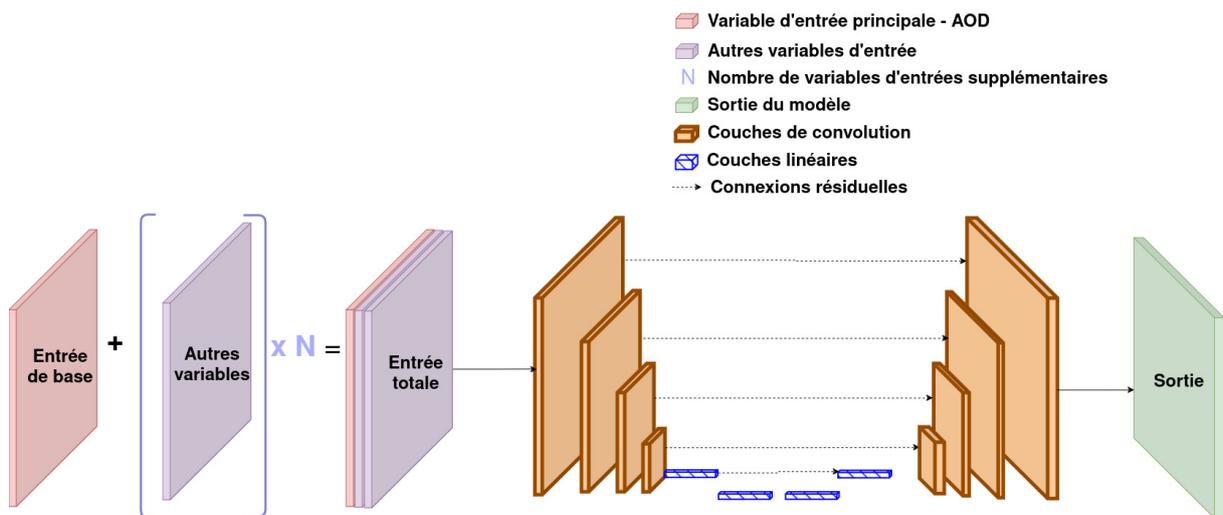
#### **4.1.2 Fusion de données / Concaténation de Canaux (CC)**

L’idée générale derrière cette stratégie est d’utiliser plusieurs données d’entrée différentes pour construire une nouvelle donnée d’entrée, plus grande, plus complète et possiblement plus utile.

Les données utilisées dans ce contexte étant des images, la manière la plus simple de mettre cette méthode en pratique est de réaliser ce qui est appelé ici de la Concaténation de Canaux (CC). Il s’agit alors de considérer ces différentes images comme des canaux différents d’une seule et même image. Cette approche est la plus simple et directe des trois.

En termes d’architecture des ANN utilisés, cela implique simplement l’utilisation de davantage de filtres de convolution. L’architecture du UNet avec fusion de données est illustré par la figure 4.1. L’architecture du discriminateur du GAN avec fusion de données est illustré par la figure A.1.

Cette architecture repose sur les relations locales (plutôt que globales) entre les différentes variables utilisées en entrée. Une hypothèse sur ses performances est qu’elle fonctionne mieux si les schémas de variation locaux dans une image correspondent à ceux dans une autre image.



**Figure 4.1:** Architecture du UNet avec l’approche de fusion de données. Elle correspond aussi à l’architecture du générateur du BC-GAN avec la même approche.

### 4.1.3 Fusion de Représentations (FR)

L’une des architectures les plus communes parmi les ANN pour les tâches de Vision Artificielle est l’encodeur-décodeur. L’idée générale est que l’encodeur produit ce qui est appelé une représentation latente (ou caractéristique). Celle-ci prend généralement la forme d’un vecteur (dit alors vecteur latent), mais peut techniquement aussi prendre celle d’une matrice (latente). L’idée est cependant, en général, que cette caractéristique doit être de dimension inférieure à celle de la donnée d’entrée qu’elle représente. Elle est supposée contenir toutes les informations pertinentes vis-à-vis de la tâche concernée contenues dans l’entrée. En d’autres termes, elle décrit la donnée d’entrée suffisamment bien pour être suffisante à l’accomplissement de la tâche. Durant l’étape suivante, le décodeur utilise cette représentation en entrée et produit la sortie du modèle.

L’idée derrière les méthodes de fusion de représentations est donc d’obtenir une

telle représentation pour chaque variable d’entrée, et de les fusionner (par exemple en les concaténant simplement) pour obtenir une super-caractéristique, capable de représenter l’ensemble des variables d’entrée. Ce type de méthode est souvent appliqué à des tâches de Vision Artificielle, par exemple lors de l’utilisation d’images hyperspectrales complexes en tant que données d’entrée [78]. D’après Sun et al. [79], deux manières intéressantes de fusionner des vecteurs latents sont la fusion en série (basée sur un vecteur union) et la fusion en parallèle (basée sur un vecteur complexe). Cependant les mêmes auteurs proposent une nouvelle méthode basée sur l’Analyse Canonique des Corrélations (ACC).

Quelle que soit la méthode utilisée pour cette fusion, la représentation unique ainsi obtenue est ensuite utilisée par le décodeur pour produire la sortie du modèle. En termes d’architecture, cela implique l’utilisation d’autant d’encodeurs que de variables d’entrée, mais d’un seul décodeur (car une seule sortie est désirée).

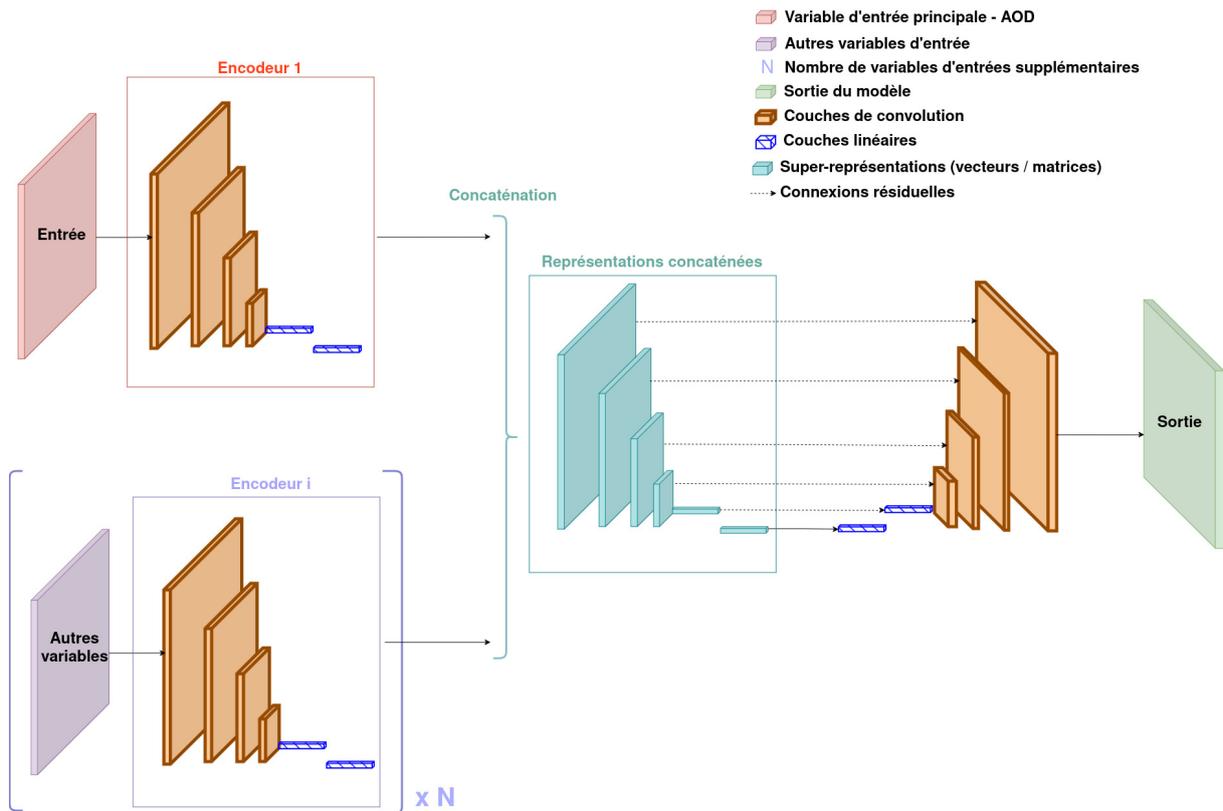
Cependant le UNet est un type particulier d’encodeur-décodeur, car il contient des connexions résiduelles. Il est de plus dans notre cas symétrique : les couches du décodeur reflètent celles de l’encodeur. Après chaque couche de l’encodeur, la caractéristique obtenue est envoyée à la couche correspondante dans le décodeur, grâce justement à ces connexions résiduelles. Cela permet au décodeur d’avoir accès à plusieurs représentations, à différents niveaux, au lieu d’avoir simplement accès à la plus petite.

Implémenter une stratégie de fusion avec un UNet n’est, par conséquent, pas trivial. Pour chaque variable d’entrée, des caractéristiques de différents niveaux sont obtenues et envoyées au décodeur. De cette observation découle une question : à quel niveau la fusion devrait-elle être réalisée ?

Ici, une méthode qui pourrait être appelée Fusion Multiple de Représentations est utilisée. Le principe de la Fusion de Représentation est appliqué à chaque niveau : une super-représentation, contenant des informations provenant de chacune des variables d’entrée, est donc obtenue pour chaque niveau. Ces représentations fusionnées sont alors envoyées au décodeur, notamment à travers les connexions résiduelles.

L’architecture du UNet avec fusion de représentation est illustrée par la figure 4.2. L’architecture du discriminateur du GAN avec fusion de représentation est illustré par la figure A.2.

Cette architecture compte davantage sur les relations globales entre les variables d’entrée. L’obtention de représentations exploite le concept de réduction de la dimen-



**Figure 4.2:** Architecture du UNet avec l’approche de fusion de représentations. Elle correspond aussi à l’architecture du générateur du BC-GAN avec la même approche.

sionalité. Par conséquent, plus le niveau considéré est profond, plus la caractéristique obtenue à ce niveau représente l’entrée de manière plus globale, et moins elle contient d’informations sur les variations locales. Cette architecture repose donc sur l’idée que chaque variable d’entrée contient (en plus de variations locales) une information globale, non localisée, qui peut être utile pour l’estimation de la concentration en particules fines.

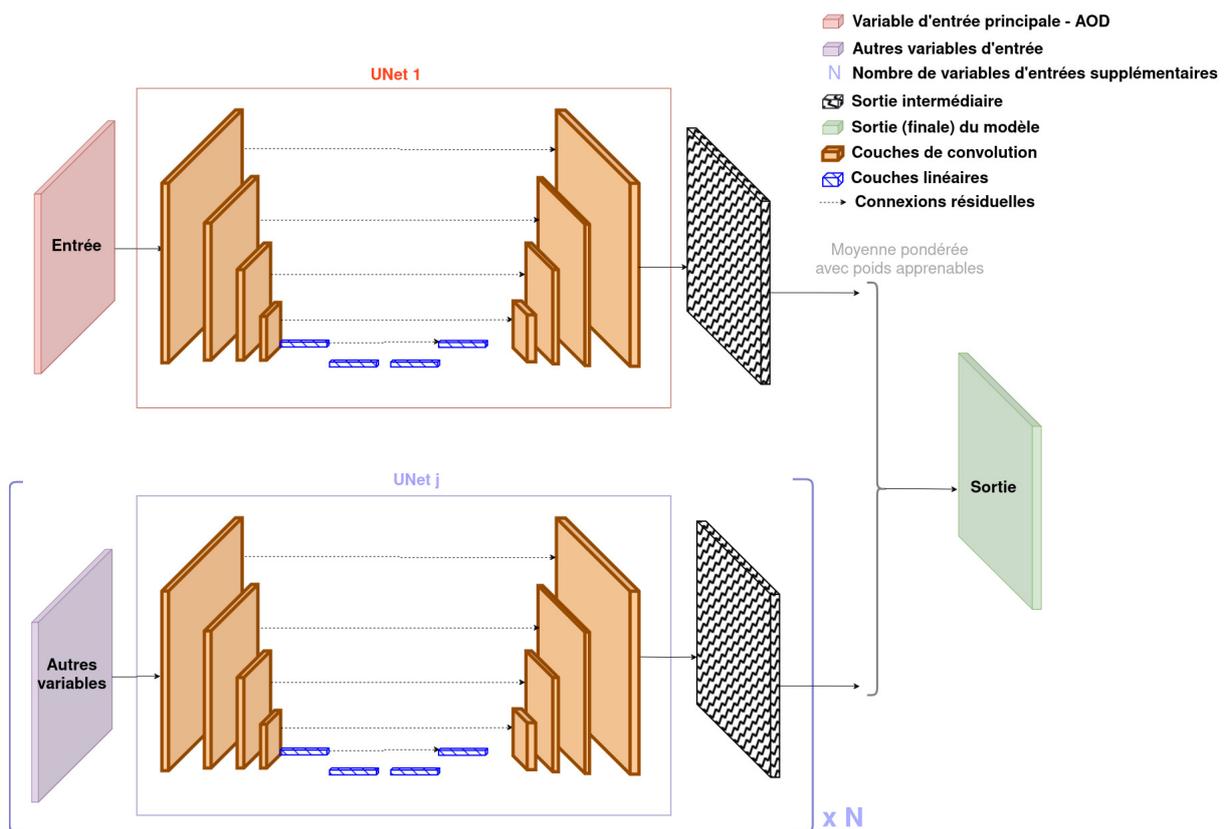
#### 4.1.4 Fusion de Décisions (FD)

L’idée derrière la fusion de décisions est d’utiliser un modèle différent pour chaque variable d’entrée, d’obtenir ainsi une sortie pour chacune de ces variables, et enfin de les fusionner pour obtenir une décision finale, supposément meilleure que les sorties individuelles. Pour les tâches de classification, la décision est la classe prédite. Pour les tâches de régression, comme celle considérée dans ces travaux, elle représente l’estimation proposée. Les valeurs des fonctions de coût sont ensuite calculées à partir de cette sortie finale. Le processus rétro-propagation a lieu dans le modèle entier (donc

dans chacun des modèles individuels qui le composent).

Il existe différentes manières de fusionner les décisions, comme les méthodes linéaires et logarithmiques [80], qui correspondent respectivement à une somme et à un produit pondérés. Des algorithmes de vote peuvent même être utilisés pour des tâches de classification [80]. Dans ces travaux, une méthode linéaire est utilisée. Plus précisément, une moyenne pondérée est appliquée aux sorties du modèle pour obtenir la décision finale. Les poids sont des paramètres apprenables, ce qui permet au modèle d’apprendre quelles sorties individuelles (et par extension quelles variables d’entrée) sont les plus pertinentes. Pour cette raison, cette approche est supposée obtenir de meilleurs résultats quand toutes les variables d’entrée sont utilisées.

La figure 4.3 illustre l’architecture d’un UNet avec Fusion de Décisions, tandis que l’architecture du discriminateur d’un GAN avec Fusion de Décisions est représentée par la figure A.3.



**Figure 4.3:** Architecture du UNet avec l’approche de fusion de décisions. Elle correspond aussi à l’architecture du générateur du BC-GAN avec la même approche.

Cette approche repose sur l’idée que chaque variable d’entrée peut individuellement

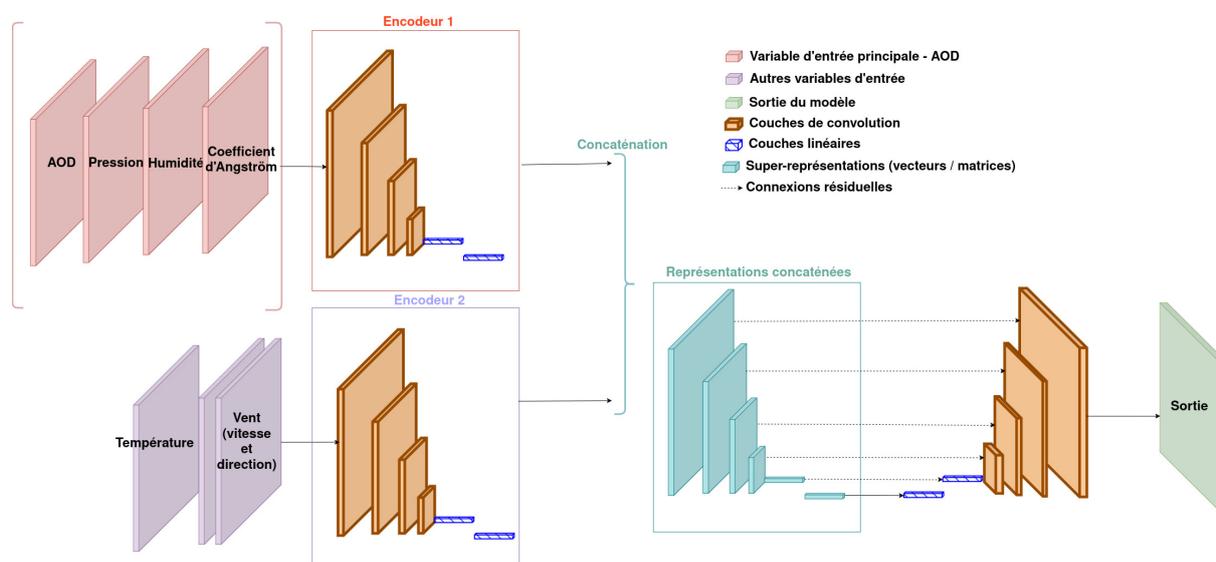
être utilisée pour produire une estimation de la concentration en aérosols, mais que cette estimation contient potentiellement des défauts. La meilleure estimation peut alors être obtenue par une combinaison (ici une moyenne pondérée) de ces estimations imparfaites. En d'autres termes, il est possible de corriger une estimation produite avec une variable d'entrée grâce la sortie obtenue avec une autre.

#### 4.1.5 Approches de fusion hybrides

Les variables pouvant être utilisées pour expliquer ou prédire la concentration en PM<sub>2.5</sub> au niveau du sol sont appelées ses prédicteurs. Leur nature physique a un impact évident sur la fonction non linéaire définissant la relation entre AOD et PM<sub>2.5</sub>. L'AOD est assez directement lié à la concentration en aérosols totale sur une colonne à un endroit donné, tandis que la pression de surface peut être indirectement reliée à la concentration en PM<sub>2.5</sub> à travers son accumulation dans la couche limite atmosphérique sous conditions stables. La vitesse du vent peut quant à elle influencer la concentration en PM<sub>2.5</sub> à grandes échelles d'espace et de temps. Nous pouvons donc distinguer des variables d'"état", capables de relier directement les PM<sub>2.5</sub> à l'AOD à travers une expression intégrale sur la colonne atmosphérique et l'utilisation de "prédicteurs indirects" agissant sur la concentration en PM<sub>2.5</sub> à différentes échelles spatio-temporelles. Dans notre analyse actuelle, les variables de vent (vitesse et direction) sortent du lot, car elles décrivent la dynamique de l'atmosphère. A l'inverse l'AOD et le coefficient d'Ångström sont clairement des variables d'état en ce qui concerne l'inférence de la concentration en PM<sub>2.5</sub>. Les variables d'Humidité, de Température et de Pression, peuvent être considérées comme étant principalement des variables d'état également, du fait de leur fort impact sur l'hygroscopicité des aérosols et par extension sur la distribution de la taille des particules qui les composent. Ils peuvent aussi influencer indirectement sur la concentration en PM<sub>2.5</sub> au niveau de la surface en favorisant l'accumulation de leurs particules sous des conditions atmosphériques stables ou au contraire leur élimination par les processus de déposition sèche ou de wet scavenging.

La performance des réseaux et leur robustesse vis-à-vis du bruit est bien entendu influencée par leur architecture. De plus, comme montré par les auteurs de [81], leurs performances peuvent être améliorées au-delà de l'état de l'art quand le réseau et son architecture sont bien alignés avec la fonction cible. Nous émettons ici l'hypothèse que pour des applications atmosphériques, l'alignement optimal de l'architecture du réseau avec la fonction cible peut dépendre de la nature des variables utilisées en entrée et de la stratégie employée pour fusionner les informations contenues dans

ces variables. Cette hypothèse amène un questionnement sur l’avantage d’appliquer différentes stratégies de fusion à différents types de variables d’entrée. A partir de cette observation, deux modèles hybrides sont proposés, utilisant différentes stratégies de fusion en fonction de la variable considérée. L’idée est d’utiliser la Fusion de Représentations (FR) pour le vent (vitesse et direction) et la température, et d’utiliser Fusion de données (ou Concaténation de Canaux (CC)) pour les autres variables (y compris l’AOD). Les figures 4.4 et 4.5 illustrent l’architecture de ces deux modèles. La différence entre ces deux modèles tient au nombre d’encodeurs utilisés. Ces approches de fusion hybrides ne sont implémentés qu’avec des UNet, et pas avec des BC-GANs.

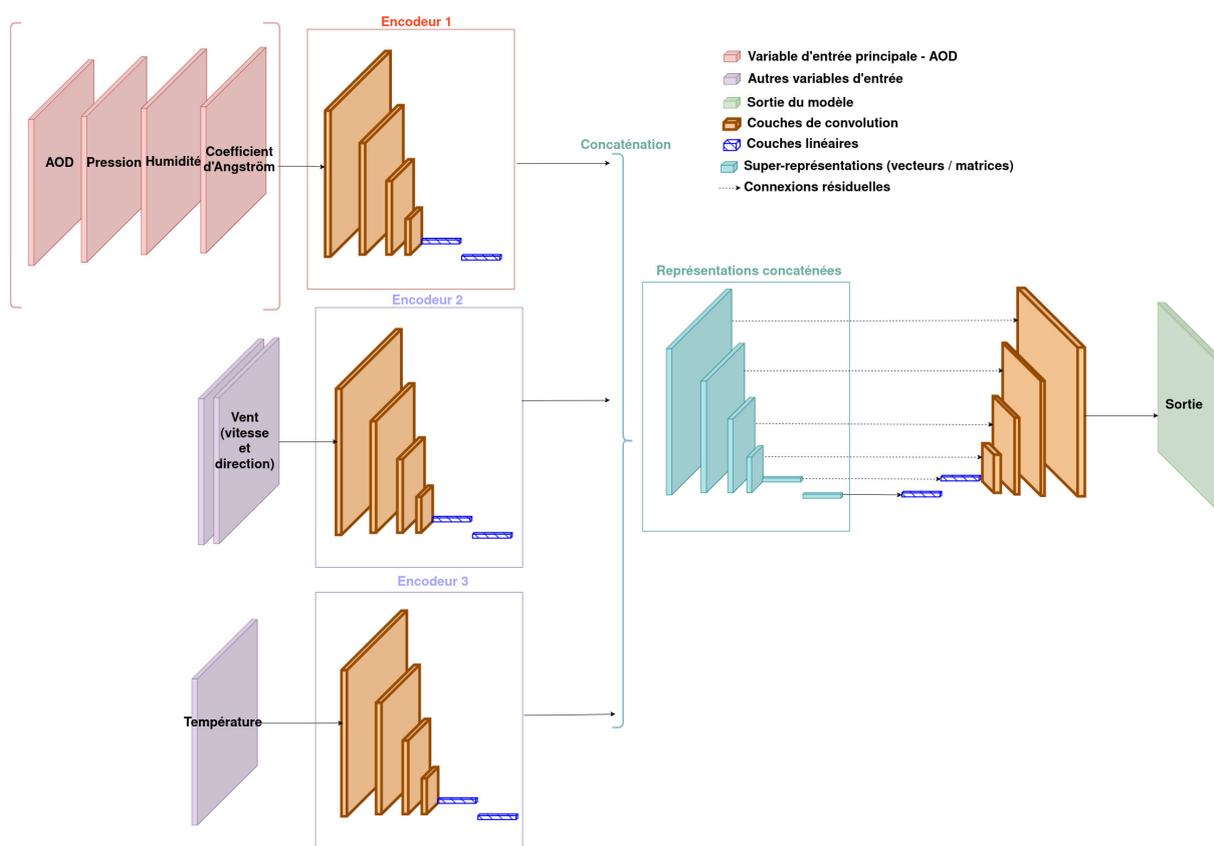


**Figure 4.4:** Le premier modèle hybride proposé, utilisant à la fois la fusion de données et la fusion de représentations.

## 4.2 Protocole expérimental

Dans le but de trouver la meilleure manière d’utiliser la variété de variables d’entrée disponibles, il est proposé d’étudier les trois principales stratégies de fusion (CC, FR, FD) décrites en section 4.1.1. Chaque approche est expérimentée sur deux types de modèles : un UNet purement supervisé, et un BC-GAN semi-supervisé utilisant des valeurs éparées de concentration en aérosols au niveau du sol comme CAL. Dans un second temps, des expériences sur les approches hybrides de fusion proposées en section 4.1.1 sont également réalisées.

Un deuxième objectif des expériences réalisées durant ces travaux est de comprendre quelles variables ont le plus fort impact sur nos résultats. En d’autres termes, quelles



**Figure 4.5:** Le deuxième modèle hybride proposé, utilisant à la fois la fusion de données et la fusion de représentations.

variables d’entrée aident les modèles conçus à mieux prédire la concentration en PM<sub>2.5</sub> au niveau du sol. C’est pourquoi, pour chaque architecture de modèle distincte (donc pour chaque couple type de modèle / stratégie de fusion), des expériences sont réalisées en utilisant différentes combinaisons de variables d’entrée additionnelles, dans le but d’étudier leur impact sur les résultats.

#### 4.2.1 Description des données

Les expériences sont conduites sur deux jeux de données distincts : l’un contient des données fournies par le modèle CAMS, et le second par le modèle ALADIN. Dans les deux cas, le type de scénario considéré est toujours le même, et il est illustré par la figure 4.6. Dans celui-ci, deux ensembles de données sont disponibles pour l’entraînement. Le premier, appelé ensemble d’entraînement complet, contient des vérités terrains complètes, et peut donc être utilisé dans le cadre d’un apprentissage strictement supervisé. Il couvre une période temporelle de 11 mois. Le deuxième, appelé ensemble d’entraînement éparé, ne couvre une période que d’un mois, et ne

contient que des vérités terrain éparses. Il ne peut donc être utilisé que dans un contexte d’apprentissage semi-supervisé. Ces vérités terrain éparses sont alors utilisées en tant que CAL, comme suggéré dans les travaux décrits dans le chapitre 3. Dans le contexte de ces travaux, elles contiennent toujours une quantité de pixels équivalente à 5% du nombre total de pixels de l’image, et ces 5% sont répartis aléatoirement. Pour reprendre les notations du chapitre 3, chaque fois que des CAL sont utilisées, leur configuration est **Aléatoire 5%**.



**Figure 4.6:** Représentation des jeux de données utilisés pour nos expériences. Le texte en noir spécifie les périodes temporelles correspondantes pour chaque jeu de données, lorsque les données sont fournies par le modèle CAMS. Le texte en gris fournit la même information, mais pour le cas où les données proviennent du modèle ALADIN.

Pour le jeu de données CAMS, un échantillon est généré toutes les trois heures. Ces échantillons prennent la forme de matrices de taille 241x480. En fonction du nombre de variables d’entrées choisi, chaque entrée de modèle peut être composée d’une à sept de ces matrices. Cela est d’ailleurs vrai pour les deux jeux de données. Pour en revenir au jeu de données CAMS, l’ensemble d’entraînement complet contient 2680 échantillons, l’ensemble d’entraînement épars en contient 240, et l’ensemble de test 2920.

En ce qui concerne le jeu de données ALADIN, un échantillon est généré toutes les heures. Les matrices qui composent ces échantillons sont de taille 405x613. L’ensemble d’entraînement complet contient 8040 échantillons et l’ensemble d’entraînement épars en contient 720. Lors de la création de l’ensemble de test, seule une image toutes les trois heures est utilisée afin qu’il contienne le même nombre d’échantillons que l’ensemble de test issu de CAMS, c’est-à-dire 2920.

#### 4.2.2 Pré-traitement des données

Les modèles météorologiques présentés en section 1.5 peuvent fournir des valeurs pour la majorité des variables d’intérêt présentées en section 1.4. Pour rendre l’utilisation de certaines de ces variables météorologiques plus aisée, un pré-traitement est nécessaire.

Les valeurs de pression sont converties de Pascal à atmosphères, et les valeurs de température de Kelvin à degrés Celsius. Les valeurs d’humidité ne nécessitent aucun pré-traitement particulier.

Les données disponibles contiennent des valeurs de vitesse du vent en direction de l’Est ainsi qu’en direction du Nord. Plutôt que d’utiliser ces valeurs telles qu’elles sont présentées, la transformation décrite par l’équation 4.1 leur est appliquée. Dans cette équation,  $U$  et  $V$  représentent respectivement la vitesse du vent en direction de l’Est et du Nord. À l’issue de cette transformation, deux matrices sont obtenues. La première contient des valeurs de vitesse du vent absolue. La deuxième contient des valeurs de direction du vent, en degrés, par rapport au Nord. Chaque fois qu’une expérience fait usage des "valeurs du vent", ces deux matrices sont en réalité utilisées.

$$\begin{aligned} \text{norme} &= \sqrt{U^2 + V^2} \\ \text{direction} &= \arctan\left(\frac{V}{U}\right) \cdot \left(\frac{180}{\pi}\right) \end{aligned} \quad (4.1)$$

Ces mêmes données ne contiennent, à l’origine, pas de valeurs du coefficient d’Ångström. Cependant il est possible de les calculer aisément à partir de valeurs de l’AOD fournies à différentes longueurs d’onde, de la manière décrite par l’équation 1.1. Dans cette équation,  $\lambda_1$  représente la longueur d’onde de l’AOD original (celui utilisé comme variable d’entrée principale). Elle est de  $550nm$  dans toutes nos expériences, quelle que soit la source des données.  $\lambda_2$  représente la longueur d’onde d’un second AOD, utilisé uniquement pour le calcul du coefficient d’Ångström. Le modèle ALADIN nous fournit, en plus de l’AOD original à  $550nm$ , un autre AOD à  $1000nm$  : c’est donc à cette longueur d’onde que correspondra  $\lambda_2$  pour le modèle ALADIN. Le modèle CAMS quant à lui, fournit des valeurs de l’AOD à différentes longueurs d’onde, mais pas à  $1000nm$ . La longueur d’onde la plus proche exploitable avec les données du modèle CAMS est de  $865nm$ .  $\lambda_2$  vaudra donc  $865nm$  pour les expériences réalisées avec ces données.

Comme dans le chapitre précédent, des valeurs d’AOD et de concentration en  $PM_{2.5}$  sont bien entendues utilisées. Elles subissent le pré-traitement décrit en section 3.2.2.

### 4.2.3 Description des expériences

Dans toutes les expériences réalisées, l’AOD est systématiquement utilisé en entrée. Il s’agit de la donnée d’entrée principale, et ne fait pas partie des variables d’entrée "additionnelle". Celles-ci sont en effet composées de l’humidité relative, de la pression, de la température, des vitesse et direction du vent, ainsi que du coefficient d’Ångström. L’un des objectifs de ces expériences, comme expliqué précédemment, est de comprendre, voir d’évaluer, l’impact de ces valeurs sur les résultats. Il est donc nécessaire d’utiliser différentes combinaisons de variables d’entrée pour atteindre ce but. Quelle que soit la combinaison de variables considérée, les variables de vitesse et de direction du vent sont toujours utilisées ensemble : soit les deux variables sont utilisées, soit aucune des deux. A cette exception près, des expériences sont réalisées avec chaque combinaison possible de ces variables (incluant l’utilisation d’aucune et de toutes ces variables). Cette ensemble d’expérience est reproduit avec chaque architecture distincte proposée, autrement dit avec chaque couple type de modèle et stratégie de fusion.

L’étude exhaustive décrite dans le paragraphe précédent est conduite sur le jeu de données issu du modèle CAMS. Elle nécessite la réalisation de 192 expériences différentes. Seules les expérimentations menant aux meilleurs résultats sont reproduites sur le jeu de données fourni par le modèle ALADIN. Cela fournira ainsi des connaissances sur l’impact sur les résultats des caractéristiques de chaque ensemble de données.

Dans un second temps, des expériences sont réalisées sur les approches hybrides. Ces expériences n’utilise que des modèles UNet (donc pas de BC-GANs). De plus, toutes les variables d’entrée y sont chaque fois utilisées. Ces expériences sont conduites sur les deux ensembles de données.

### 4.2.4 Métrique supplémentaire : la FSIM

Pour l’évaluation de nos modèles, les métriques décrites en section 3.3 sont utilisées. Cependant la figure 3.6 ainsi que la section 3.5 ont mis en évidence le besoin d’utiliser une métrique additionnelle. Celle-ci doit être représentative de la conservation par le modèle des formes et structures fines de l’image. La meilleure candidate pour cela est une métrique appelée index de SIMilarité de Caractéristiques ou de Représentations (Feature SIMilarity index) (FSIM), exprimée en %.

Cette métrique a été proposée par Zhang et al. [82], en tant que métrique d’évaluation de la qualité d’une image. Elle repose sur les concepts de congruence de phase et de magnitude de gradient (au sens du gradient d’image). La congruence de phase est également utilisée pour calculer le poids de chaque pixel dans le calcul de la similarité entre les deux images (prédiction et vérité terrain). Cela amène à donner un poids significatif aux bords, formes et autres structures dans l’image. Cette métrique permet donc une meilleure représentation de la conservation par le modèle de ces formes, bords et structures.

En plus d’être utilisée comme métrique par tous les modèles, elle est également utilisée comme fonction de coût additionnelle (en plus de la traditionnelle  $MSE$ ) par les UNets utilisés dans ces travaux. La section A.3 montre l’efficacité de cette méthode.

#### 4.2.5 Scores

Le grand nombre de métriques utilisées peut complexifier la comparaison entre modèles (ou entre expériences). Afin de la faciliter, trois scores différents, calculés à partir des métriques décrites en sections 3.3 et 4.2.4, sont utilisés.

- **Score total** : ce score est calculé à partir de toutes les métriques, temps d’inférence inclus. Son calcul suit l’équation 4.2.

$$Score\ total = \frac{\frac{0.05s - \text{Temps d'inférence}}{0.05s} + (1 - rMAE) + \frac{3-QE}{3} + (1 - rMBE) + (1 - FSIM)}{5} \quad (4.2)$$

- **Score intemporel** : ce score est très similaire au premier, mais il exclut le temps d’inférence. Cela permet d’identifier les modèles qui se montreraient les plus performants dans un contexte où le temps d’inférence n’est pas un facteur prédominant. Son calcul suit l’équation 4.3.

$$Score\ intemporel = \frac{(1 - rMAE) + \frac{3-QE}{3} + (1 - rMBE) + (1 - FSIM)}{4} \quad (4.3)$$

- **Score réduit** : ce score est calculé en utilisant exclusivement la  $rMAE$  et la  $FSIM$ . Parmi les quatre métriques utilisées pour le calcul du **score intemporel**, ces deux-là sont les plus pertinentes pour le domaine de la Vision Artificielle. Ce score permet donc une comparaison des modèles et expériences de ce point de vue, tandis que les autres scores mettent davantage en avant les besoins et

objectifs de la communauté de physique atmosphérique. Il n’utilise pas non plus le temps d’inférence, et son calcul suit l’équation 4.4.

$$Score\ réduit = \frac{(1 - rMAE) + (1 - FSIM)}{2} \quad (4.4)$$

#### 4.2.6 Disponibilité du code et des données

Le code utilisé pour les expériences présentées dans ce chapitre est disponible dans une archive Zenodo.<sup>1</sup>

Les données issues du modèle CAMS utilisées pour ces expériences sont disponibles dans une autre archive Zenodo.<sup>2</sup>

Enfin les données issues du modèle ALADIN sont extraites d’un jeu de données librement accessible.<sup>3</sup>

### 4.3 Résultats

Dans un premier temps, un aperçu global des résultats est proposé. Cet aperçu permet notamment d’aboutir à quelques conclusions sur les choix d’architecture et les variables à utiliser. L’étape suivante est l’identification des meilleurs résultats parmi les expériences réalisées sur le jeu de données CAMS. Cela permet alors de reproduire ces expériences sur le jeu de données ALADIN. La comparaison de ces résultats permet de mieux comprendre l’impact des caractéristiques (notamment le domaine géographique et la résolution) de ces deux jeux de données sur les résultats.

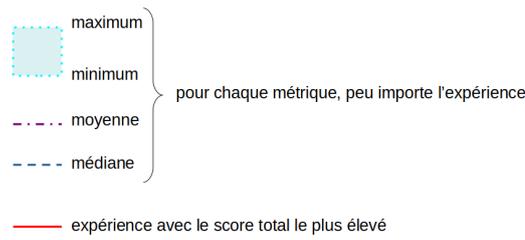
#### 4.3.1 Aperçu

Un grand nombre d’expériences ayant été réalisées, leurs résultats sont résumés sous la forme de graphes radar. Cinq métriques sont représentées sur ces graphes : le temps d’inférence  $t$ , la  $rMAE$ , la  $rMBE$ , la  $QE$  et la  $FSIM$ . Toutes ces métriques sont exprimées en % pour faciliter la construction des graphes. De plus ces métriques ont été inversées pour faciliter l’interprétation : elles sont donc à maximiser. La figure 4.7 présente la légende commune pour tous ces graphes.

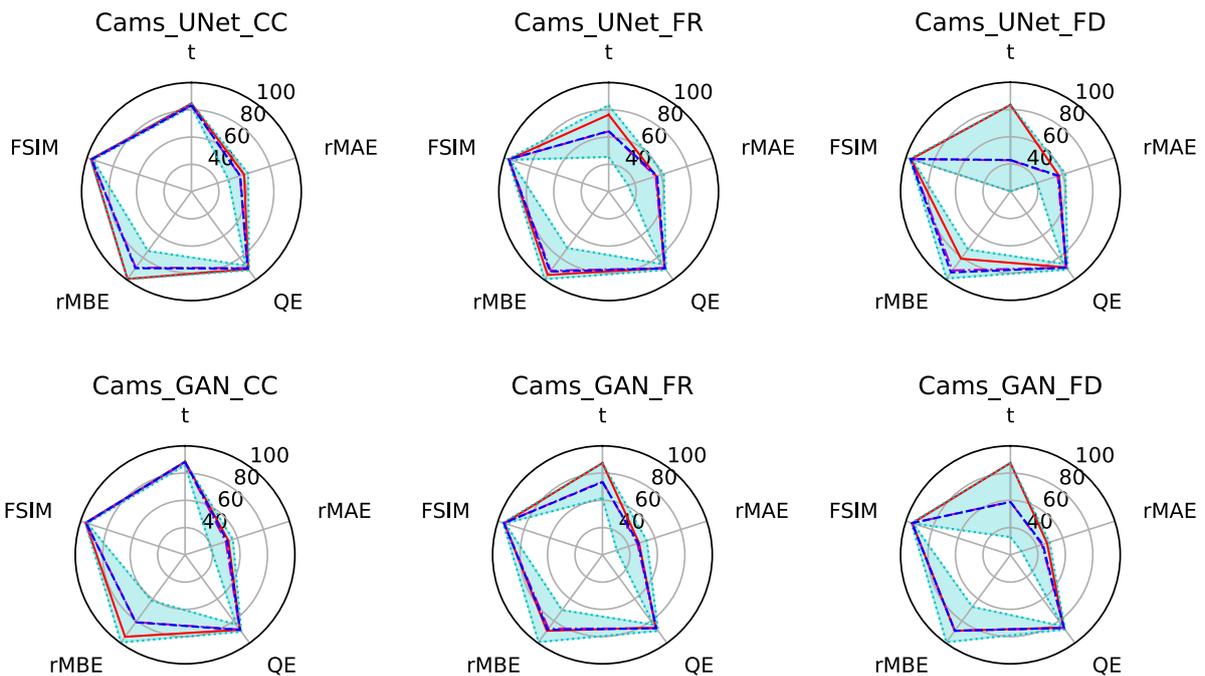
<sup>1</sup><https://doi.org/10.5281/zenodo.13947256>

<sup>2</sup><https://doi.org/10.5281/zenodo.13929498>

<sup>3</sup><https://doi.org/10.25326/703>



**Figure 4.7:** Légende commune à tous les graphes radar.



**Figure 4.8:** Aperçu de la performance de chaque couple modèle / stratégie de fusion. Les noms des graphes donnent des informations sur le jeu de données utilisé, le modèle, et le type de fusion. Pour rappel les types de fusion sont Concaténation de Canaux (CC), Fusion de Représentations (FR), et Fusion de Décisions (FD).

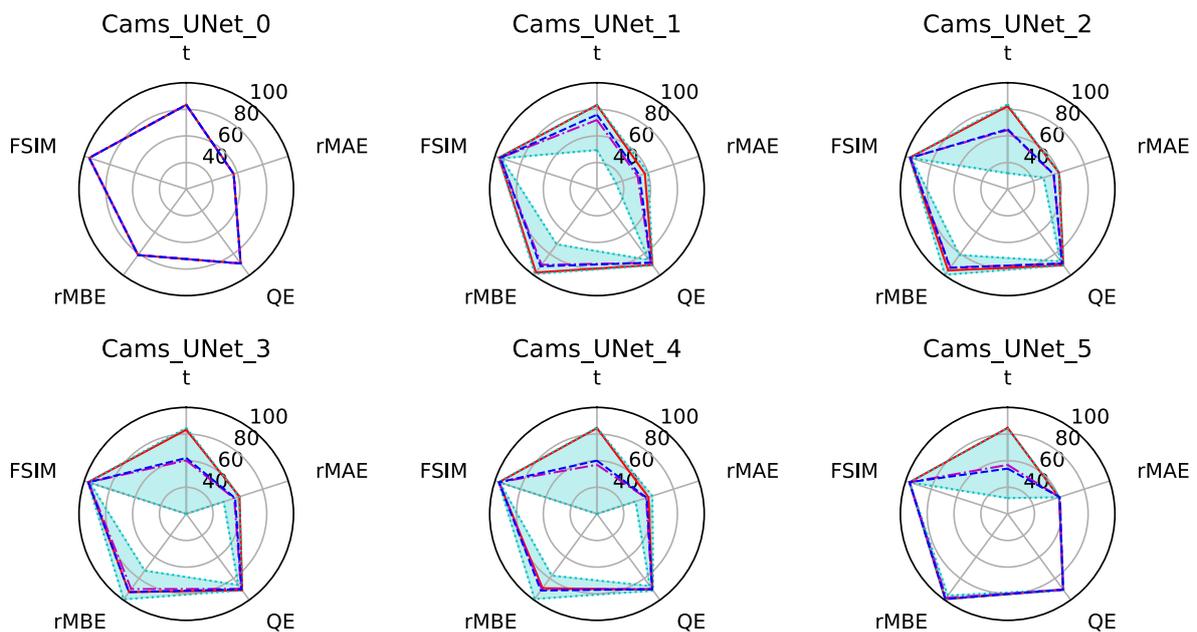
La figure 4.8 donne quant à elle un aperçu de la performance de chaque couple modèle - stratégie de fusion. Elle montre qu’en moyenne, les modèles utilisant la Fusion de Décisions (FD) ont les plus longs temps d’inférence, tandis que les modèles utilisant la Concaténation de Canaux (CC) sont les plus rapides. Cette observation n’est pas surprenante, car elle est corrélée avec le nombre de paramètres des modèles, illustré par la figure A.4 en section A.4.

Il apparaît aussi que les BC-GANs souffrent généralement de moins bonnes performances en  $rMAE$  et en  $rMBE$  que les UNets. L’interprétation proposée est que

la proportion de l’ensemble d’entraînement complet, réservé à l’apprentissage purement supervisé, sur l’ensemble d’entraînement total (composé des ensembles complet et épars) est suffisamment importante pour que les méthodes purement supervisées soient performantes. Inversement, la part de l’ensemble épars est faible. Or l’intérêt des GANs dans ce contexte réside justement dans leur capacité à réaliser de l’apprentissage semi-supervisé, donc d’utiliser l’ensemble d’entraînement épars. Cela rend donc l’intérêt et la pertinence de l’utilisation des GANs (voire de toute méthode d’apprentissage semi-supervisé) dans ce contexte plus limités. Par ailleurs, dans le chapitre 3, l’ensemble épars représente la moitié de l’ensemble total, et les BC-GANs présentent de meilleurs résultats que les UNets.

Il est intéressant de noter que les métriques *QE* et *FSIM* semblent impactées moins intensément par ce phénomène.

Enfin, la stratégie de Concaténation de Canaux (CC) semble mener à des résultats plus stables que les deux autres stratégies. Cela peut encore une fois être lié à la différence de complexité des modèles, comme montré en section A.4.



**Figure 4.9:** Aperçu de l’évolution de la performances des UNets lorsque le nombre de variables d’entrées augmente. Le vent est compté comme une seule variable bien qu’il contienne deux canaux (un pour la vitesse et un second pour sa direction), car ces deux canaux sont toujours utilisés ensemble.

La figure 4.9 montre l’évolution des performances des UNets lorsque le nombre de variables d’entrée augmente. Sans surprise, lorsque ce nombre augmente, le temps d’inférence augmente (ce qui est témoigné par une diminution de la variable  $t$  sur la figure). Cela est attendu, puisque la figure A.4 montre que la complexité des modèles augmente avec la quantité de variables d’entrée. A l’inverse les autres métriques, et en particulier la  $rMAE$  montrent que la performance des modèles a tendance à augmenter avec le nombre de variables d’entrée. De plus les meilleurs résultats (du point de vue de ces autres métriques) sont généralement produits par les méthodes utilisant toutes ces variables, et ce quelque soit la stratégie de fusion utilisée.

Cependant cette augmentation de la performance n’est pas linéaire. Ainsi pour une expérience individuelle donnée l’ajout d’une variable d’entrée supplémentaire ne garantit pas l’obtention de meilleurs résultats. La comparaison entre les expériences utilisant deux et trois variables d’entrée additionnelles montre qu’avec trois, les valeurs de  $rMBE$  sont moins stables (donc moins prévisibles). Comme le nombre d’expériences dans ces deux cas est le même, ce n’est donc pas ce qui cause cette différence.

### 4.3.2 Meilleurs résultats

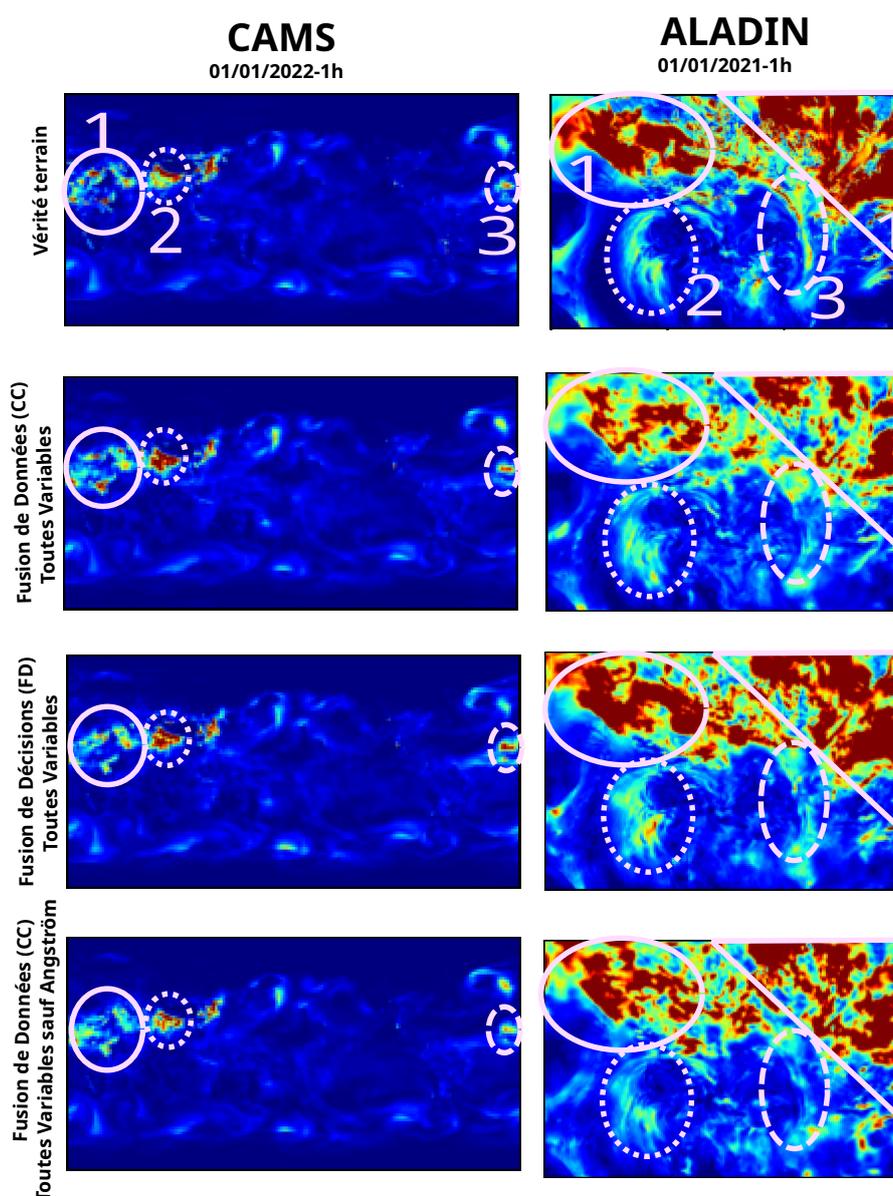
La table 4.1 montre les meilleurs résultats selon chacun des trois scores décrits en section 4.2.5. La première ligne correspond au meilleur Score Total, la deuxième au meilleur Score Intemporel, et la troisième au meilleur Score Réduit. La figure 4.10 montre les sorties de ces modèles pour un échantillon donné.

**Table 4.1:** Meilleurs résultats sur les données CAMS. Dans la colonne "Vars" figure l’initiale de chaque variable utilisée : V pour Vent (vitesse et direction), H pour Humidité, T pour Température, A pour Coefficient d’Ångström. L’AOD est toujours utilisé comme donnée d’entrée.

Modèle	Fusion	Vars	t	MAE	QE	MBE	FSIM	Scores (✓)		
								Total	Intemporel	Réduit
UNet	CC	VHPTA	0.0077	4.38	0.29	-0.13	3.10%	86.28%	86.7%	78.61%
UNet	FD	VHPTA	0.0341	4.33	0.29	0.05	2.95%	75.8%	86.81%	78.93%
UNet	CC	VHPT	0.0086	4.04	0.26	-2	2.8%	83.38%	83.52%	80.33%

Cette table montre qu’utiliser toutes les variables disponibles amène généralement aux meilleurs résultats, excepté en ce qui concerne le Coefficient d’Ångström sur la dernière ligne. Cela montre aussi que la méthode de Fusion de Décisions souffre de temps d’inférence sensiblement plus longs, spécialement comparé à la stratégie de Concaténation de Canaux.

Ces résultats amènent à trois recommandations principales dépendant du contexte et des performances désirées. Si la *MBE* (autrement dit le biais) n’est pas un facteur important, alors le modèle recommandé est un UNet avec la stratégie de Concaténation de Canaux, utilisant toutes les variables d’entrée disponibles à l’exception du Coefficient d’Ångström. Si le temps d’inférence n’est pas un facteur important, alors l’utilisation d’un UNet avec la stratégie de Fusion de Décisions et utilisant toutes les variables disponibles est conseillé. Enfin, le modèle aboutissant aux résultats les plus équilibrés est un UNet avec la stratégie de Concaténation de Canaux, utilisant toutes les variables.



**Figure 4.10:** Sorties produites par les modèles présentés en table 4.1, pour les deux jeux de données, et pour un échantillon donné à chaque fois.

La colonne gauche de la figure 4.10 donne un aperçu plus détaillé des résultats avec les données CAMS. Pour cet échantillon, il semblerait que le modèle utilisant la Concaténation de Canaux ainsi que toutes les variables à l’exception du coefficient d’Ångström soit celui fournissant la meilleure prédiction pour la zone n°2 de l’image. Les autres modèles semblent surestimer la concentration en aérosols dans cette région. Le fait que ce modèle ait la pire *MBE* et que celle-ci soit négative, montre une tendance à la sous-estimation. Cette dernière hypothèse et les observations proposées plus haut dans ce paragraphe sont mutuellement cohérentes.

Le modèle utilisant la même stratégie de fusion mais utilisant toutes les variables (y compris le coefficient d’Ångström) fournit quant à lui de bonnes prédictions pour les zones n°1 et 3 de l’image. Enfin le modèle utilisant la stratégie de FD (et toutes les variables) sous-estime la concentration dans la région n°1 et la surestime dans les zones n°2 et 3. Ce dernier modèle a la meilleure *MBE*, mais pas la meilleure *MAE*. Une hypothèse proposée est qu’il surestime la concentration dans certaines zones et la sous-estime ailleurs, ce qui induit une compensation des biais locaux et aboutit à un biais global relativement faible, et ce bien qu’un autre modèle le surclasse en ce qui concerne la *MAE*.

La table 4.2 montre les résultats obtenus par les mêmes modèles, mais appliqués cette fois aux données ALADIN. La colonne droite de la figure 4.10 montre les sorties de ces modèles, sur ce jeu de données, pour un échantillon donné.

**Table 4.2:** Résultats sur les données ALADIN. Dans la colonne "Vars" figure l’initiale de chaque variable utilisée : V pour Vent (vitesse et direction), H pour Humidité, T pour Température, A pour Coefficient d’Ångström. L’AOD est toujours utilisé comme donnée d’entrée.

Modèle	Fusion	Vars	t	MAE	QE	MBE	FSIM	Scores (↗)		
								Total	Intemporel	Réduit
UNet	CC	VHPTA	0.0064	8.37	0.35	-3.92	4.37%	84.03%	83.24%	80.02%
UNet	FD	VHPTA	0.0364	<b>8.29</b>	<b>0.34</b>	1.65	<b>3.97%</b>	73.41%	84.96%	79.99%
UNet	CC	VHPT	0.0075	8.76	0.37	<b>-1.04</b>	4.64%	<b>85.39%</b>	<b>85.49%</b>	78.79%

Cette table montre une différence importante entre les résultats obtenus avec les données CAMS et ALADIN. Cependant, bien que les mêmes métriques soient utilisées dans les deux cas, les résultats obtenus avec l’un des jeux de données ne sont pas aisément comparables avec ceux obtenus avec le second. En effet, le jeu de données issu du modèle ALADIN contient des images de bien plus haute résolution que celles du jeu de données CAMS, ces images ne représentent pas le même domaine géographique (l’Europe pour ALADIN, le monde pour CAMS), et ces jeux de données ne correspondent pas à la même période temporelle (Juillet 2020 à Juin 2022 pour CAMS,

Juillet 2019 à Juin 2021 pour ALADIN). Cela explique d’ailleurs les différences dans les valeurs de concentration en aérosols dans les deux jeux de données. En effet pour les données CAMS, les valeurs de concentration sont comprises entre 0 et 34 425  $\mu\text{g}/\text{m}^3$  avec une moyenne de 11.02  $\mu\text{g}/\text{m}^3$ , tandis que pour les données ALADIN, elles sont comprises entre 0 et 6 774  $\mu\text{g}/\text{m}^3$  avec une moyenne de 23.17  $\mu\text{g}/\text{m}^3$ .

La figure 4.10 montre également une différence dans les résultats obtenus avec les données CAMS et ALADIN. Le modèle utilisant la stratégie de Concaténation de Canaux avec toutes les variables sous-estime la concentration en aérosols dans les régions n°1 et 3. Cela est consistant avec le fait que ce modèle a la plus basse *MBE* parmi les trois. Le modèle utilisant la même stratégie de fusion mais sans le coefficient d’Ångström sous-estime également la concentration dans ces mêmes zones, mais commet une erreur moins importante. Ce modèle a également la deuxième *MBE* la plus basse, ce qui est cohérent avec ces observations. Enfin le modèle utilisant la stratégie de Fusion de Décisions et toutes les variables sous-estime également la concentration dans la région n°3 mais pas dans la n°1, et la surestime dans la zone n°2. Des trois modèles, c’est celui avec la *MBE* la plus élevée, ainsi que la meilleure *MAE*.

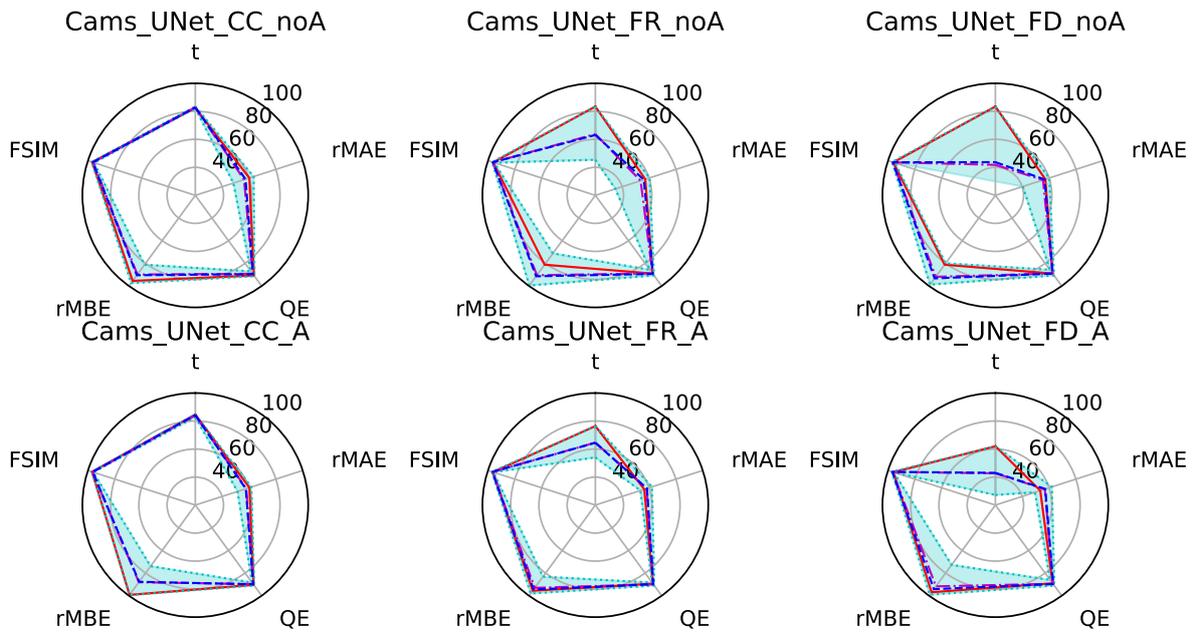
La comparaison entre les résultats issus des deux jeux de données reste intéressante. En effet les méthodes menant aux meilleures performances avec les données CAMS n’ont pas la même supériorité relative lorsqu’elles sont appliquées aux données ALADIN.

L’attention du lecteur est dirigée vers la table 4.1, et en particulier vers les résultats obtenus avec la méthode de Concaténation de Canaux. L’un de ces résultats est obtenu en utilisant toutes les variables en entrée, et l’autre en les utilisant toutes sauf le coefficient d’Ångström. En se basant ces deux seuls résultats obtenus avec les données CAMS, il semble pertinent de supposer que l’utilisation du coefficient d’Ångström en tant que variable d’entrée additionnelle amène à une meilleure *MBE*, mais à de moins bonnes performances selon toutes les autres métriques (à l’exception du temps d’inférence). Cependant, comme montré par la table 4.2, les résultats obtenus avec les mêmes modèles et méthodes appliqués aux données ALADIN semble mener à supposer l’inverse. Autrement dit, à supposer que l’utilisation du coefficient d’Ångström amène à de meilleures performances pour toutes les métriques excepté la *MBE*.

Cela montre que l’impact de l’utilisation d’une variable d’entrée individuelle sur les résultats des modèles proposés ne peut pas être aisément interprété. Cela est dû à l’interaction entre les variables d’entrée elles-mêmes, et la nature de boîte noire

des ANN.

L’analyse de la figure 4.11 peut permettre de tenter de comprendre l’impact du coefficient d’Ångström sur les résultats obtenus sur le jeu de données issu du modèle CAMS. Cette figure montre que les deux métriques les plus impactées par le



**Figure 4.11:** Aperçu des résultats des expériences avec (première ligne) et sans (seconde ligne) l’utilisation du coefficient d’Ångström en tant que variable d’entrée. Les noms des graphes donnent des informations sur le jeu de données utilisé, le modèle, et le type de fusion. Pour rappel les types de fusion sont Concaténation de Canaux (CC), Fusion de Représentations (FR), et Fusion de Décisions (FD).

coefficient d’Ångström sont la *MAE* et la *MBE* (et leurs versions relatives, la *rMAE* et la *rMBE*). L’utilisation de cette variable d’entrée semble mener à une valeur de *MAE* minimale plus élevée. En d’autres termes, cela aide à éviter l’obtention des pires résultats (au regard de la *MAE*). Les meilleures valeurs de *MBE* sont obtenues avec le coefficient d’Ångström en tant que variable d’entrée. Son utilisation semble donc mener à un biais plus bas.

Encore une fois, ces observations sont valides en ce qui concerne le jeu de données CAMS, et pour la période temporelle considérée. Aucune conclusion générale sur l’utilisation du coefficient d’Ångström en tant que variable d’entrée ne peut être faite à partir de ces seules observations. En particulier, celles-ci sont cohérentes avec les résultats montrés en table 4.1 (obtenus avec les données CAMS), mais pas avec

ceux de la table 4.2 (obtenus avec les données ALADIN). Ces observations au sujet du coefficient d'Ångström pour le jeu de données CAMS ne peuvent donc pas être automatiquement supposées vraies pour le jeu de données ALADIN aussi.

### 4.3.3 Résultats des méthodes hybrides de fusion

Les tables 4.3 et 4.4 présentent les résultats obtenus avec les deux modèles hybrides décrits en section 4.1.5, pour les données CAMS et ALADIN, respectivement. La figure 4.12 montre les sorties produites par ces modèles pour un échantillon donné. Toutes les variables d'entrée disponibles sont utilisées par ces deux modèles. D'après ces résultats,

**Table 4.3:** Résultats des modèles hybrides avec le jeu de données CAMS. La colonne "Fusion" contient "Hybride 1" pour le modèle représenté par la figure 4.4, et "Hybride 2" pour celui de la figure 4.5.

Modèle	Fusion	t	MAE	QE	MBE	FSIM	Scores (↗)		
							Total	Intemporel	Réduit
UNet	Hybride 1	<b>0.0098</b>	4.39	0.28	<b>0.2</b>	2.96%	<b>85.25%</b>	<b>86.46%</b>	78.67%
UNet	Hybride 2	0.0116	<b>4.1</b>	<b>0.27</b>	-0.86	<b>2.9%</b>	84.09%	85.91%	<b>80.03%</b>

et du point de vue de la Vision Artificielle, le deuxième modèle hybride est meilleur. Cependant le premier est plus équilibré, et est recommandé dans toute situation où le biais (autrement dit la *MBE*) et le temps d'inférence sont des métriques importantes.

La performance de ces modèles, bien que relativement satisfaisante, l'est moins que les résultats présentés en table 4.1. Par conséquent l'utilisation de ces modèles avec les données CAMS n'est pas recommandée.

La colonne gauche de la figure 4.12 montre que les deux modèles produisent une estimation de la concentration dans la zone n°3 plutôt adéquate, la sous-estiment dans les régions n°1 et la surestiment dans la zone n°2. Cela peut s'expliquer par le fait que ces deux modèles ont des performances assez proches pour la plupart des métriques, et ont tous deux une *MBE* proche de zéro.

**Table 4.4:** Résultats des modèles hybrides avec le jeu de données ALADIN. La colonne "Fusion" contient "Hybride 1" pour le modèle représenté par la figure 4.4, et "Hybride 2" pour celui de la figure 4.5.

Modèle	Fusion	t	MAE	QE	MBE	FSIM	Scores (↗)		
							Total	Intemporel	Réduit
UNet	Hybride 1	0.0108	<b>7.69</b>	<b>0.33</b>	<b>-1.42</b>	<b>4%</b>	<b>85.21%</b>	<b>86.91%</b>	<b>81.52%</b>
UNet	Hybride 2	<b>0.0103</b>	7.9	<b>0.33</b>	-2.87	4.2%	83.91%	85.04%	81.05%

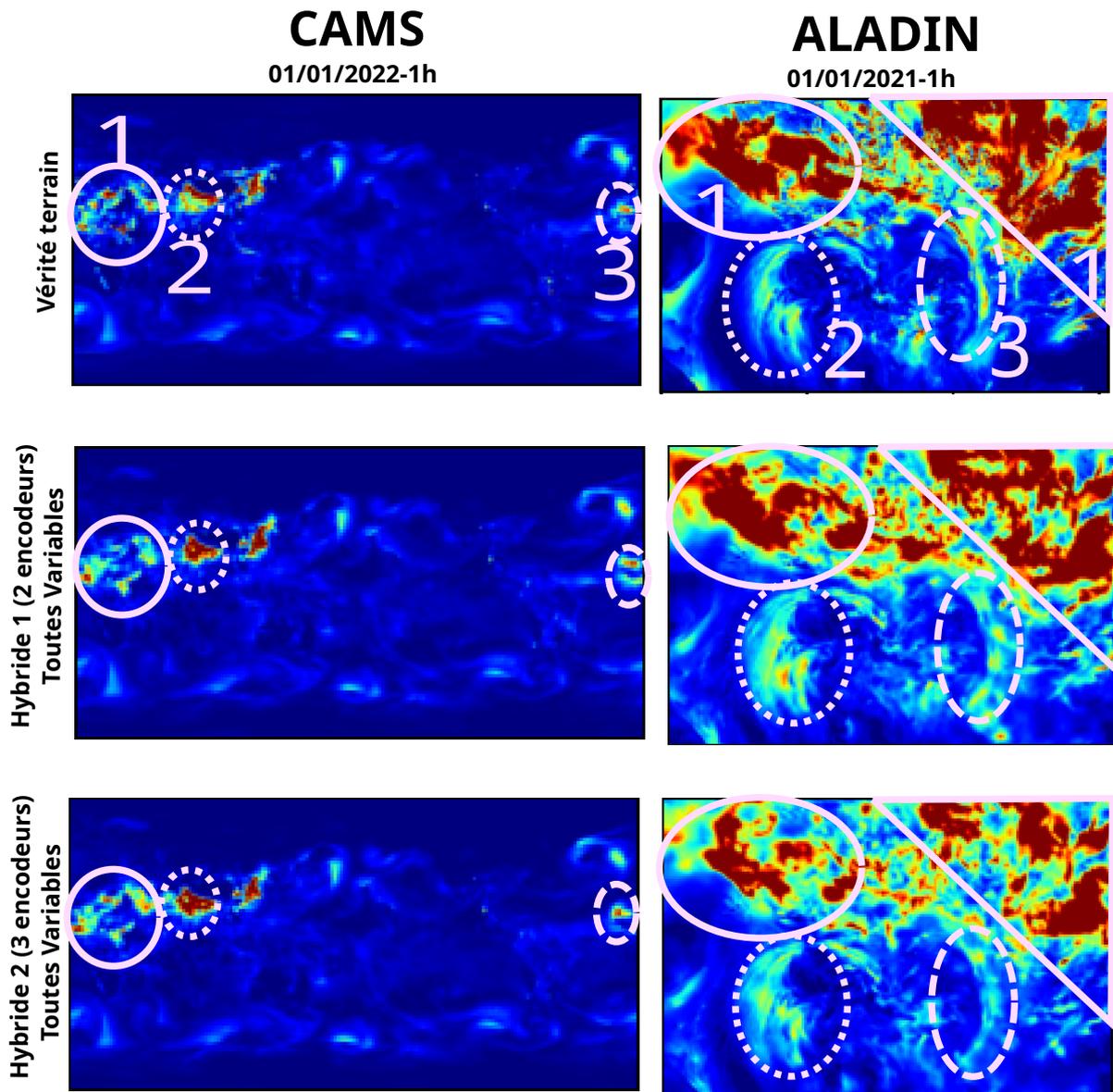


Figure 4.12: Sorties produites par les modèles hybrides, pour les deux jeux de données, et pour un échantillon donné à chaque fois.

Ces résultats montrent qu’avec les données ALADIN, le premier modèle hybride amène à de meilleures performances que le second, sur toutes les métriques à l’exception du temps d’inférence.

Les performances de ce modèle surpassent également celles présentées en table 4.2. Cependant, les modèles présentés dans cette table sont ceux ayant produit les meilleurs résultats avec le jeu de données CAMS. Par conséquent, il n’est pas possible de déduire de ces seuls résultats que ce modèle hybride est le plus performant sur le

jeu de données ALADIN. Pour arriver à une telle conclusion, une étude exhaustive serait nécessaire. Une telle étude devrait prendre en compte les trois principales stratégies de fusion, les deux modèles implémentés (GANs et UNets) ainsi que les six variables d’entrée disponibles.

La colonne droite de la figure 4.12 montre que le premier modèle hybride sous-estime légèrement la concentration dans les zones n°1, et le deuxième la sous-estime davantage. Le premier modèle hybride surestime légèrement la concentration en région n°2, tandis que le second fournit une estimation plus proche de la réalité. Ces observations peuvent s’expliquer par le fait que ces deux modèles ont une *MBE* relativement faible, et le second a la plus basse des deux. Il est aussi intéressant de noter que les deux modèles hybrides proposent une meilleure estimation de la concentration dans la zone n°3 que les modèles présentés en figure 4.10.

#### 4.3.4 Comparaison avec les méthodes proposées comme base de comparaison

La table 4.5 montre une comparaison des meilleurs résultats présentés dans cette section avec les méthodes proposées en section 2.2, le tout sur le jeu de données CAMS.

Les temps d’inférence de l’algorithme Random Forest et de la méthode du krigeage sont trop élevés pour permettre le calcul du score total avec la méthode proposée en section 4.2.5 : cela amènerait à un score négatif. Nous choisissons donc de considérer leurs temps d’inférence comme étant de 0.05s, le plus haut temps d’inférence pris en compte par le calcul du score total tel qu’exprimé en section 4.2.5.

Il est important de noter que les algorithmes Random Forest et d’Interpolation Polynômiale utilisent exclusivement l’AOD comme donnée d’entrée. Par ailleurs la méthode du krigeage utilise comme entrée exclusivement des valeurs éparses de la concentration en aérosol.

La méthode d’Interpolation Polynômiale a un temps d’inférence significativement plus court que toute autre méthode présentée dans ce chapitre. Cependant il s’agit de la seule métrique sur laquelle l’une des méthodes de la section 2.2 surpasse les résultats présentés dans cette section. En effet, à cette exception près, ces méthodes sont surpassées sur toutes les métriques et tous les scores par chacun des modèles proposés dans ce chapitre, et ce avec une marge significative.

**Table 4.5:** Comparaison des méthodes présentées en section 2.2 avec les meilleurs résultats sur le jeu de données CAMS, présentés en table 4.1. "Interp. Poly. Deg. 3" signifie "Interpolation Polynômiale de degré 3". "Krig. Ord. + Variog. ET" signifie "Krigage Ordinaire avec Variogramme à Effets de Trous".

Modèle	Fusion	Vars	t	MAE	QE	MBE	FSIM	Scores (↗)		
								Total	Intemporel	Réduit
Interp. Poly. Deg. 3		AOD seul	<b>0.0007</b>	6	0.41	-3	4.96%	79.62%	74.87%	70.34%
Random Forest		AOD seul	1.1535	6.01	0.41	-2.63	4.95%	60.56%	75.7%	70.33%
Ord. Krig. + Variog. ET		CAL seules	40.3732	6.03	0.38	-2.99	7.67%	62.41%	78.01%	73.51%
UNet	CC	VHPTA	0.0077	4.38	0.29	-0.13	3.10%	<b>86.28%</b>	86.7%	78.61%
UNet	FD	VHPTA	0.0341	4.33	0.29	<b>0.05</b>	2.95%	75.8%	<b>86.81%</b>	78.93%
UNet	CC	VHPT	0.0086	<b>4.04</b>	<b>0.26</b>	-2	<b>2.8%</b>	83.38%	83.52%	<b>80.33%</b>

## 4.4 Conclusion

Dans ce chapitre, une étude extensive sur l'utilisation de plusieurs variables météorologiques et mesures optiques en tant qu'entrées pour un ANN est proposée. Elle inclut également l'utilisation de trois stratégies de fusion différentes pour l'exploitation de ces variables d'entrée, dans le but de trouver la meilleure manière de fusionner toutes les informations utilisées. Des méthodes de fusion hybrides sont aussi proposées, implémentées et étudiées. Des expériences ont également été réalisées sur un nouveau jeu de données, issu du modèle ALADIN (au lieu de CAMS), afin d'étudier l'impact des caractéristiques de ces ensembles de données sur les résultats.

Les expériences ainsi réalisées montrent que, dans le contexte considéré, les UNets surclassent les GANs. Cependant il est suggéré que cela est strictement dû à une caractéristique des jeux de données utilisés. En effet cette caractéristique réduit le besoin, l'utilité et la pertinence de l'apprentissage semi-supervisé. Il est donc également suggéré que dans un contexte différent, une conclusion opposée aurait pu être atteinte.

Elles démontrent aussi l'intérêt d'utiliser davantage de variables d'entrée que l'AOD seul, quand celles-ci sont sélectionnées avec soin. En ce qui concerne celles choisies pour les expériences réalisées dans ce chapitre, plus le nombre de ces variables d'entrées utilisées est grand, plus les résultats tendent à s'améliorer. Cependant cela reste une tendance et pas une garantie, notamment car des exceptions peuvent être observées dans des cas spécifiques, dépendant de l'architecture et de la stratégie de fusion utilisées. En particulier la table 4.1 présente un cas dans lequel l'utilisation du coefficient d'Ångström comme variable d'entrée supplémentaire amène à une moins bonne performance selon toutes les métriques à l'exception du biais. Cette tendance suit la croyance générale selon laquelle, en Apprentissage Profond, utiliser davantage de données (lorsqu'elles sont pertinentes) amène à l'obtention de meilleurs résultats.

A l’inverse, ces exceptions s’en écartent, et soulignent l’intérêt d’étudier l’impact des architectures de réseaux de neurones pour les applications atmosphériques.

Ces expériences montrent aussi l’importance du jeu de données choisi et de ses caractéristiques (en l’occurrence les couverture et résolution spatiales). En effet cela impacte non seulement les résultats mais aussi les déductions et conclusions qui peuvent en être tirées. Cela est d’autant plus important en sciences atmosphériques car les variables géophysiques ont différentes échelles de variabilité. Or les architectures de réseaux doivent idéalement être alignées avec les caractéristiques spatiales des variables d’entrée qu’ils utilisent. Ces observations suggèrent donc que davantage de travaux sont nécessaires pour comprendre l’impact de l’architecture des réseaux sur leur capacité à exploiter les caractéristiques spatiales spécifiques aux sciences de l’atmosphère dans leur totalité.

Bien qu’il serait utile de pouvoir identifier précisément l’impact de chaque variable sur les performances, les observations issues des expériences proposées dans ce chapitre mettent en lumière la difficulté d’une telle tâche.

Les deux stratégies de fusion amenant aux meilleurs résultats sont la Fusion de Données (ou Concaténation de Canaux) et la Fusion de Décisions. D’après les expériences présentées dans ce chapitre, cette première méthode semble produire des résultats plus stables. Elle permet par ailleurs l’implémentation de modèles plus petits, ce qui amène donc à un temps d’inférence, et donc un temps d’apprentissage, plus courts. De plus ces architectures démontrent des performances en matière d’inférence de la concentration en PM<sub>2.5</sub> comparables à des modèles spécialisés trouvés dans la littérature [17, 23]. Nous suggérons aussi que la performances des modèles proposés ici pourrait être améliorée en augmentant la couverture temporelle de l’ensemble d’entraînement.

Les expériences réalisées sur les modèles de fusion hybride ne montrent aucune preuve claire de leur avantage comparé aux autres modèles, malgré des performances comparables.

Par conséquent, la stratégie de Fusion de Données est celle à recommander dans le cas général. Bien entendu, cette recommandation dépend du contexte général et plus précisément de la définition de la solution désirée. Par exemple, utiliser d’autres métriques pour mesurer la performance pourrait mener à une recommandation différente.

Ce chapitre a démontré l’intérêt d’utiliser un certain nombre de variables d’entrée additionnelles, du point de vue des performances du modèle. Ces variables sont

soigneusement choisies. En effet différentes études en physique atmosphérique [23] les désignent comme ayant un impact important sur la concentration en aérosols (qui constitue la variable estimée par les modèles proposés). Le chapitre 3 a quant à lui montré l’intérêt de fournir au modèle des informations a priori sous forme de CAL, souvent utiles pour des modèles physiques. De plus ce chapitre a également montré, dans le contexte qu’il propose, l’intérêt de l’apprentissage semi-supervisé.

La suite logique de ces observations est la conception d’un modèle capable d’utiliser différentes variables d’entrée, connues pour leur impact sur la concentration en aérosols. Ce modèle doit aussi être capable de prendre en compte des informations a priori sur la physique du problème. Autrement dit, des informations typiquement utiles pour résoudre ce type de problème par une méthode purement physique. Enfin, la capacité à réaliser de l’apprentissage semi-supervisé serait également un avantage. Or ces caractéristiques sont typiques d’un Réseaux Informés par la Physique (Physics-Informed Neural Network) (PINN). La suite logique des travaux proposés dans les chapitres 3 et 4 est donc la conception d’un PINN adapté au contexte décrit dans le chapitre 2.

# 5

## Modèle informé par la physique pour la prédiction de la concentration en aérosols au niveau du sol

### Sommaire

---

5.1	Proposition . . . . .	113
5.2	Protocole expérimental . . . . .	116
5.2.1	Données . . . . .	116
5.2.2	Pré-traitement des données . . . . .	117
5.2.3	Evaluation des modèles . . . . .	117
5.2.4	Conditions Aux Limites : configurations et méthodes . . . . .	118
5.2.5	Disponibilité du code et des données . . . . .	119
5.3	Résultats . . . . .	119
5.3.1	Comparaison à l'état de l'art . . . . .	119
5.3.2	Impact de la fonction de coût informée par la physique . . . . .	120
5.3.3	Impact des CAL . . . . .	121
5.3.4	Impact de la proportion d'entraînement non supervisé . . . . .	123

Chapitre 5 – Modèle informé par la physique pour la prédiction de la concentration en  
aérosols au niveau du sol

5.3.5	Comparaison visuelle . . . . .	125
5.4	Conclusion . . . . .	<b>125</b>

---

Les Réseaux de Neurones Convolutifs Informés par la Physique (Physics-Informed Convolutional Neural Networks) (PICNN) sont des algorithmes nés de l'intersection de deux champs de recherche : la Vision Artificielle et les Réseaux Informés par la Physique (Physics-Informed Neural Network) (PINN). Ils ont à la fois la capacité d'apprendre les dépendances et motifs locaux dans une image ou une matrice, et d'utiliser des informations *a priori* sur la physique du problème pour guider l'apprentissage.

Comme expliqué dans le chapitre 1, les variables atmosphériques auxquelles nous nous intéressons sont souvent représentées sous forme de matrices. Celles-ci représentent les valeurs de ces variables sur un domaine géographique défini, qui peut d'ailleurs correspondre à la Terre entière. L'analyse de ce type de matrices peut justement révéler des motifs ou structures locales. Cela rend donc l'utilisation de techniques de Vision Artificielle pour ce type de problème particulièrement prometteur. Cela reste vrai bien que, dans le cadre de ces travaux, les principaux objets d'intérêts sont des panaches d'aérosols, qui sont des objets extrêmement déformables comparés aux objets considérés dans des jeux de données plus traditionnels [37, 36, 38].

En ce qui concerne la prédiction de valeurs de concentration en aérosols, plusieurs variables sont pertinentes. Les valeurs de certaines propriétés optiques (comme l'AOD ou le coefficient d'Ångström) sont utiles pour obtenir la concentration en Particules de Matière ayant un diamètre inférieur à  $2.5 \mu m$  (PM<sub>2.5</sub>). Cependant certains travaux montrent l'intérêt de l'utilisation d'autres variables également [23, 18, 17]. Les principaux modèles de transport suggèrent l'importance de considérer le vent ainsi que les termes des sources et de pertes d'aérosols [8].

Cependant il n'est pas aisé d'obtenir ces termes de sources et de pertes, ce qui impacte négativement notre capacité à utiliser ces modèles de transport pour construire une fonction de coût (et par extension un réseau de neurones) basée sur la physique. La disponibilité de ces valeurs est similaire à la disponibilité des vérités terrain utilisées durant l'entraînement. Or les PINN sont particulièrement utiles dans le contexte de l'apprentissage semi-supervisé, lorsqu'une fonction de coût basée sur la physique peut être utilisée pour compenser le manque de vérités terrains. Le fait que ces termes de sources et de pertes soient nécessaire à la conception d'une fonction de coût basée sur la physique, ainsi que leur manque de disponibilité, constituent un obstacle scientifique. Le besoin d'une solution permettant de dépasser celui-ci est donc réel.

De plus, au vu du grand nombre de variables d'intérêt dans les problèmes physiques,

il est raisonnable de supposer que ce type d'obstacle n'est pas nécessairement rare dans la communauté de sciences atmosphériques. Une solution permettant de dépasser cet obstacle précis pourrait donc potentiellement ouvrir la voie et permettre d'en surmonter d'autres du même type.

Les propositions de ce chapitre sont résumées de la manière suivante :

- la conception d'une nouvelle architecture permettant au modèle de prédire non seulement la concentration en aérosols, mais également les termes de sources et de pertes, ainsi qu'une preuve empirique de son efficacité
- une fonction de coût informée par la physique, spécifiquement conçue pour répondre au problème présenté dans ce chapitre, guidant l'apprentissage semi-supervisé du modèle proposé
- une étude sur l'impact de différents paramètres sur ce modèle : quantité de vérités terrain complètes disponibles, utilisation et configuration des CAL, utilisation de la fonction de coût informée par la physique

La section 5.1 décrit en détail le modèle proposé dans ce chapitre. Le protocole expérimental employé pour évaluer les performances de ce modèle est exposé en section 5.2. Les résultats de ces évaluations sont présentés dans la section 5.3. Enfin la section 5.4 résume les déductions et conclusions les plus intéressantes issues de ces travaux.

## 5.1 Proposition

Le but de ces travaux est de concevoir un PINN pour la prédiction de la concentration en PM<sub>2.5</sub> dans les basses couches de l'atmosphère, à échelle globale. Son aspect informé par la physique permet au modèle de surclasser les modèles et méthodes de l'état de l'art dans un contexte semi-supervisé.

Plus précisément, l'entraînement purement supervisé est réalisable uniquement sur une portion de notre jeu de données. Sur le reste, un entraînement non supervisé (ou guidé par la physique) est possible. La figure 5.3 illustre ces caractéristiques de notre jeu de données. Les figures 5.1 et 5.2 montrent respectivement l'entraînement supervisé et non supervisé du modèle. Le modèle proposé est capable de surmonter l'obstacle du manque de disponibilité des termes de sources et de pertes d'aérosols. Il le fait en fournissant des estimations de ces termes ainsi que des estimations de la concentration

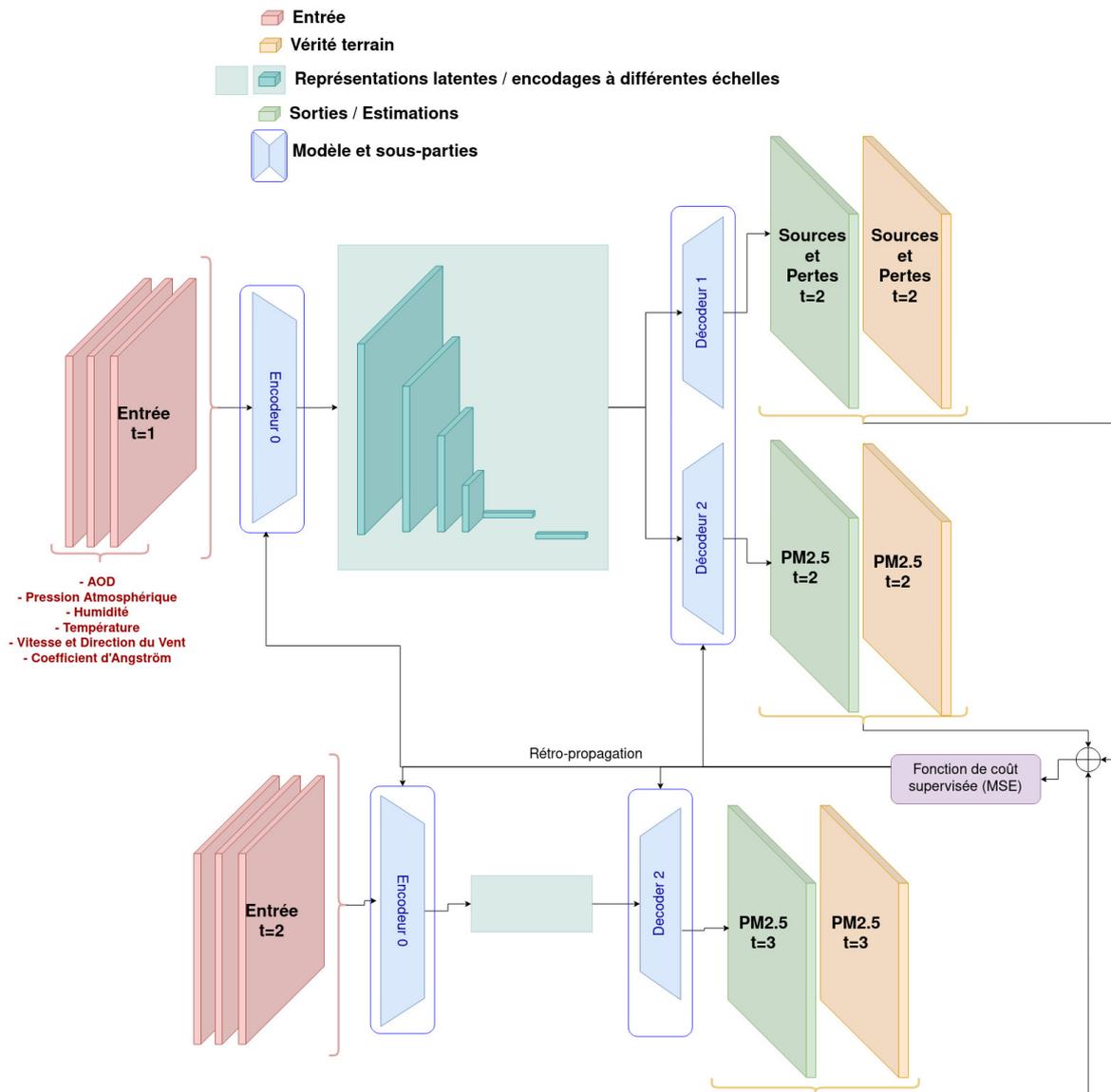
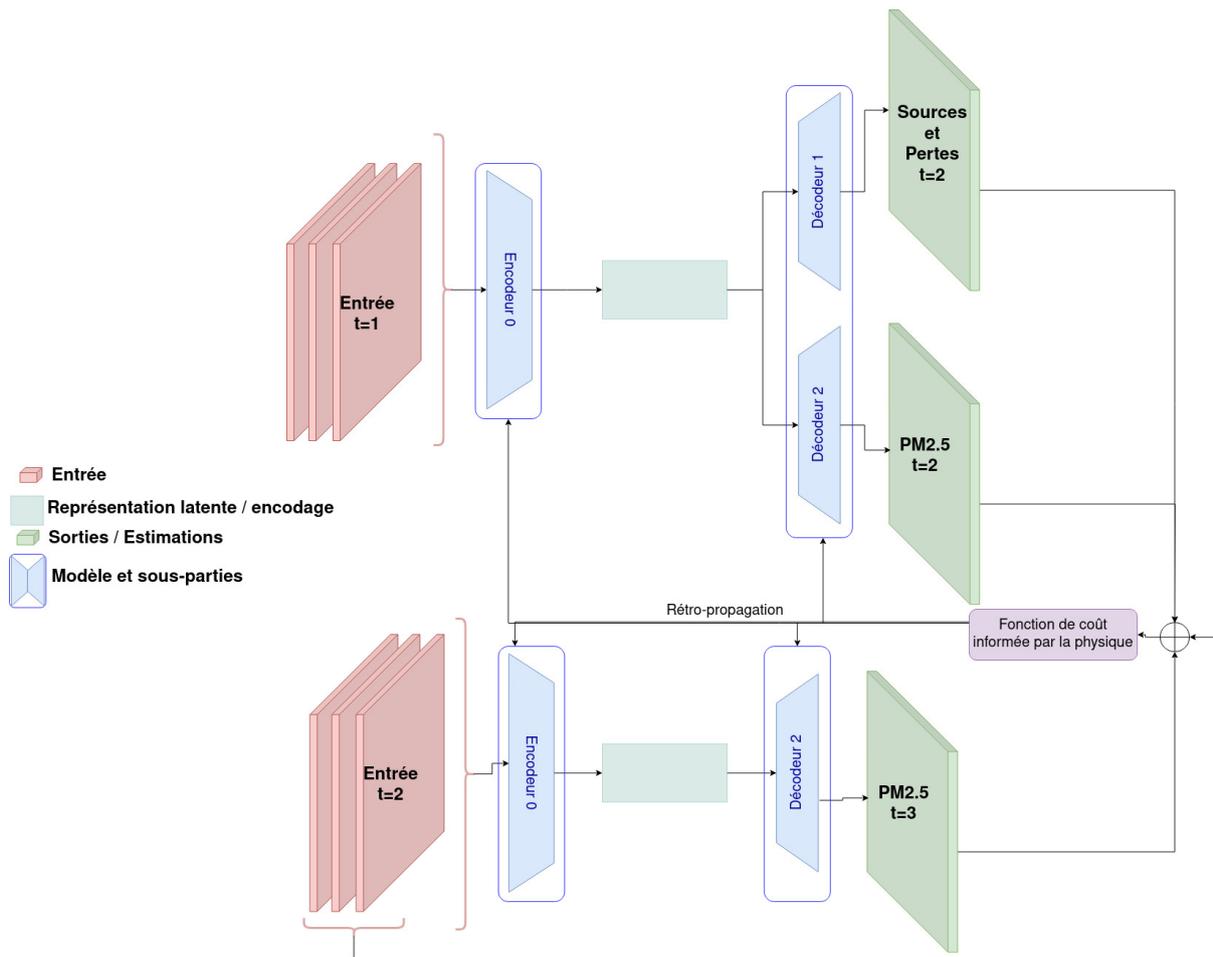


Figure 5.1: Illustration de l'entraînement supervisé de notre modèle.

en PM<sub>2.5</sub> (qui reste la principale variable d'intérêt). Cela est nécessaire car ces termes apparaissent dans le modèle physique à l'origine de la fonction de coût informée par la physique. Ce modèle est décrit par l'équation 5.1 [8]. Dans cette équation,  $C$  représente la concentration en aérosols,  $t$  le temps,  $K_m$  est un coefficient (considéré constant ici) de diffusion des aérosols, et  $\vec{V}$  est le vecteur du vent (qui contient des informations à la fois sur sa vitesse et sa direction).  $S$  et  $P$  sont respectivement les termes de sources et de pertes d'aérosols. Notons que par soucis de simplicité afin de ne pas nécessiter l'utilisation de champs météorologiques en trois dimensions et respecter la géométrie du problème de conversion image à image, nous avons considéré ici uniquement



**Figure 5.2:** Illustration de l’entraînement non supervisé (guidé par la physique) de notre modèle.

l’équation de dispersion en deux dimensions dans la couche proche de la surface.

$$\frac{\partial C}{\partial t} + \text{div}(C \vec{V}) - \text{div}(K_m \overrightarrow{\text{grad}}(C)) - (S - P) = 0 \quad (5.1)$$

Comme une dérivée de la concentration en aérosols en fonction du temps apparaît dans cette équation, il est nécessaire d’utiliser des valeurs de cette concentration à deux pas de temps différents. Cela permet en effet de calculer cette dérivée, et par extension la valeur de la fonction de coût. C’est pour cette raison que le modèle proposé fournit durant l’entraînement des estimations de la concentration en PM2.5 aux pas de temps deux et trois, comme illustré sur les figures 5.1 et 5.2.

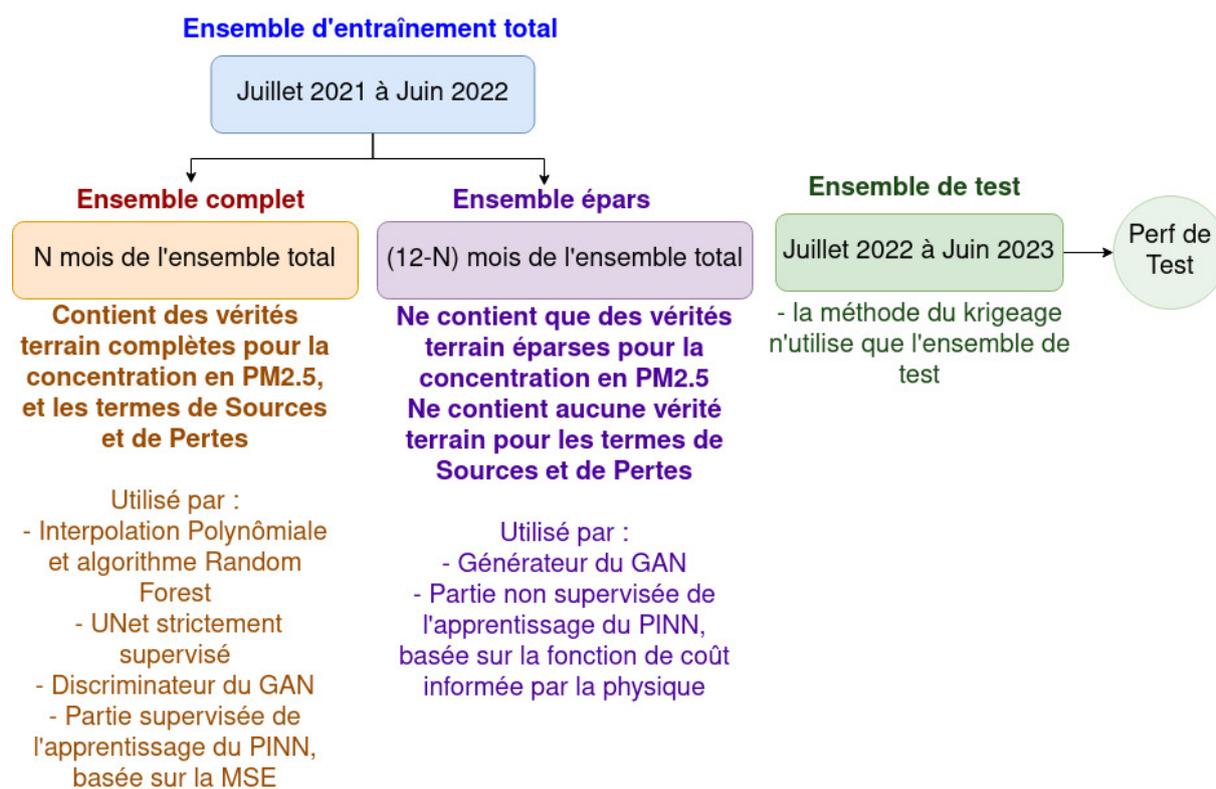
Pour ces deux prédictions, l’horizon de prédiction sera chaque fois de trois heures. Des expériences ont été réalisées en utilisant des horizons plus longs pour la deuxième

prédiction. Leurs résultats sont présentés en section A.9.

## 5.2 Protocole expérimental

### 5.2.1 Données

Les données utilisées dans ces travaux sont fournies par le modèle CAMS [6]. Comme dans les travaux précédents, ces données simulées sont utilisées pour répliquer une situation réelle, qui pourrait être obtenues avec des données réelles. Ainsi, l'ensemble d'entraînement total est divisé en deux sous-ensembles, plus petits : un ensemble d'entraînement dit complet (parce qu'il contient des cartes complètes de la concentration en aérosols) et un ensemble dit épars (qui ne contient que quelques pixels épars de ces mêmes cartes). La figure 5.3 illustre cette idée.



**Figure 5.3:** Illustration des caractéristiques du jeu de données utilisé durant toutes les expériences présentées dans ce chapitre. Le nombre  $N$  de mois couverts par l'ensemble complet est une variable dont la valeur peut être choisie arbitrairement.

L'ensemble d'entraînement total couvre une période temporelle de douze mois. Le nombre  $N$  de mois couverts par l'ensemble complet est une variable dont la valeur

peut être choisie arbitrairement. Cela permet la réalisation d'une courte étude sur l'impact de la proportion non supervisée de l'apprentissage sur les résultats.

L'ensemble de test couvre lui aussi une période temporelle d'un an. Il est important de noter qu'il n'y a aucun recouvrement entre les jeux de données (ni entre leurs couvertures temporelles respectives). Le modèle CAMS fournit une image toutes les trois heures pour chaque variable. Par conséquent ces jeux de données contiennent 2920 échantillons chacun. Chaque image est de taille 451x900. Les variables considérées ici sont : l'AOD à 550nm, le coefficient d'Ångström (calculé à partir de valeurs de l'AOD à 550 et 865nm), les pression, température et humidité de surface, les vitesse et direction du vent, les termes de sources et de pertes d'aérosols, ainsi que la concentration en PM<sub>2.5</sub> au niveau du sol.

Afin de raccourcir le temps d'inférence et d'accélérer l'entraînement, seule une petite portion de chaque image est utilisée à la place de la carte entière. Cette portion est la même pour toutes les images : elle est de taille 225x450, et ne couvre que l'Europe au lieu du monde entier. Nous émettons l'hypothèse que la plupart des conclusions dérivées des résultats obtenus avec ce jeu de données géographiquement restreint, pourraient également être dérivées d'expérimentations réalisées sur l'ensemble de données original, contenant des images couvrant la planète entière.

### 5.2.2 Pré-traitement des données

Pour ces travaux, le processus de pré-traitement appliqué aux données est identique à celui décrit dans la section 4.2.2. La seule exception est que des termes de source et de pertes d'aérosols sont cette fois utilisés, mais ils ne nécessitent aucun pré-traitement particulier.

Les valeurs de vitesse et direction du vent obtenues à l'issue de ce pré-traitement sont utilisées parmi les entrées du modèle. Cependant pour le calcul de la fonction de coût basée sur la physique, les valeurs originales de vitesse du vent dans les directions du Nord et de l'Est, sont utilisées.

### 5.2.3 Evaluation des modèles

Pour l'évaluation de nos modèles, nous utilisons les métriques décrites en section 3.3 ainsi que la métrique FSIM, introduite en section 4.2.4.

## 5.2.4 Conditions Aux Limites : configurations et méthodes

L'un des objectifs de nos expérimentations sur les PINNs est d'étudier l'impact que les CAL (Conditions Aux Limites), leur quantité et leur distribution, pourraient avoir sur les résultats. Ces différentes caractéristiques des CAL sont appelées leur configuration. Il s'agit d'un paramètres importants de ces expériences. Cinq configurations différentes de CAL ont été proposées en section 3.2.3.3. Elles sont réutilisées dans ce chapitre :

**Sans** : aucune CAL (cas habituel de l'entraînement semi-supervisé)

**Aléatoire 1%** : 1% des pixels de la vérité terrain sont utilisés en tant que CAL, et ils sont localisés à des positions aléatoires

**Aléatoire 5%** : même idée que pour **Aléatoire 1%** mais 5% de pixels sont utilisés au lieu de 1%

**Aléatoire Terre 1%** : même idée que pour **Aléatoire 1%**, mais les pixels choisis en tant que CAL sont toujours géographiquement positionnés sur terre (donc pas dans la mer ou l'océan). L'idée est de simuler les positions géographiques que pourraient avoir des capteurs de concentration en aérosols, or ceux-ci ne peuvent être situés que sur terre.

**Aléatoire Terre 5%** : même idée que pour **Aléatoire Terre 1%** mais avec 5% de pixels au lieu de 1%

Un autre objectif de ces expérimentations est de comparer deux méthodes pour la gestion des CAL, et de déterminer laquelle amène aux meilleures performances avec notre PINN. Sun et al. [49] appellent ces méthodes les contraintes (ou méthodes) dure et douce, respectivement.

La section 3.1 décrit ces deux méthodes en grand détail. En résumé, la contrainte dure consiste à appliquer une transformation directement sur les sorties du modèle, afin de le forcer à respecter les CAL. A l'inverse la contrainte douce consiste à concevoir une fonction de coût pour permettre au modèle d'apprendre explicitement à respecter les CAL. Cette fonction est essentiellement une MSE localisée, appliquée uniquement sur les localisations correspondant aux CAL.

Dans le chapitre 3, nous avons appliqué ces méthodes à un GAN en apprentissage semi-supervisé, et en avons déduit que la méthode douce amenait à de meilleures performances avec ce modèle. Nous utiliserons donc ce modèle comme base de

comparaison avec le PINN développé dans ce chapitre.

### 5.2.5 Disponibilité du code et des données

Le code utilisé pour les expériences présentées dans ce chapitre est disponible via une archive Zenodo.<sup>1</sup>

Les données utilisées pour ces expériences ne peuvent pas, pour des raisons pratiques, être partagées aisément. Cependant ces données peuvent être téléchargées depuis le système de stockage de données atmosphériques du programme Copernicus,<sup>2</sup> ou depuis le centre de données et de services ICARE.<sup>3</sup>

## 5.3 Résultats

### 5.3.1 Comparaison à l'état de l'art

La table 5.1 compare les résultats des méthodes proposées en section 2.2 avec les méthodes proposées dans ce chapitre.

Il est important de noter que l'interpolation polynômiale et l'algorithme Random Forest utilisent exclusivement l'AOD comme donnée d'entrée, tandis que la méthode du krigeage utilise uniquement les valeurs éparées de concentration en aérosols au niveau du sol. A l'inverse, nos modèles (y compris les UNets et GANs) utilisent toutes les variables d'entrée décrites en section 1.4, comme suggéré dans le chapitre 4. Pour tous les résultats présentés en table 5.1, les ensembles d'entraînement complets et éparés couvrent chacun une période de six mois (en d'autres termes,  $N = 6$ ).

La table 5.1 montre que notre PINN surclasse tous les autres modèles testés, sur toutes les métriques à l'exception du temps d'inférence et de la *MBE* (qui représente le biais). Cependant le temps d'inférence reste suffisamment bas pour que ce PINN soit une solution exploitable.

Il peut être remarqué que l'écart de performance entre le UNet et notre PINN est faible. La section 5.3.4 discute en plus grand détail des différences de performance des modèles UNet, GAN et PINN en fonction de la quantité de données annotées disponible. Elle propose également une réponse à cet argument.

---

<sup>1</sup><https://doi.org/10.5281/zenodo.14041853>

<sup>2</sup><https://ads.atmosphere.copernicus.eu/!/home>

<sup>3</sup><https://www.icare.univ-lille.fr/>

**Table 5.1:** Comparaison des méthodes présentées en section 2.2 avec les méthodes proposées dans ce chapitre. "IP D<sub>3</sub>" signifie Interpolation Polynômiale de Degré 3, "KO HE" signifie Krigeage Ordinaire avec variogramme "Hole-Effect". Pour tous les résultats présentés dans cette table,  $N = 6$ . "CAL-A" signifie "CAL Aléatoire", "CAL-AT" signifie "CAL Aléatoire sur Terre émergée", "CS" "Contrainte Soft (Douce)" et "CD" "Contrainte Dure".

Modèle	t	MAE	QE	MBE	FSIM
IP D <sub>3</sub>	<b>0.0006</b>	4.46	0.6	-1.44	6.49%
Random Forest	1.8525	4.54	0.59	-0.45	7.51%
KO HE	69.053	12.76	1.11	-4.58	11.29%
UNet	0.0062	3.74	0.49	-1.41	6.19%
GAN 5% CAL-A-CS	0.006	6.72	0.95	-6.68	6.63%
<b>PINN 5% CAL-A-CS</b>	0.0083	3.17	0.4	-2.13	<b>5.15%</b>
<b>PINN 5% CAL-AT-CD</b>	0.008	<b>2.71</b>	<b>0.33</b>	<b>0.11</b>	5.38%

Enfin la section 5.3.3 détaille l'impact des CAL sur la performance des PINNs, et propose une interprétation pour la différence de performance entre les deux dernières lignes de la table 5.1 (PINN 5% CAL-A-CS et PINN 5% CAL-AT-CD).

### 5.3.2 Impact de la fonction de coût informée par la physique

Des expériences ont été réalisées avec et sans l'utilisation de la fonction de coût informée par la physique, avec la même architecture de modèle. Le but est de comprendre si la performance du modèle est plutôt due à son architecture ou à l'utilisation de cette fonction de coût.

**Table 5.2:** Résultats de notre modèle avec et sans utilisation de la fonction de coût informée par la physique. Pour tous les résultats présentés dans cette table,  $N = 6$ , la configuration de CAL est Aléatoire 5%, et la contrainte douce est chaque fois utilisée.

Coût Physique	Temps d'Inférence	MAE	QE	MBE	FSIM
<b>Oui</b>	0.0083	<b>3.17</b>	<b>0.4</b>	<b>-2.13</b>	<b>5.15%</b>
Non	<b>0.008</b>	3.62	0.45	<b>-2.13</b>	5.53%

La table 5.2 montre que la fonction de coût basée sur la physique amène effectivement à une augmentation de la performance pour notre modèle. En d'autres termes, sa performance n'est pas seulement dûe à son architecture, mais également à l'utilisation de sa fonction de coût basée sur la physique.

En observant les tables 5.1 et 5.2, il apparaît clairement que même sans utilisation de la fonction de coût basée sur la physique, notre modèle surclasse toujours l'ensemble

des méthodes présentées dans la table 5.1 (à l'exception du PINN). Cela montre ainsi que bien que l'architecture de notre modèle a été conçue pour permettre l'utilisation de la fonction de coût basée sur la physique, elle amène tout de même par elle-même une augmentation de la performance.

La caractéristique principale de l'architecture de notre modèle est qu'elle permet l'inférence simultanée, à partir de la même entrée, de valeurs de concentration en aérosols et de termes de sources et de pertes d'aérosols. Ces résultats montrent donc qu'apprendre à inférer des termes de sources et de pertes d'aérosols aide le modèle à apprendre à inférer des valeurs de concentration en PM<sub>2.5</sub>, ce qui est intéressant.

Une interprétation possible à ce phénomène est que cette architecture particulière applique en réalité un type de contrainte physique sur le modèle, en le forçant à produire des représentations latentes adaptées à la fois pour l'inférence de valeurs de concentrations en aérosols et de termes de sources et de pertes. De plus, il existe une idée selon laquelle un modèle entraîné pour réaliser une tâche particulière peut voir ses performances augmenter s'il est entraîné sur une tâche supplémentaire en lien avec la première. Cette idée n'est pas nouvelle, et est à la base du domaine de l'Apprentissage Multi-Tâches [83]. On peut considérer qu'il existe un lien entre les tâches de prédiction des valeurs de concentrations en aérosols et de termes de sources et de pertes, car ces deux dernières variables ont un impact sur la première.

La section A.6 propose une analyse des résultats de notre modèle sur sa tâche secondaire de prédiction des termes de sources et de pertes.

### 5.3.3 Impact des CAL

Comme expliqué en section 5.2.4, notre PINN est capable d'exploiter les CAL en utilisant la contrainte dure ou douce. Des expériences sont réalisées avec ces deux méthodes et avec les cinq configurations de CAL présentées dans la même section. Le but est de mieux comprendre l'impact des CAL sur les résultats. Plusieurs observations intéressantes peuvent être faites sur ces résultats. Premièrement, pour la configuration "Aléatoire 5%", la contrainte douce surclasse la contrainte dure. De plus, la contrainte douce avec cette configuration de CAL surclasse les résultats obtenus avec la configuration sans BC, ce qui n'est pas le cas pour la contrainte dure.

Pour chaque autre configuration de CAL, l'observation opposée peut être faite : les résultats obtenus avec la contrainte dure surclassent ceux obtenus sans CAL, qui eux-

**Table 5.3:** Impact des configurations de CAL sur les résultats de nos modèles PINN. Pour tous les résultats de cette table,  $N = 6$ . "A" signifie Aléatoire et AT "Aléatoire Terre".

Méthode	Config de CAL	t	MAE	QE	MBE	FSIM
Aucune	Sans BC	0.0081	3.4	0.4	0.96	6.15%
Dure	A 1%	0.0081	2.8	0.34	-0.55	5.16%
Dure	A 5%	0.008	3.88	0.43	1.99	<b>5.11%</b>
Dure	AT 1%	0.0081	3.35	0.43	<b>-0.02</b>	6.18%
Dure	AT 5%	<b>0.008</b>	<b>2.71</b>	<b>0.33</b>	0.11	5.38%
Douce	A 1%	0.0081	4.56	0.6	-4.01	5.54%
Douce	A 5%	0.0083	3.17	0.4	-2.13	5.15%
Douce	AT 1%	0.0084	3.5	0.39	1	5.15%
Douce	AT 5%	0.0081	4.11	0.53	-3.62	5.83%

mêmes surclassent ceux obtenus avec la contrainte douce. Cela montre que, comme montré par Sun et al. [49], le PINN a tendance à obtenir de meilleures performances avec la contrainte dure qu'avec la contrainte douce. En l'occurrence, l'écart de performance est important. Une interprétation possible pour ceci est qu'avec la contrainte douce, une fonction de coût supplémentaire est utilisée, et des effets de compétition entre celle-ci et la fonction de coût basée sur la physique pourraient apparaître. Nous recommandons par conséquent l'utilisation de la contrainte dure avec le PINN.

Les meilleurs résultats obtenus avec la contrainte douce correspondent à la configuration de CAL "Aléatoire 5%". Cela est peu surprenant car cette configuration fournit à la fois le plus grande quantité de valeurs éparses de vérités terrain, et la répartition la plus uniforme. En effet la contrainte douce implique l'utilisation d'une fonction de coût additionnelle, et localisée (il s'agit d'ailleurs d'une MSE localisée). Cette fonction de coût se concentre donc sur les performances locales, plus spécifiquement aux performances à l'endroit où se situent les valeurs de CAL. Il semble donc logique qu'un nombre plus important de CAL, ainsi qu'une distribution plus uniforme, amène à de meilleurs résultats quand la contrainte douce est utilisée.

Cependant lorsque la contrainte dure est utilisée, les meilleurs résultats obtenus correspondent à la configuration "Aléatoire Terre 5%" et les deuxième meilleurs avec la configuration "Aléatoire 1%", ce qui semble donc plus surprenant. Cependant l'idée derrière la contrainte dure est simplement de remplacer quelques pixels par leurs valeurs connues (les CAL). Cette méthode se repose sur la ou les fonctions de coût existantes pour permettre au modèle de produire des sorties conformes à ces valeurs.

Dans notre cas, la fonction de coût existante est globale et non locale. Cela signifie que la contrainte dure ne se concentre pas sur la performance locale, mais plutôt sur l'augmentation de la performance générale du modèle via l'utilisation de valeurs locales. Il ne semble donc pas évident que l'utilisation de davantage de CAL amènerait à de meilleurs résultats avec la contrainte dure : cela peut dépendre des valeurs de CAL elles-mêmes, et de la pertinence de l'information qu'elles représentent.

En effet, il semblerait que certaines valeurs de vérités terrain sont plus utiles que d'autres pour l'entraînement du modèle, en fonction de leur localisation et du phénomène météorologique en jeu au moment considéré. Cela semble vrai en général et en particulier pour les PINNs, car leur fonction de coût contient des informations sur les relations locales entre les valeurs de concentration à prédire. Cela signifie qu'un changement local dans les valeurs éparses de vérités terrain pourrait avoir un impact important sur les sorties produites par le modèle.

Il est finalement important de noter que la majorité des sources d'aérosols sont situées sur des terres émergées. Cela signifie que les valeurs de concentration en aérosols situées sur ces terres pourraient en moyenne contenir des informations plus pertinentes, et donc être plus utiles au modèle, que des valeurs situées au-dessus de la mer ou de l'océan. Comme expliqué précédemment, cela est particulièrement vrai lorsque la contrainte dure est utilisée. Cela pourrait d'ailleurs expliquer pourquoi la configuration "Aléatoire Terre 5%" amène à de meilleurs résultats avec cette méthode que la configuration "Aléatoire 5%".

Quoiqu'il en soit, le fait que notre PINN atteigne de bonnes performances avec différentes configurations de CAL reste prometteur.

Enfin, même sans utiliser de CAL, notre PINN surclasse tout de même toutes les autres méthodes présentées dans la table 5.1.

#### **5.3.4 Impact de la proportion d'entraînement non supervisé**

Comme expliqué en section 5.2, nos jeux d'entraînement complets et sparse couvrent des périodes de respectivement  $N$  et  $12 - N$  mois,  $N$  étant un nombre arbitrairement choisi entre 0 et 12. Changer la valeur de  $N$  signifie changer la proportion d'entraînement non supervisé dans le processus d'entraînement semi-supervisé.

Des expériences sont réalisées avec les modèles UNet, GAN et PINN, avec différentes valeurs de  $N$ . En ce qui concerne l'entraînement du UNet, comme ce modèle est

entraîné de manière purement semi-supervisée, changer la valeur de  $N$  signifie simplement changer la quantité de données à laquelle ce modèle a accès. Il est donc attendu que les performances du UNet augmentent simplement avec les données auxquelles il a accès. La table 5.4 montre que plus la valeur de  $N$  est basse, plus le

**Table 5.4:** Impact de la taille du jeu de données épars sur les résultats de plusieurs modèles. L'ensemble complet couvre une période de  $N$  mois, et l'ensemble sparse  $12 - N$  mois. La configuration de CAL pour les GAN et PINN est "Aléatoire 5%" pour tous les résultats de cette table, et la contrainte douce est chaque fois utilisée.

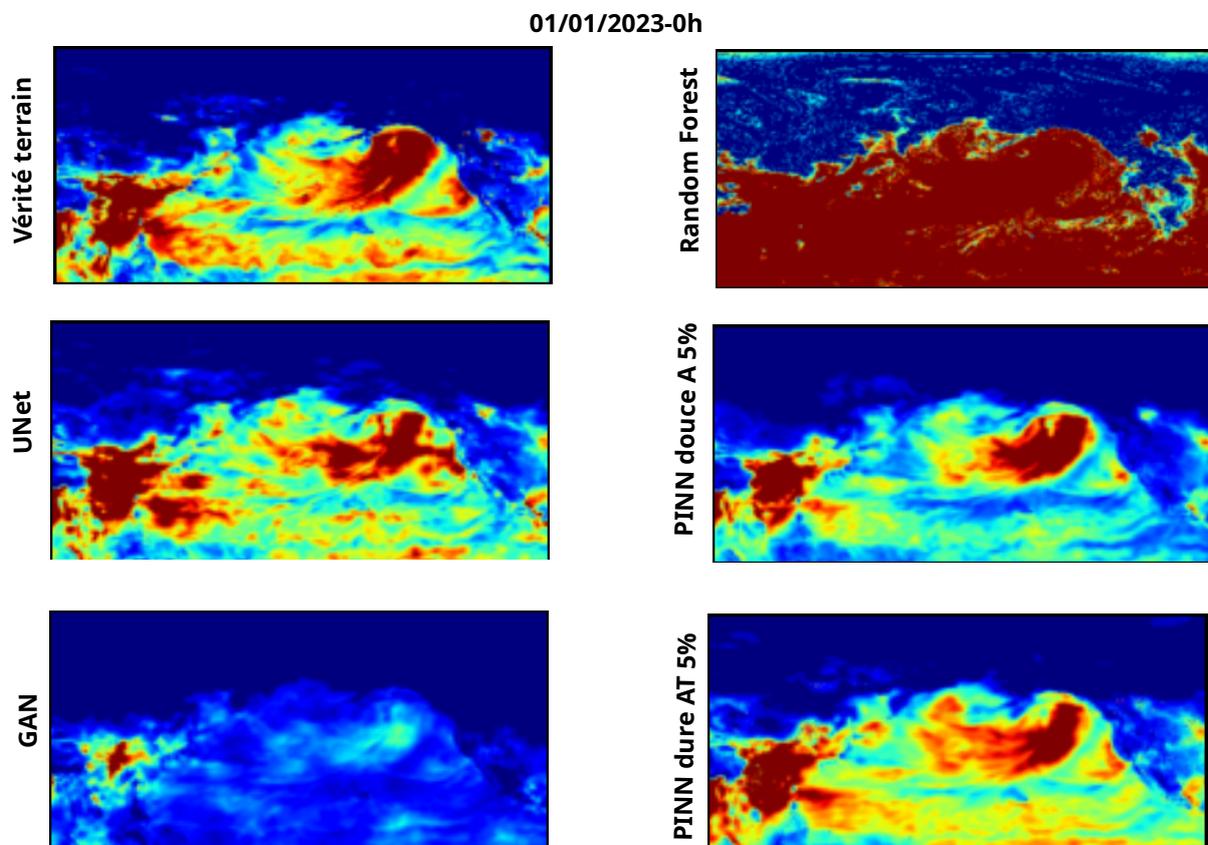
N	Modèle	t	MAE	QE	MBE	FSIM
9	UNet	<b>0.006</b>	4.29	0.51	2.39	<b>5.27%</b>
9	GAN	0.0063	9.44	0.75	6.15	8.88%
9	<b>PINN</b>	0.0081	<b>3.35</b>	<b>0.42</b>	<b>-0.61</b>	6.10%
6	UNet	0.0062	3.74	0.49	<b>-1.41</b>	6.19%
6	GAN	<b>0.006</b>	6.72	0.95	-6.68	6.63%
6	<b>PINN</b>	0.0083	<b>3.17</b>	<b>0.4</b>	-2.13	<b>5.15%</b>
3	UNet	0.006	4.6	0.61	-3.82	6.26%
3	GAN	<b>0.0059</b>	3.94	0.53	0.97	6.22%
3	<b>PINN</b>	0.008	<b>3.08</b>	<b>0.37</b>	<b>-0.37</b>	<b>5.94%</b>

GAN est performant. Cela est peu surprenant car ce modèle est conçu pour réaliser de l'apprentissage semi-supervisé. Il surclasse le UNet purement supervisé pour  $N = 3$ , en d'autres termes quand la proportion d'entraînement non supervisé dans le processus d'entraînement semi-supervisé est de 75%.

Elle montre également que, surprenamment, le UNet semble atteindre de meilleures performances quand  $N = 6$  que quand  $N = 9$ , bien que davantage de données soient accessibles dans le deuxième cas. Notre hypothèse est que ces données d'entraînement additionnelles sont moins représentatives du jeu de test, ce qui amène à une diminution de la performance sur l'ensemble de test choisi.

Enfin, cette table montre également que pour toutes les valeurs de  $N$  testées, notre PINN surclasse le UNet et le GAN, sur toutes les métriques à l'exception du temps d'inférence et de la *MBE*. L'écart de performance est particulièrement important pour  $N = 9$ . Elle montre aussi que plus la valeur de  $N$  est basse, plus les performances du PINN sont élevées. C'est aussi le cas du GAN, qui est lui aussi conçu pour réaliser de l'apprentissage semi-supervisé. Le fait que la meilleure performance du PINN dans cette table est obtenue avec la plus faible quantité de données annotées montre l'efficacité de la fonction de coût basée sur la physique.

### 5.3.5 Comparaison visuelle



**Figure 5.4:** Comparaison visuelle des sorties produites par le UNet, le GAN, le PINN et l'algorithme Random Forest. "douce A 5%" signifie que la contrainte dure est utilisée, et que la configuration de CAL utilisée est "Aléatoire 5%". "dure AT 5%" signifie que la contrainte dure est utilisée, et que la configuration de CAL utilisée est "Aléatoire Terre 5%".

La figure 5.4 nous permet de comparer les sorties produites par le GAN, le UNet et le PINN présentés dans ce chapitre. Elle montre clairement que le GAN est surclassé par les deux autres modèles. La sortie produite par le UNet est plus détaillée que celles produites par le PINN. Cependant il semble que le PINN (spécialement lorsque la contrainte dure est utilisée) respecte davantage certains motifs et structures observés dans les vérités terrain.

## 5.4 Conclusion

Dans ce chapitre, un nouveau modèle semi-supervisé informé par la physique pour l'estimation de la concentration en PM<sub>2.5</sub> au niveau du sol a été proposé. Ce modèle

utilise une équation physique comme base pour l'une de ses fonctions de coût, et est capable de produire lui-même certaines des valeurs nécessaires à l'utilisation de cette fonction de coût.

Nous avons montré que ce modèle surclasse toutes les méthodes d'état de l'art et de Deep Learning présentées, incluant les méthodes présentées en section 2.2, mais non limité à ces dernières. Les études d'ablation proposées ont montré que cela était dû à la fois à son architecture inhabituelle et à sa fonction de coût informée par la physique.

Il s'est aussi montré capable de bénéficier de l'utilisation de Conditions Aux Limites (sous la forme de matrices éparses de vérité terrain) en tant qu'information a priori. De plus, la contrainte dure a été montrée comme étant plus efficace que la douce pour la gestion des CAL avec ce PINN, ce qui est cohérent avec les conclusions d'autres travaux [49] ainsi que du chapitre 3.

Enfin, il a été montré que ce modèle continue de surclasser les méthodes de référence lors de l'utilisation de différentes quantités de vérités terrains complètes et éparses. Cela démontre sa robustesse et son efficacité dans le contexte de l'entraînement semi-supervisé.

Ce modèle ayant été entraîné et testé sur des données simulées, l'adapter et l'appliquer à des données réelles est une possibilité intéressante de travaux futurs. Son application à des images représentant un domaine géographique plus important est également considérée.

Ce réseau de neurones informé par la physique contient un nombre significatif de paramètres. Une autre possibilité de travaux futurs consiste à réduire ce nombre de paramètres, ce qui faciliterait un éventuel futur déploiement.

# 6

## Conclusion et Travaux Futurs

### Sommaire

---

6.1	Bilan des Contributions . . . . .	128
6.1.1	GAN avec utilisation de vérités terrain éparses sous forme de CAL . . . . .	128
6.1.2	Variables d'entrée et choix d'architecture . . . . .	129
6.1.3	PINN pour la prédiction des PM <sub>2.5</sub> au niveau du sol . . . . .	130
6.2	Travaux futurs . . . . .	131
6.2.1	Des métriques spécialisées pour la prédiction des termes de sources et de pertes . . . . .	131
6.2.2	La robustesse face aux changements de résolution et de couverture spatiales . . . . .	132
6.2.3	Une meilleure prise en compte de l'aspect temporel . . . . .	132
6.2.4	L'application aux données réelles . . . . .	133

---

## 6.1 Bilan des Contributions

La popularité des modèles d'Apprentissage Profond, et plus particulièrement de Vision Artificielle, s'est depuis un certain temps étendue aux domaines scientifiques et entre autres à celui de la physique atmosphérique. Ce domaine d'application comporte ses propres défis, qui concernent notamment la disponibilité des données et l'acceptabilité des modèles proposés.

Les travaux présentés dans ce manuscrit proposent de relever ces défis à travers plusieurs axes de recherche. Le premier axe s'attaque principalement à l'utilisation de vérités terrain éparses durant l'entraînement semi-supervisé d'un ANN. Le deuxième axe concerne l'exploration du lien entre les variables utilisées comme données d'entrée et l'architecture la plus efficace. Enfin le troisième et dernier axe s'attaque à l'augmentation de l'acceptabilité des modèles via l'exploitation d'informations a priori sur la physique du problème.

Ces pistes de recherche correspondent à nos contributions principales, décrites respectivement dans les chapitre 3, 4 et 5.

### 6.1.1 GAN avec utilisation de vérités terrain éparses sous forme de CAL

La première contribution que nous proposons consiste en la conception d'un *BC-GAN*, un GAN capable d'exploiter les vérités terrain éparses durant son entraînement semi-supervisé, en les considérant comme des Conditions Aux Limites (CAL). En effet les CAL représentent un type d'information souvent utile pour résoudre des problèmes physique. Les méthodes proposées pour les exploiter sont d'ailleurs inspirées de solutions utilisées en apprentissage informé par la physique.

Nos expérimentations démontrent l'intérêt de permettre l'exploitation par un GAN de ces CAL et de l'information qu'elles représentent. En effet, nos résultats montrent une supériorité de la solution proposée par rapport aux méthodes de référence de la section 2.2. Nous observons notamment une baisse de la MAE de  $0.18 \mu\text{g}/\text{m}^3$  lors de l'utilisation de notre modèle au lieu d'un UNet (la meilleure méthode de référence). Nous rappelons ici que la MAE est l'une de nos métriques principales, et qu'une diminution de sa valeur correspond à une augmentation de la performance du modèle évalué. La supériorité de notre modèle sur la méthode du krigeage est démontrée par une analyse visuelle des sorties des modèles illustrées par la figure 3.6.

Nos expériences ont aussi montré une claire supériorité d'une des deux méthodes proposées, la méthode douce, dans ce contexte. Pour la même configuration de CAL (Aléatoire 5%), nous observons une baisse de la MAE de  $6.33 \mu\text{g}/\text{m}^3$  lors de l'utilisation de la méthode douce plutôt que la méthode dure avec notre modèle. Nous suggérons cependant que la méthode à utiliser reste dépendante des spécificités du modèle concerné.

Enfin une comparaison de différentes configurations de CAL montre que la quantité de pixels contenue dans les matrices de vérité terrain éparées ainsi que l'uniformité de leur répartition ont une influence positive sur les résultats : la performance augmente généralement avec ces deux caractéristiques. Ainsi lorsque la contrainte douce est utilisée, nous remarquons une baisse de la MAE de  $5.22 \mu\text{g}/\text{m}^3$  lorsque l'on utilise la configuration Aléatoire 5% plutôt qu'Aléatoire Terre 1%, et la différence entre ces deux configurations réside justement dans la quantité de pixels contenue dans les CAL et l'uniformité de leur répartition.

### 6.1.2 Variables d'entrée et choix d'architecture

Notre deuxième contribution est une étude portant sur l'impact des variables d'entrée choisies pour leur pouvoir explicatif vis-à-vis de la concentration en aérosols au niveau du sol, qui est mis en avant par plusieurs études [23]. Elle montre que plus le nombre de variables d'entrées utilisé est important, plus la performance des modèles est importante. Si cette tendance est peu surprenante, ses quelques exceptions le sont davantage. La table 4.1 présente par exemple une situation dans laquelle l'utilisation du coefficient d'Ångström comme variable d'entrée supplémentaire amène à une diminution générale de la performance du modèle, et notamment à une augmentation de  $0.34 \mu\text{g}/\text{m}^3$  de la MAE.

Cette même étude porte également sur les stratégies de fusion à adopter en fonction des variables utilisées. Nos expérimentations désignent la stratégie de Concaténation de Canaux, la plus simple, comme la meilleure option. Elle amène en effet à de meilleurs résultats, à une plus grande stabilité dans la performance, et permet la conception de modèles plus légers. Par exemple, lorsque toutes les variables d'entrée sont utilisées, et avec les données issues du modèle CAMS, on observe une baisse de la MAE de  $0.05 \mu\text{g}/\text{m}^3$  lorsque la stratégie de Concaténation de Canaux est utilisée au lieu de celle de Fusion de Décisions. La différence est de  $0.08 \mu\text{g}/\text{m}^3$  avec les données issues du modèle ALADIN.

Des méthodes de fusion hybride, prenant en compte les caractéristiques des variables utilisées, sont également proposées. Ils font preuve de performances comparables aux autres modèles, sans toutefois montrer de supériorité ou d’avantage pratique évidents. La meilleure de ces méthodes présente une MAE supérieure de  $0.06 \mu\text{g}/\text{m}^3$  à la meilleure MAE obtenue avec un modèle utilisant la stratégie de la Concaténation de Canaux, avec les données issues du modèle CAMS.

Enfin des expérimentations sont réalisées sur différents ensembles de données. Celles-ci démontrent la forte influence de la résolution et de la couverture spatiales sur les résultats et les conclusions pouvant être tirés de ceux-ci. La variabilité spatiale des données, qui diffère selon les variables utilisées, est donc un élément à prendre en compte lors de l’élaboration d’une architecture d’ANN pour des problématiques atmosphériques. Ces expériences soulignent également la nécessité de réaliser davantage de travaux pour mieux comprendre le lien entre l’architecture des réseaux et leur habileté à prendre en compte l’ensemble des caractéristiques spatiales des différentes variables atmosphériques.

### 6.1.3 PINN pour la prédiction des PM<sub>2.5</sub> au niveau du sol

La troisième contribution proposée dans ce manuscrit est la conception d’un ANN informé par la physique pour la prédiction de la concentration en aérosols au niveau du sol. Une fonction de coût basée sur une équation physique permet au modèle de diminuer son besoin en données annotées, et donc de réaliser de l’entraînement semi-supervisé. Dans notre cas d’étude, il était de plus capable de générer des valeurs nécessaires à l’utilisation de cette fonction de coût.

Nos expérimentations montrent que ce modèle surclasse toutes nos méthodes de référence ainsi que le BC-GAN du chapitre 3. La MAE de notre meilleur modèle est en effet inférieure de  $1.03 \mu\text{g}/\text{m}^3$  à la MAE du UNet (le meilleur modèle de référence). Elles montrent aussi que ce modèle doit ses performances à son architecture ainsi qu’à sa fonction de coût informée par la physique, basée sur l’équation 6.1. En effet, l’utilisation de cette fonction de coût amène à une diminution de la MAE de  $0.45 \mu\text{g}/\text{m}^3$ .

$$\frac{\partial C}{\partial t} + \text{div}(C\vec{V}) - \text{div}(K_m \vec{\text{grad}}(C)) - (S - P) = 0 \quad (6.1)$$

Ce modèle démontre une capacité à bénéficier de l’utilisation de CAL, via l’utilisation de la méthode dure plutôt que la douce comme pour le BC-GAN. En effet pour la

même configuration de CAL (Aléatoire Terre 5%), l'utilisation de la méthode dure plutôt que la méthode douce amène à une diminution de la MAE de  $1.4 \mu\text{g}/\text{m}^3$ . Cela est cohérent avec les conclusions des auteurs de [49] et avec des suggestions faites dans le chapitre 3 selon lesquelles le choix de la méthode de prise en compte des CAL à utiliser dépend du type de modèle concerné.

De plus sa performance relative à l'état de l'art reste globalement identique même lors de l'augmentation et de la diminution de la quantité de données annotées, démontrant ainsi une certaine robustesse. Plus précisément, lorsque l'on augmente de 50% la quantité de données annotées, la MAE du modèle reste inférieure à celle du UNet de  $0.94 \mu\text{g}/\text{m}^3$ . Quand à l'inverse cette quantité est réduite de 50%, la MAE du modèle est inférieure à celle du BC-GAN (qui dans cette configuration est plus performant que le UNet) de  $0.86 \mu\text{g}/\text{m}^3$ .

## 6.2 Travaux futurs

Durant ces travaux, diverses possibilités d'amélioration de nos méthodes ont été identifiées. Certaines ont été appliquées, d'autres sont décrites ici afin de servir d'inspiration pour d'éventuelles futures contributions dans le domaine.

### 6.2.1 Des métriques spécialisées pour la prédiction des termes de sources et de pertes

Le modèle proposé dans le chapitre 5 a la particularité de proposer la prédiction à la fois de valeurs de concentration et de termes de sources et de pertes d'aérosols. Les résultats présentés dans ce chapitre démontrent que cette architecture particulière permet au modèle de surpasser l'état de l'art en matière de prédiction de la concentration en PM<sub>2.5</sub> au niveau du sol.

Ses capacités de prédiction des termes de sources et de pertes sont, elles, évaluées en section A.6. Cependant cette évaluation laisse à désirer. Deux éléments sont nécessaires pour réaliser cette évaluation de manière satisfaisante. Le premier est un ensemble de métriques adaptées à l'évaluation de la prédiction de ces termes. Cela n'est pas le cas de toutes les métriques proposées en section 3.3, comme exprimé en section A.6. En effet, lorsqu'elles sont appliquées aux termes de source et de perte, elles ne permettent pas l'obtention de valeurs interprétables, ce qui interdit donc la comparaison efficace entre différents modèles. Le deuxième élément est la constitution d'un ensemble de

méthodes de référence pour cette tâche spécifique, afin de positionner les méthodes proposées vis-à-vis de l'état de l'art.

### **6.2.2 La robustesse face aux changements de résolution et de couverture spatiales**

Tous les modèles présentés durant ces travaux ont un point commun : ils sont chaque fois entraînés sur un domaine spatial précis. Ce domaine est défini par une couverture et une résolution spatiales. Pour un modèle donné, les données utilisées peuvent correspondre à différentes variables et à différents pas de temps, mais le même domaine spatial est conservé tout au long de l'apprentissage comme lors de l'inférence. Tout changement de domaine spatial nécessite d'entraîner un nouveau modèle à part entière. L'une des raisons identifiées est que la déformation spatiale, due à la représentation sous forme d'une matrice en deux dimensions de variables représentées naturellement en coordonnées sphériques (puisqu'elles sont mesurées à la surface de la Terre), dépend de la résolution et de la couverture spatiales de cette matrice.

Il serait par conséquent intéressant de concevoir un modèle capable de fonctionner dans n'importe quel domaine spatial. Il est nécessaire pour cela de représenter les données sous un format capable de prendre en compte la position des valeurs dans le domaine réel, qui n'a en l'occurrence pas la forme d'une image en deux dimensions. Des modèles comme PhyGeoNet proposé par les auteurs de [67], ou encore des solutions basées sur le concept de CNNs sphériques [84, 85], sont des pistes intéressantes de solutions à ce défi. Une méthode exploitant à la fois ces principes et ceux développés dans les travaux de cette thèse est donc à envisager.

### **6.2.3 Une meilleure prise en compte de l'aspect temporel**

Durant ces travaux, des CNNs sont souvent utilisés pour leur capacité à exploiter les relations locales et spatiales entre les valeurs. C'est d'ailleurs avec ce type de modèle que nous obtenons nos meilleurs résultats. Cependant, les relations entre les valeurs météorologiques que nous utilisons ne sont pas exclusivement spatiales : elles sont aussi temporelles. Il serait donc intéressant de concevoir un modèle capable de prendre en compte ces relations temporelles, et donc de prendre des séries temporelles en entrée et de les exploiter efficacement.

Zhou et al. [27] montrent que les solutions basées sur des Réseaux de Neurones

Récurrent (Recurrent Neural Network) (RNN) et notamment sur des Longue Mémoire à Court Terme (Long Short-Term Memory) (LSTM), se révèlent utiles pour l'estimation de la concentration en PM<sub>2.5</sub> au niveau du sol. De plus les Transformers proposés par Vaswani et al. [29] et leur adaptation à la Vision Artificielle [30] sont eux aussi des solutions potentielles pour la prise en compte des relations temporelles entre les variables concernées. L'application de ces méthodes à notre concept, et conjointement aux solutions déjà proposées, est donc aussi à étudier.

#### **6.2.4 L'application aux données réelles**

L'ensemble des méthodes proposées dans ces travaux sont appliquées à des données issues de simulations. L'application de ces modèles à des données réelles contient son propre lot de défis supplémentaires.

Par exemple, dans le cas où les données nécessaires sont bien disponibles sous forme de matrices, il arrive que ces matrices ne représentent pas l'intégralité du domaine spatial disponible. Notamment, une même variable peut être produite par différents instruments situés sur différents satellites, couvrant chacun une portion du domaine spatial. Le recouvrement des matrices produites par ces satellites devrait permettre l'obtention d'une matrice couvrant la totalité du domaine spatial, cependant des écart de pas de temps entre les variables obtenues par les différents instruments sont à prévoir. De plus la présence de nuages peut dans certains cas gêner l'obtention de certaines valeurs (notamment l'AOD). Le dépassement de l'ensemble de ces obstacles est un sujet de recherche à part entière.

# A

## Annexes

### Sommaire

---

A.1	Procédure d'égalisation des valeurs de concentration . . . . .	135
A.2	Architecture des discriminateurs des GANs pour chaque stratégie de fusion . . . . .	136
A.2.1	Fusion de données / Concaténation de Canaux . . . . .	136
A.2.2	Feature Fusion . . . . .	136
A.2.3	Decision Fusion . . . . .	136
A.3	Utilisation de la FSIM en tant que fonction de coût additionnelle . . .	138
A.4	Complexité des modèles en fonction de la stratégie de fusion . . . . .	138
A.5	Performances d'encodeurs-décodeurs convolutifs et linéaires . . . . .	139
A.6	Performance de notre PINN pour la prédiction des termes de Sources et de Pertes . . . . .	140
A.7	Performance de notre PINN sur les valeurs extrêmes . . . . .	141
A.8	Etude d'ablation des différentes couches du UNet . . . . .	142
A.9	Performance de notre PINN avec d'autres horizons de prédiction . .	143

---

## A.1 Procédure d'égalisation des valeurs de concentration

La table A.1 présente les résultats de quelques modèles avec et sans l'application de la fonction  $\ln(1+x)$  aux vérités terrain avant l'apprentissage. Avant le calcul des métriques, si la fonction logarithme a été utilisée, la fonction  $\exp(x) - 1$  est appliquée aux vérités terrain et aux sorties du modèle comme suggéré en section 3.2.2.

En dehors de ces considérations, ces résultats sont obtenus avec le protocole expérimental décrit en section 3.2. On peut voir sur cette table que l'utilisation de la fonction

**Table A.1:** Comparaison des résultats de l'interpolation polynômiale, de l'algorithme Random Forest ainsi que du UNet avec et sans application de la fonction logarithme comme suggéré en section 3.2.2.

Log	Modèle	MAE ( $\mu\text{g}/\text{m}^3$ )	rMAE (%)	MBE ( $\mu\text{g}/\text{m}^3$ )	rMBE (%)	SE (no units)
Non	Interpolation Polynômiale de Degré 3	7.95 $\pm 1.67$	60.12% $\pm 7.47\%$	0.74 $\pm 2.27$	5.29% $\pm 16.41\%$	<b>0.44</b> $\pm 0.03$
Non	Random Forest	<b>7.89</b> $\pm 1.57$	<b>59.68%</b> $\pm 6.66\%$	<b>0.49</b> $\pm 2.07$	<b>3.45%</b> $\pm 15.02\%$	<b>0.44</b> $\pm 0.02$
Non	UNet + Couches Linéaires	11.72 $\pm 6.04$	87.95% $\pm 40.09\%$	6.41 $\pm 7.53$	47.85% $\pm 53.39\%$	0.85 $\pm 0.43$
Oui	Interpolation Polynômiale de Degré 3	7.14 $\pm 1.04$	54.12% $\pm 2.69\%$	-2.76 $\pm 1.51$	-21.19% $\pm 11.62\%$	0.41 $\pm 0.03$
Oui	Random Forest	7.17 $\pm 1.07$	54.34% $\pm 2.89\%$	-2.7 $\pm 1.43$	-20.85% $\pm 11.27\%$	0.41 $\pm 0.03$
Oui	UNet + Couches Linéaires	<b>6.69</b> $\pm 1.03$	<b>50.71%</b> $\pm 2.97\%$	<b>-1.36</b> $\pm 0.7$	<b>-10.21%</b> $\pm 4.79\%$	<b>0.39</b> $\pm 0.02$

logarithme amène à une augmentation de la performance de l'ensemble des modèles présentés, sur presque toutes ces métriques. La seule exception à cette augmentation générale de la performance est que l'Interpolation Polynômiale et l'algorithme Random Forest présentent un moins bon biais après application de cette fonction.

Le UNet semble particulièrement bénéficier de cette méthode, ses performances augmentant de manière significative. Ainsi lorsque cette fonction est appliquée, le UNet surpasse la méthode d'interpolation polynômiale et l'algorithme Random Forest, ce qui n'est pas le cas lorsqu'elle n'est pas utilisée.

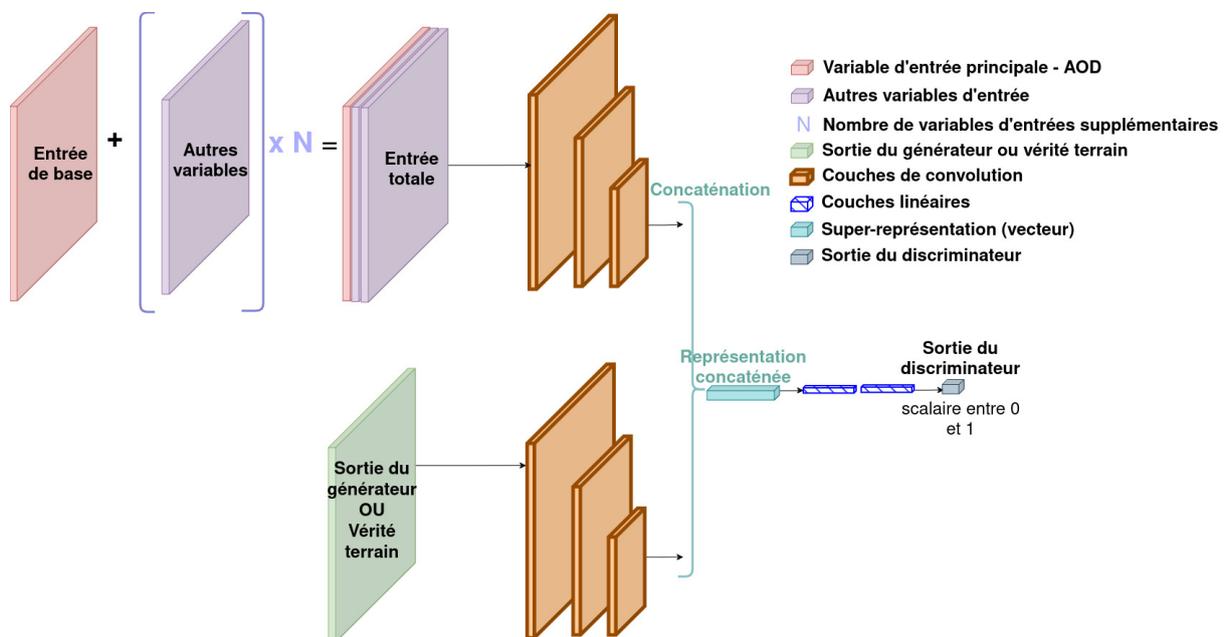
Pour cette raison, l'utilisation de la fonction  $\ln(1+x)$  aux vérités terrain avant l'apprentissage, puis de la fonction  $\exp(x) - 1$  aux vérités terrain et sorties du modèle avant calcul des métriques, est fortement conseillé.

## A.2 Architecture des discriminateurs des GANs pour chaque stratégie de fusion

Cette section présente l'architecture des discriminateurs des GANs en fonction de la stratégie de fusion utilisée.

### A.2.1 Fusion de données / Concaténation de Canaux

La figure A.1 montre l'architecture du discriminateur d'un GAN avec la stratégie de fusion de données.



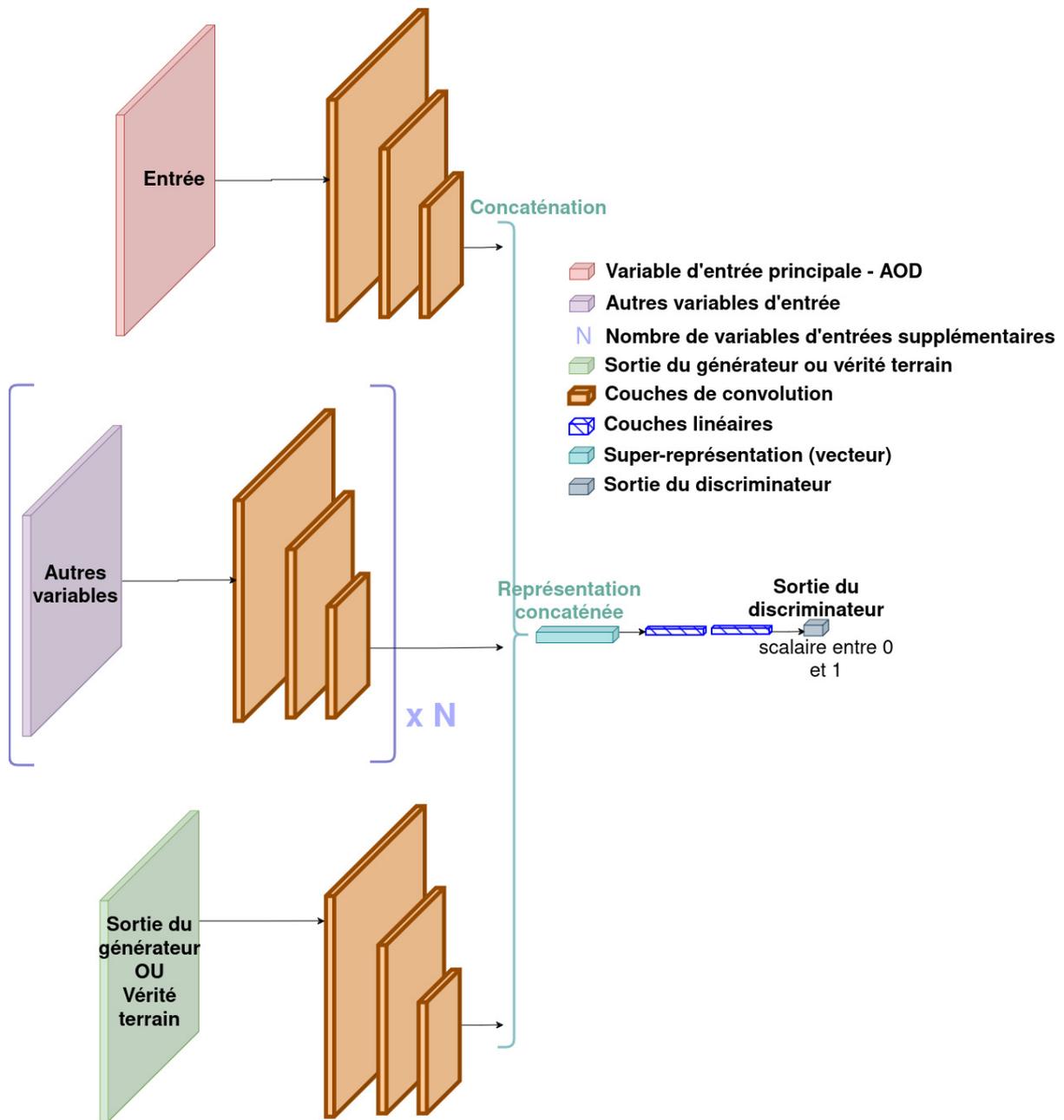
**Figure A.1:** Architecture du discriminateur du GAN avec la stratégie de fusion de données.

### A.2.2 Feature Fusion

La figure A.2 montre l'architecture du discriminateur d'un GAN avec la stratégie de fusion de représentations.

### A.2.3 Decision Fusion

La figure A.3 montre l'architecture du discriminateur d'un GAN avec la stratégie de fusion de décisions. Elle est très similaire à l'architecture d'un UNet avec fusion de



**Figure A.2:** Architecture du discriminateur du GAN avec la stratégie de fusion de représentations.

décisions, comme illustré par la figure 4.3 en section 4.1.4. La principale différence réside dans le format de la sortie du modèle. Il s'agit d'images pour le UNet, et de scalaires entre 0 et 1 pour le discriminateur.

### A.3 Utilisation de la FSIM en tant que fonction de coût additionnelle

La *FSIM*, présentée en section 4.2.4, peut être utilisée non seulement comme métrique, mais également comme fonction de coût additionnelle. Des expériences ont été réalisées pour vérifier l'efficacité de cette méthode. La table A.2 présente les résultats de ces expériences. Quand la *FSIM* est utilisée pour l'apprentissage en plus de la *MSE*, la fonction de coût totale est calculée selon l'équation A.1. Les données utilisées sont celles présentées en section 4.2.1. Les métriques utilisées sont présentées en section 3.3, et les scores en section 4.2.5.

$$CostTotal = \frac{3MSE + FSIM}{4} \quad (A.1)$$

**Table A.2:** Résultat avec et sans utilisation de la *FSIM* comme fonction de coût additionnelle. Le modèle est un UNet, la stratégie de fusion utilisée est la fusion de données ou Concaténation de Canaux. La seule variable d'entrée utilisée est l'AOD.

Avec FSIM	t	MAE	QE	MBE	FSIM	Scores (↗)		
						Total	Intemporel	Réduit
Non	0.0064	5.13	0.37	-2.08	4.66%	80.96%	79.4%	74.38%
Oui	0.0083	4.72	0.33	-2.09	3.92%	81.38%	80.87%	76.64%

La table A.2 montre bien que l'utilisation de la *FSIM* comme fonction de coût additionnelle amène à de meilleurs résultats, sur toutes les métriques exceptées le temps d'inférence et la *MBE*. Le temps d'inférence avec *FSIM* reste cependant suffisamment faible pour permettre l'obtention d'une solution exploitable. Quant à la *MBE*, l'écart de performance pour cette métrique entre les résultats avec et sans utilisation de la *FSIM* comme fonction de coût supplémentaire est exceptionnellement faible.

### A.4 Complexité des modèles en fonction de la stratégie de fusion

La figure A.4 montre le nombre de paramètres de nos modèles en fonction de la stratégie de fusion et du nombre de variables d'entrées utilisées. Elle présente le nombre de paramètres de nos UNets mais aussi de nos GANs, entraînés avec les données des modèles CAMS et ALADIN.

## A.5 Performances d’encodeurs-décodeurs convolutifs et linéaires

Li et al. [28] proposent et évaluent plusieurs modèles d’encodeur-décodeur linéaire avec connexions résiduelles, suivant l’idée illustrée par la figure 2.3. Nous nous intéressons à l’un de ces modèles en particulier. L’encodeur et le décodeur de ce modèle contiennent chacun six couches. Les couches de l’encodeur sont composées respectivement de 128, 96, 64, 32, 16 et 8 neurones. Le décodeur lui est exactement symétrique. Dans cette section, nous proposons de comparer les performances de cet encodeur-décodeur avec un UNet convolutif dont l’architecture est illustrée par la figure 3.5. Ce dernier modèle est inspiré par les travaux de Ronneberger, Fischer et Brox [43] et Li et al. [28].

Les données utilisées sont celles présentées en section 4.2.1. Seul l’AOD est utilisé comme donnée d’entrée pour ces expériences. Le pré-traitement appliqué aux données est celui décrit en section 3.2.2, et les métriques utilisées pour l’évaluation des performances sont celles décrites en sections 3.3 et 4.2.4. Seule la *MSE* est utilisée comme fonction de coût. La table A.3 présente les résultats de ces expériences.

**Table A.3:** Comparaison des performances d’un encodeur-décodeur proposé par les auteurs de [28] avec celles du UNet décrit par la figure 3.5.

Modèle	t	MAE	QE	MBE	FSIM
[28]	0.7108	7.59	0.65	-3.94	7.96%
Figure 3.5	<b>0.0064</b>	<b>5.13</b>	<b>0.37</b>	<b>-2.08</b>	<b>4.66%</b>

Ces résultats montrent sans appel la supériorité du UNet convolutif proposé sur le modèle d’ANN de Li et al. [28], dans ce contexte. En effet il est peu surprenant qu’un modèle contenant des couches de convolution présente de meilleures performances sur une tâche de translation image-vers-image, surtout au vu de la taille des images utilisées. En revanche Li et al. [28] montrent dans leurs travaux l’intérêt et l’efficacité des connexions résiduelles pour la prédiction des PM<sub>2.5</sub>. Cela est à nouveau démontré dans notre contexte dans le chapitre 3, et mis en œuvre à travers l’ensemble de notre manuscrit.

## A.6 Performance de notre PINN pour la prédiction des termes de Sources et de Pertes

Le modèle décrit dans le chapitre 5 permet non seulement la prédiction de la concentration en aérosols, mais également celles des termes de Sources et de Pertes. Nous rappelons ici que notre modèle est conçu avant tout pour la prédiction de valeurs de concentration en aérosols : la prédiction des termes de sources et de pertes est une tâche secondaire, ayant à l'origine pour seul but d'augmenter la performance du modèle vis-à-vis de sa tâche principale. Néanmoins, nous proposons dans cette section d'évaluer la performance de ce modèle sur cette tâche secondaire. La table A.4 présente les résultats du modèle, et la figure A.5 propose une analyse visuelle de ces sorties du modèle.

**Table A.4:** Résultats de notre modèle pour la prédiction des termes de Sources et de Pertes, avec utilisation de la fonction de coût informée par la physique durant l'apprentissage non supervisé. Pour tous les résultats présentés dans cette table, la configuration de CAL est Aléatoire 5%, et la contrainte douce est chaque fois utilisée.

N	Temps d'Inférence	MAE	QE	MBE	FSIM
9	0.0097	1.11	3	<b>0.48</b>	48.28%
6	0.0091	0.97	3.03	0.82	43.18%
3	<b>0.0083</b>	<b>0.96</b>	<b>2.97</b>	0.93	<b>40.43%</b>

Les termes de sources et de pertes d'aérosols étant situés sur leur propre échelle de valeurs, bien distincte de celle de la concentration en aérosols, il est difficile d'interpréter la MAE et la MBE obtenues lors de ces expérimentations. En revanche la QE et la FSIM sont indépendantes de cette échelle de valeurs. En ce qui concerne ces deux métriques, notre modèle montre de bien moins bonnes performances sur cette tâche que pour la prédiction de valeurs de concentration en PM<sub>2.5</sub>.

Ce résultat est également cohérent avec ce que nous observons sur la figure A.5. En effet, les sorties du modèle semblent relativement éloignées des vérités terrain, pour les termes de sources comme de pertes, aussi bien du point de vue des valeurs elles-mêmes que des formes et structures locales. En revanche, nous notons une ressemblance entre les sorties du modèle pour cette tâche secondaire, et la vérité terrain pour la concentration en aérosols. Il semblerait donc que l'entraînement du modèle sur la tâche principale influe la manière dont il réalise la tâche secondaire. En d'autres termes, le modèle tend à reproduire les structures locales de la concentration en PM<sub>2.5</sub> lors de la prédiction des termes de sources et de pertes. Nous émettons l'hypothèse que de

plus importants écarts de valeurs sont rencontrés dans les matrices de concentration en PM<sub>2.5</sub>, ce qui les amène à avoir davantage de poids durant l'apprentissage.

Nous avons quoiqu'il en soit démontré par l'expérience (notamment avec la table 5.2) que la réalisation de cette tâche secondaire augmentait la performance du modèle sur sa tâche principale de prédiction de la concentration en PM<sub>2.5</sub>.

## A.7 Performance de notre PINN sur les valeurs extrêmes

Nous proposons dans cette section d'évaluer la performance de notre modèle sur une tâche de prédiction des valeurs extrêmes de concentration en aérosols au niveau du sol. En particulier, nous nous intéressons aux valeurs de concentration supérieures à  $125\mu\text{g}/\text{m}^3$ , car ce seuil est associé à un niveau de qualité de l'air classifié comme très dangereux, selon l'Agence de Protection de l'Environnement des Etats-Unis <sup>1</sup>.

Pour cela, nous proposons l'utilisation d'une métrique dans ce but précis, que nous appelons MAE-X pour MAE sur les valeurs extrêmes. L'idée est de calculer la MAE uniquement sur les valeurs supérieures au seuil de  $125\mu\text{g}/\text{m}^3$ , toutes les autres valeurs étant considérées comme valant  $0\mu\text{g}/\text{m}^3$ .

La table A.5 présente les performances de trois modèles, à savoir un UNet, notre CAL-GAN et notre PINN. Ces performances sont évaluées via les métriques présentées en section 3.3 ainsi que via la MAE-X. Le protocole expérimental pour ces expériences correspond à celui décrit en section 5.2. L'ensemble d'entraînement complet et l'ensemble épars couvrent une période temporelle de six mois chacun. Le PINN et le GAN utilisent la configuration de CAL "Aléatoire sur Terre émergée, 5%", et la contrainte douce.

**Table A.5:** Comparaison des performances du UNet, de notre CAL-GAN et de notre PINN. On s'intéresse ici en particulier à la valeur de MAE-X.

Modèle	t	MAE	QE	MBE	FSIM	MAE-X
UNet	0.0062	3.74	0.49	-1.41	6.19%	240
CAL-GAN	0.006	6.72	0.95	-6.68	6.63%	<b>203</b>
<b>PINN</b>	0.0083	3.17	0.4	-2.13	<b>5.15%</b>	<b>203</b>

<sup>1</sup>Final Updates to the Air Quality Index (AQI) for Particulate Matter, Fact Sheet and Common Questions - U.S. Environmental Protection Agency (EPA) - Février 2024

<https://www.epa.gov/system/files/documents/2024-02/pm-naaqs-air-quality-index-fact-sheet.pdf>

On observe que le CAL-GAN et le PINN présentent une meilleure MAE-X que le UNet. Il semble raisonnable d’attribuer la meilleure performance de ces modèles sur cette métrique à leur capacité à réaliser de l’apprentissage semi-supervisé. C’est en effet là la principale différence entre le modèle UNet et les deux autres modèles. Il est intéressant de noter que le CAL-GAN présente une bonne performance, équivalente à celle du PINN, en ce qui concerne la prédiction de valeurs extrêmes.

## A.8 Etude d’ablation des différentes couches du UNet

Dans cette section nous proposons une étude d’ablation des couches du modèle UNet présenté et utilisé comme base de comparaison tout au long de ces travaux. Chaque couche produisant une représentation de la donnée à une échelle différente, cette étude d’ablation a également pour objectif d’identifier l’importance relative de ces représentations pour la prédiction de la concentration en aérosols au niveau du sol.

L’encodeur et le décodeur du UNet étant symétriques, lorsque pour une expérience donnée l’une des couches de l’encodeur est absente, la couche correspondante du décodeur est également absente.

La figure A.6 permet notamment d’identifier les noms des différentes couches du modèle. Ces noms sont réutilisés dans la table A.6, qui présente les résultats de cette étude d’ablation. Les autres caractéristiques de ce UNet sont illustrées par la figure 3.5.

Ces expériences sont réalisées avec le protocole expérimental décrit en section 5.2. L’ensemble d’entraînement complet et l’ensemble épars couvrent une période temporelle de six mois chacun.

**Table A.6:** Résultats de l’étude d’ablation sur les couches du UNet. Les noms des couches sont illustrés par la figure A.6.

Couche manquante	t	MAE	QE	MBE	FSIM
Aucune	0.0062	<b>3.74</b>	<b>0.49</b>	-1.41	<b>6.19%</b>
C1	0.0054	7.17	1.03	-7.07	6.92%
C2	<b>0.0051</b>	6.16	0.86	-5.86	6.76%
C3	0.0058	5.88	0.84	<b>-0.42</b>	8.75%
C4	0.0052	7.41	1.06	-7.22	6.88%
L5	0.006	6.01	0.84	-5.66	7.36%
L6	0.0063	7.42	1.08	-7.37	7.29%

Les couches *C2* et *C3* semblent avoir un effet moins important sur la performance du UNet que les autres couches de convolution. Cela peut s’expliquer par le fait que ces deux couches contiennent des noyaux de même taille (7). On peut raisonnablement penser que l’absence de ces deux couches à la fois aurait, en revanche, entraîné une diminution plus importante de la performance du modèle.

La couche *L6* semble quant à elle avoir un effet plus important sur la performance du modèle que la couche *L5*. La différence principale entre ces deux couches réside notamment dans la dimension du vecteur qu’elles produisent. *L5* produit un vecteur de taille 256 et *L6* un vecteur de taille 128. Il semblerait ainsi que réduire la dimension de la représentation latente jusqu’à 128 plutôt que 256 amène à de meilleures performances.

Il est intéressant de rappeler ici que l’analyse des résultats présentés en table 3.1, section 3.4 a montré que l’utilisation de couches complètement connectées amenait à une amélioration des performances dans le cas d’un UNet, mais pas pour un Encodeur-Décodeur classique.

Enfin la table A.6 nous permet d’observer une différence significative systématique de performance entre le UNet complet et le UNet privé d’une de ses couches, et ce peu importe la couche manquante. Cela démontre l’impact non négligeable sur la performance de chacune des couches du UNet illustré par la figure A.6.

## **A.9 Performance de notre PINN avec d’autres horizons de prédiction**

Les résultats présentés dans la section 5.3 ont été obtenus avec un horizon de prédiction de trois heures. Dans cette section, nous présentons les résultats de deux modèles, à savoir le UNet utilisé comme base de comparaison tout au long de nos travaux et notre PINN, avec des horizons de prédictions plus importants. Le UNet est choisi pour être comparé au PINN sur cette expérience car, comme le montre la table 5.1, il s’agit de la méthode la plus performante au sein de notre base de comparaison.

La table A.8 présente les résultats de ces modèles avec un horizon de prédiction de 24h, et la table A.9 avec un horizon de prédiction de 168h (ce qui correspond à une semaine). La table A.7 rappelle la performance de ces deux modèles pour un horizon de prédiction de 3h.

Le protocole expérimental pour ces expériences correspond à celui décrit en section 5.2. L'ensemble d'entraînement complet et l'ensemble épars couvrent une période temporelle de six mois chacun. Le PINN utilise la configuration de CAL "Aléatoire sur Terre émergée, 5%", et la contrainte douce.

**Table A.7:** Performance des modèles UNet et PINN avec un horizon de prédiction de 3h.

Modèle	t	MAE	QE	MBE	FSIM
UNet	<b>0.0062</b>	3.74	0.49	<b>-1.41</b>	6.19%
PINN	0.0083	<b>3.17</b>	<b>0.4</b>	-2.13	<b>5.15%</b>

**Table A.8:** Performance des modèles UNet et PINN avec un horizon de prédiction de 24h.

Modèle	t	MAE	QE	MBE	FSIM
UNet	<b>0.0068</b>	4.4	0.57	<b>-1.62</b>	9.72%
PINN	0.0084	<b>3.81</b>	<b>0.48</b>	-2.21	<b>8.64%</b>

**Table A.9:** Performance des modèles UNet et PINN avec un horizon de prédiction de 168h.

Modèle	t	MAE	QE	MBE	FSIM
UNet	<b>0.0068</b>	6.46	0.57	2.02	9.58%
PINN	0.019	<b>4.49</b>	<b>0.56</b>	<b>0.7</b>	<b>9.34%</b>

Dans ces trois cas, on peut observer que le PINN présente une meilleure performance que le UNet, sur toutes les métriques à l'exception de la MBE. Avec un horizon de prédiction de 168h, on remarque une différence plus importante entre les MAE des deux modèles. En revanche l'écart entre leurs valeurs de QE et de FSIM est plus faible dans ce cas.

Ces résultats démontrent bien la supériorité de notre PINN sur le UNet pour la prédiction de valeurs de concentration en aérosols au niveau du sol, et ce pour différents horizons de prédiction.

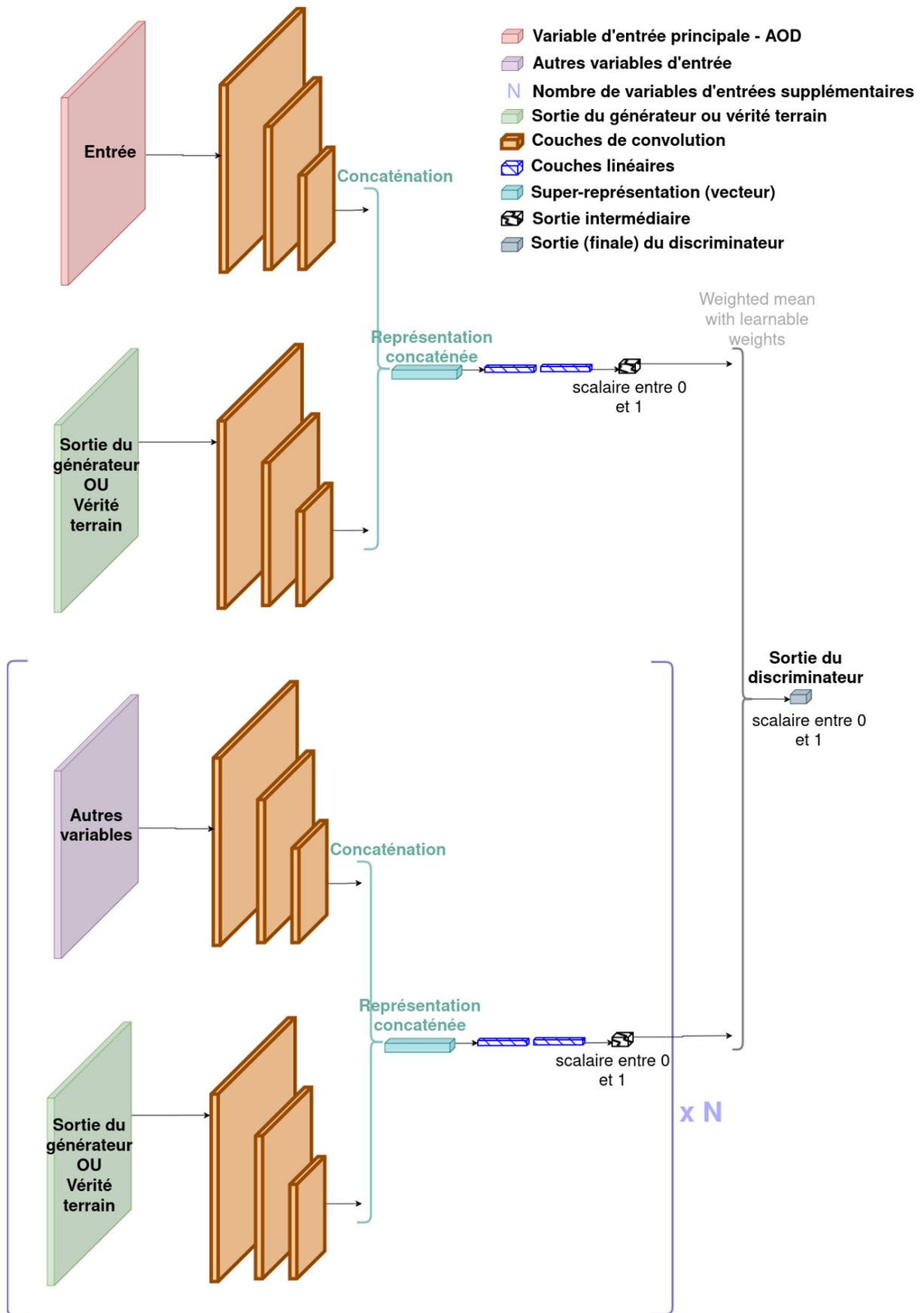
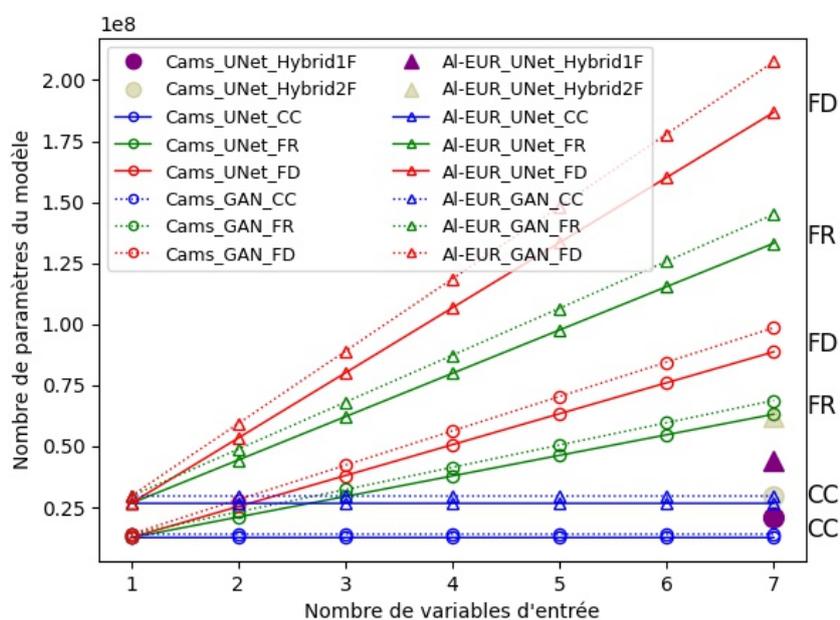
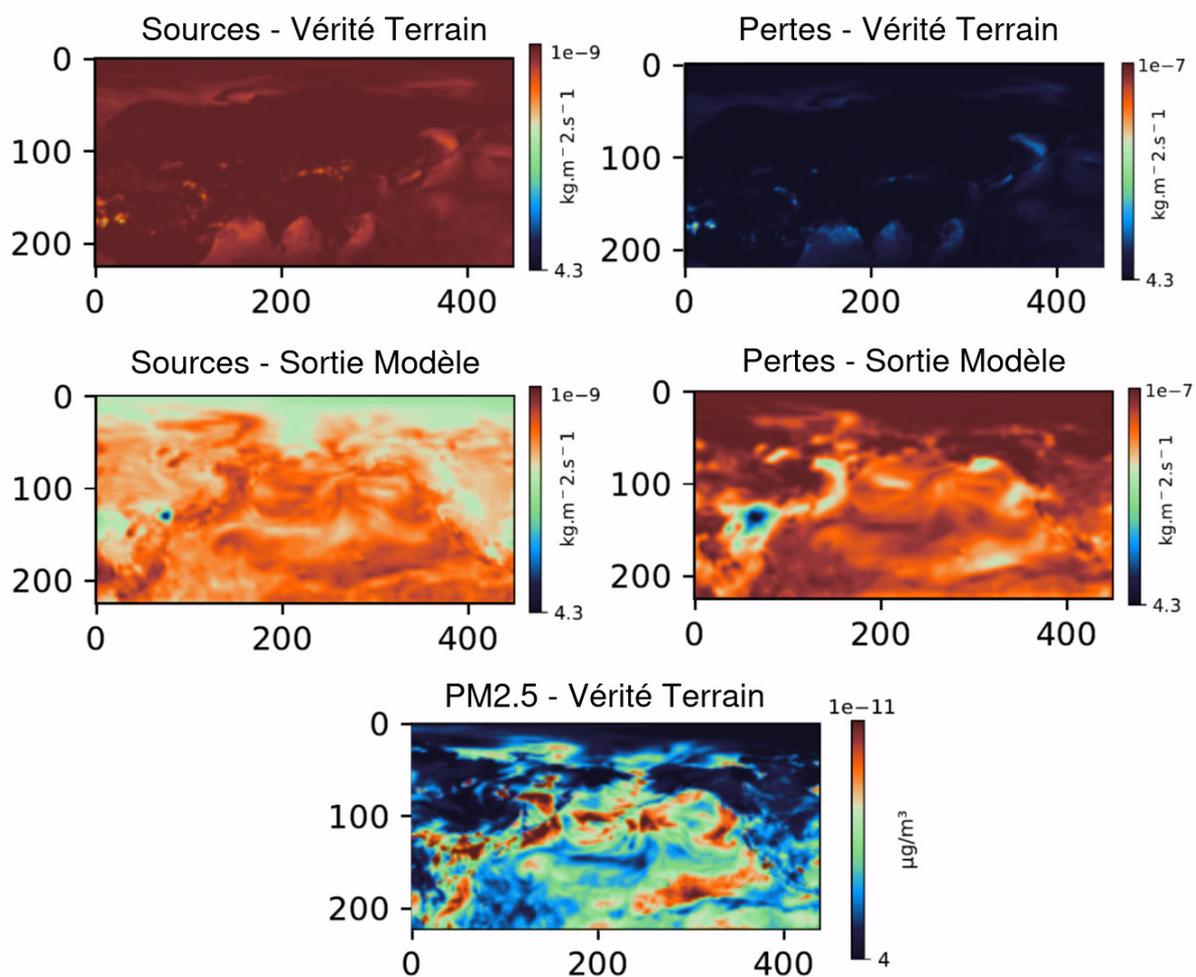


Figure A.3: Architecture du discriminateur du GAN avec la stratégie de fusion de décisions.



**Figure A.4:** Nombre de paramètres de chaque modèle en fonction du nombre de variables d'entrées. Pour rappel les types de fusion sont Concaténation de Canaux (CC), Fusion de Représentations (FR), et Fusion de Décisions (FD).



**Figure A.5:** Vérités terrain et sorties du modèle pour la prédiction de termes de Sources et de Pertes d'Aérosols

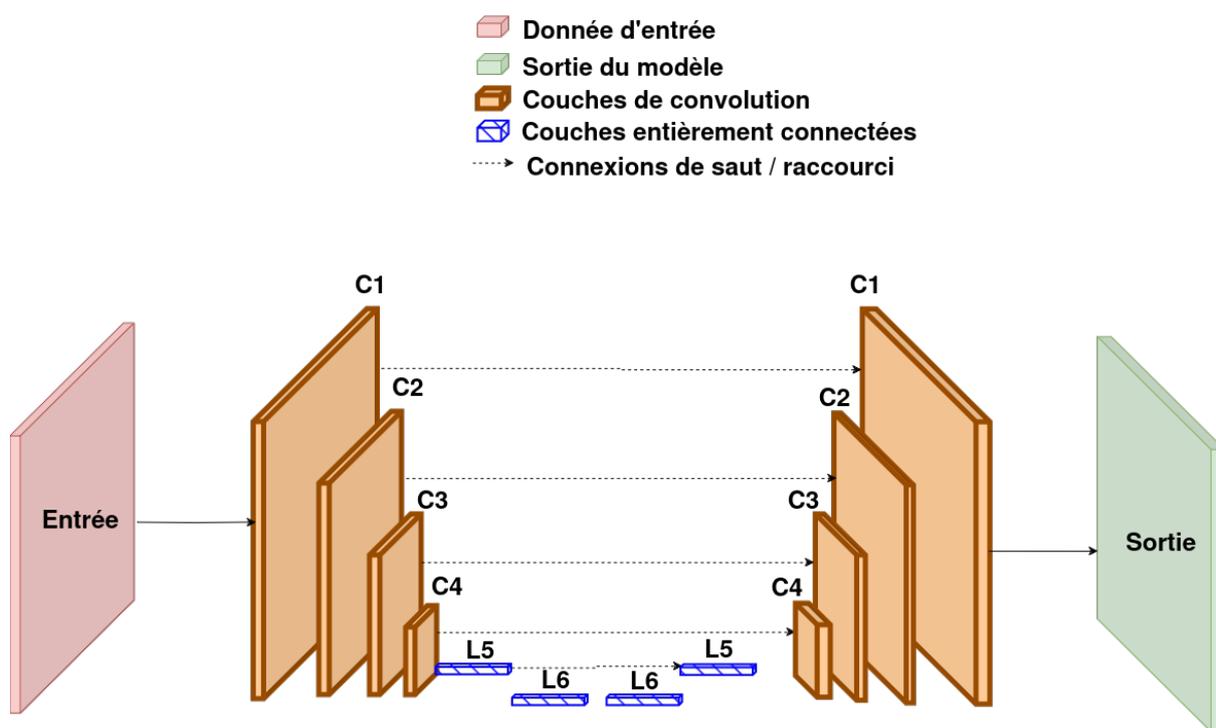


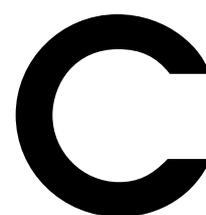
Figure A.6: Architecture du UNet avec noms des couches.

**B**

Financement

Ces travaux ont été financés par l'Agence Nationale de la Recherche, la région Hauts-de-France ainsi que le programme AI\_PhD@Lille de l'ISite.





# Bibliographie

## Sommaire

---

Références . . . . .	152
Publications . . . . .	161

---

## Références

- [1] S. Bose et al. “Indoor Particulate Matter Associated with Systemic Inflammation in COPD”. In: *Journal of Environmental Protection* 6.5 (2015), pp. 566–572. URL: <https://www.scirp.org/journal/paperinformation?paperid=56751>.
- [2] Madrigano Jaime et al. “Long-term Exposure to PM<sub>2.5</sub> and Incidence of Acute Myocardial Infarction”. In: *Environmental Health Perspectives* 121.2 (Feb. 2013), pp. 192–196. URL: <https://doi.org/10.1289/ehp.1205284>.
- [3] Andreas M. Neophytou et al. “Marginal Structural Models in Occupational Epidemiology: Application in a Study of Ischemic Heart Disease Incidence and PM<sub>2.5</sub> in the US Aluminum Industry”. In: *American Journal of Epidemiology* 180.6 (Sept. 2014), pp. 608–615. URL: <https://doi.org/10.1093/aje/kwu175>.
- [4] Aaron J Cohen et al. “Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015”. In: *The Lancet* 389.10082 (May 2017), pp. 1907–1918. URL: <https://www.sciencedirect.com/science/article/pii/S0140673617305056>.
- [5] Antje Inness et al. “The CAMS reanalysis of atmospheric composition”. In: *Atmos. Chem. Phys.* 19.6 (Mar. 20, 2019), pp. 3515–3556. URL: <https://acp.copernicus.org/articles/19/3515/2019/>.
- [6] Xavier Ceamanos, Bruno Six et Jérôme Riedi. “Quasi-Global Maps of Daily Aerosol Optical Depth From a Ring of Five Geostationary Meteorological Satellites Using AERUS-GEO”. In: *Journal of Geophysical Research: Atmospheres* 126.20 (2021), e2021JD034906. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1029/2021JD034906>.
- [7] B. Fougnie et al. “The multi-viewing multi-channel multi-polarisation imager—Overview of the 3MI polarimetric mission for aerosol and cloud characterization”. In: *Journal of Quantitative Spectroscopy and Radiative Transfer (JQRST)* 219 (2018), pp. 23–32.

- [8] Olivier Boucher. *Atmospheric Aerosols*. Dordrecht: Springer Netherlands, 2015. URL: <http://link.springer.com/10.1007/978-94-017-9649-1>.
- [9] Pawan Gupta et Sundar A. Christopher. "Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach". In: *Journal of Geophysical Research: Atmospheres* 114 (D14 2009). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1029/2008JD011496>.
- [10] G. Snider et al. "SPARTAN: a global network to evaluate and enhance satellite-based estimates of ground-level particulate matter for global health applications". In: *Atmospheric Measurement Techniques (AMT)* 8 (2015), pp. 505–521. URL: <https://doi.org/10.5194/amt-8-505-2015>.
- [11] Hanyu Xiang et al. "Deep learning for image inpainting: A survey". In: *Pattern Recognition* 134 (Feb. 1, 2023), p. 109046. URL: <https://www.sciencedirect.com/science/article/pii/S003132032200526X>.
- [12] L. Menut et al. "CHIMERE 2013: a model for regional atmospheric composition modelling". In: *Geoscientific Model Development (GMD)* 6.4 (July 22, 2013), pp. 981–1028. URL: <https://gmd.copernicus.org/articles/6/981/2013/>.
- [13] B. Lyu et al. "Deep-learning spatial principles from deterministic chemical transport models for chemical reanalysis: an application in China for PM<sub>2.5</sub>". In: *Geoscientific Model Development (GMD)* 15.4 (2022), pp. 1583–1594. URL: <https://gmd.copernicus.org/articles/15/1583/2022/>.
- [14] Yi Xiao et al. "Generating a long-term (2003–2020) hourly 0.25° global PM<sub>2.5</sub> dataset via spatiotemporal downscaling of CAMS with deep learning (Deep-CAMS)". In: *Science of The Total Environment* 848 (2022), p. 157747. URL: <https://www.sciencedirect.com/science/article/pii/S004896972204846X>.
- [15] Aaron van Donkelaar, Randall V. Martin et Rokjin J. Park. "Estimating ground-level PM<sub>2.5</sub> using aerosol optical depth determined from satellite remote sensing". In: *Journal of Geophysical Research: Atmospheres* 111 (D21 2006). URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006996>.
- [16] Guannan Geng et al. "Estimating long-term PM<sub>2.5</sub> concentrations in China using satellite-based aerosol optical depth and a chemical transport model". In: *Remote Sensing of Environment* 166 (2015), pp. 262–270. URL: <https://www.sciencedirect.com/science/article/pii/S0034425715300171>.
- [17] Zongwei Ma et al. "A review of statistical methods used for developing large-scale and long-term PM<sub>2.5</sub> models from satellite data". In: *Remote Sensing of Environment* 269 (Feb. 2022), p. 112827. URL: <https://www.sciencedirect.com/science/article/pii/S0034425721005472>.

- [18] Xinghan Xu, Chengkun Zhang et Yi Liang. "Review of satellite-driven statistical models PM<sub>2.5</sub> concentration estimation with comprehensive information". In: *Atmospheric Environment* 256 (July 1, 2021), p. 118302. URL: <https://www.sciencedirect.com/science/article/pii/S1352231021001205>.
- [19] Xuefei Hu et al. "Estimating ground-level PM<sub>2.5</sub> concentrations in the Southeastern United States using MAIAC AOD retrievals and a two-stage model". In: *Remote Sensing of Environment* 140 (Jan. 1, 2014), pp. 220–232. URL: <https://www.sciencedirect.com/science/article/pii/S0034425713002903>.
- [20] Trevor Hastie et Robert Tibshirani. "Generalized Additive Models: Some Applications". In: *Journal of the American Statistical Association* 82 (398 June 1, 1987), pp. 371–86. URL: <https://doi.org/10.1080/01621459.1987.10478440>.
- [21] Bin Zou et al. "Satellite Based Mapping of Ground PM<sub>2.5</sub> Concentration Using Generalized Additive Modeling". In: *Remote Sensing* 9 (1 Jan. 2017). URL: <https://doi.org/10.3390/rs9010001>.
- [22] Christopher Irrgang et al. "Towards neural Earth system modelling by integrating artificial intelligence in Earth system science". In: *Nature Machine Intelligence* 3.8 (Aug. 2021), pp. 667–674. URL: <https://doi.org/10.1038/s42256-021-00374-3>.
- [23] Mitra Unik et al. "PM<sub>2.5</sub> Estimation using Machine Learning Models and Satellite Data: A Literature Review". In: *International Journal of Advanced Computer Science and Applications* 14.5 (May 2023).
- [24] Seunghee Lee et al. "Air Quality Forecasts Improved by Combining Data Assimilation and Machine Learning With Satellite AOD". In: *Geophysical Research Letters* 49.1 (Jan. 2022).
- [25] Caiyi Jin et al. "An optimized semi-empirical physical approach for satellite-based PM<sub>2.5</sub> retrieval: embedding machine learning to simulate complex physical parameters". In: *Geoscientific Model Development (GMD)* 16.14 (July 2023), pp. 4137–4154.
- [26] Tongwen Li et al. "Point-surface fusion of station measurements and satellite observations for mapping PM<sub>2.5</sub> distribution in China: Methods and assessment". In: *Atmospheric Environment* 152 (2017), pp. 477–489. URL: <https://www.sciencedirect.com/science/article/pii/S1352231017300043>.
- [27] Shiyun Zhou et al. "Deep-learning architecture for PM<sub>2.5</sub> concentration prediction: A review". In: *Environmental Science and Ecotechnology* 21 (Sept. 2024), p. 100400. URL: <https://www.sciencedirect.com/science/article/pii/S2666498424000140>.

- [28] Lianfa Li et al. "Encoder–Decoder Full Residual Deep Networks for Robust Regression and Spatiotemporal Estimation". In: *IEEE Transactions on Neural Networks and Learning Systems* 32.9 (Sept. 2021), pp. 4217–4230. URL: <https://ieeexplore.ieee.org/document/9186306/?arnumber=9186306>.
- [29] Ashish Vaswani et al. "Attention is All you Need". In: *Advances in Neural Information Processing Systems (NeurIPS)* 30 (2017). Ed. by Inc. Curran Associates. URL: [https://proceedings.neurips.cc/paper\\_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html](https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html).
- [30] Li Yuan et al. "Tokens-to-Token ViT: Training Vision Transformers from Scratch on ImageNet". In: *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Oct. 2021, pp. 538–547. URL: <https://ieeexplore.ieee.org/document/9710747/?arnumber=9710747>.
- [31] Mohammed A. A. Al-qaness et al. "ResInformer: Residual Transformer-Based Artificial Time-Series Forecasting Model for PM<sub>2.5</sub> Concentration in Three Major Chinese Cities". In: *Mathematics* 11.2 (Jan. 2023), p. 476. URL: <https://www.mdpi.com/2227-7390/11/2/476>.
- [32] Debao Chen et al. "Spatiotemporally continuous PM<sub>2.5</sub> dataset in the Mekong River Basin from 2015 to 2022 using a stacking model". In: *Science of The Total Environment* 914 (Mar. 1, 2024), p. 169801. URL: <https://www.sciencedirect.com/science/article/pii/S0048969723084334>.
- [33] Katharina Beckh et al. *Explainable Machine Learning with Prior Knowledge: An Overview*. May 21, 2021. arXiv: 2105.10172[cs]. URL: <http://arxiv.org/abs/2105.10172>.
- [34] Yongbee Park et al. "Estimating PM<sub>2.5</sub> concentration of the conterminous United States via interpretable convolutional neural networks". In: *Environmental Pollution* 256 (Jan. 2020).
- [35] Sebastian Bach et al. "On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation". In: *PLOS ONE* 10.7 (July 2015), e0130140. URL: <https://doi.org/10.1371/journal.pone.0130140>.
- [36] Li Deng. "The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web]". In: *IEEE Signal Processing Magazine* 29.6 (Nov. 2012), pp. 141–142. URL: <https://ieeexplore.ieee.org/document/6296535>.
- [37] Jia Deng et al. "ImageNet: A large-scale hierarchical image database". In: *2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2009, pp. 248–255. URL: <https://ieeexplore.ieee.org/document/5206848>.

- [38] Tsung-Yi Lin et al. *Microsoft COCO: Common Objects in Context*. Feb. 20, 2015.
- [39] Tin Kam Ho. "Random decision forests". In: *Proceedings of 3rd International Conference on Document Analysis and Recognition 1* (Aug. 1, 1995), pp. 278–82. URL: <https://doi.org/10.1109/ICDAR.1995.598994>.
- [40] Georges Matheron. "Principles of geostatistics". In: *Economic Geology* 58.8 (Dec. 1, 1963), pp. 1246–1266. URL: <https://doi.org/10.2113/gsecongeo.58.8.1246>.
- [41] Yves Gratton. *Le Krigeage : la méthode optimale d'interpolation spatiale*. 2002.
- [42] Sophie Baillargeon. *Le krigeage : revue de la théorie et application à l'interpolation spatiale de données de précipitations*. 2005. URL: <https://corpus.ulaval.ca/jspui/handle/20.500.11794/18036>.
- [43] Olaf Ronneberger, Philipp Fischer et Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015, pp. 234–241.
- [44] Jesper E. van Engelen et Holger H. Hoos. "A survey on semi-supervised learning". In: *Mach Learn* 109.2 (Feb. 2020), pp. 373–440. URL: <http://link.springer.com/10.1007/s10994-019-05855-6>.
- [45] Ali Reza Sajun et Imran Zuolkernan. "Survey on Implementations of Generative Adversarial Networks for Semi-Supervised Learning". In: *Applied Sciences* 12.3 (Feb. 7, 2022), p. 1718. URL: <https://www.mdpi.com/2076-3417/12/3/1718>.
- [46] Manan Oza, Himanshu Vaghela et Sudhir Bagul. "Semi-Supervised Image-to-Image Translation". In: *2019 International Conference of Artificial Intelligence and Information Technology (ICAIIIT)*. Mar. 2019, pp. 16–20.
- [47] Jose Eusebio, Hemanth Venkateswara et Sethuraman Panchanathan. "Semi-supervised Adversarial Image-to-Image Translation". In: *Smart Multimedia*. Ed. by Anup Basu et Stefano Berretti. Vol. 11010. Cham: Springer International Publishing, 2018, pp. 334–344. URL: [http://link.springer.com/10.1007/978-3-030-04375-9\\_28](http://link.springer.com/10.1007/978-3-030-04375-9_28).
- [48] M. Raissi, P. Perdikaris et G.E. Karniadakis. "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations". In: *Journal of Computational Physics* 378 (Feb. 2019), pp. 686–707. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0021999118307125>.
- [49] Luning Sun et al. "Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data". In: *Computer Methods in Applied Me-*

- chanics and Engineering* 361 (Apr. 2020), p. 112732. URL: <https://linkinghub.elsevier.com/retrieve/pii/S004578251930622X>.
- [50] Phillip Isola et al. “Image-to-Image Translation with Conditional Adversarial Networks”. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI: IEEE, July 2017, pp. 5967–5976. URL: <http://ieeexplore.ieee.org/document/8100115/>.
- [51] Jun-Yan Zhu et al. “Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks”. In: *2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE, Oct. 2017, pp. 2242–2251. URL: <http://ieeexplore.ieee.org/document/8237506/>.
- [52] Ian Goodfellow et al. “Generative adversarial networks”. In: *Commun. ACM* 63.11 (Oct. 22, 2020), pp. 139–144. URL: <https://dl.acm.org/doi/10.1145/3422622>.
- [53] Martin Arjovsky, Soumith Chintala et Léon Bottou. “Wasserstein Generative Adversarial Networks”. In: *Proceedings of the 34th International Conference on Machine Learning*. PMLR, July 17, 2017, pp. 214–223. URL: <https://proceedings.mlr.press/v70/arjovsky17a.html>.
- [54] Ishaan Gulrajani et al. “Improved Training of Wasserstein GANs”. In: *Advances in Neural Information Processing Systems (NeurIPS)*. Vol. 30. Curran Associates, Inc., 2017. URL: <https://proceedings.neurips.cc/paper/2017/hash/892c3b1c6dccc52936e27cbd0ff683d6-Abstract.html>.
- [55] Yaxing Wang et al. “SDIT: Scalable and Diverse Cross-domain Image Translation”. In: *Proceedings of the 27th ACM International Conference on Multimedia*. MM ’19. New York, NY, USA: Association for Computing Machinery, Oct. 15, 2019, pp. 1267–1276. URL: <https://doi.org/10.1145/3343031.3351004>.
- [56] Yaxing Wang et al. “Semi-Supervised Learning for Few-Shot Image-to-Image Translation”. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA: IEEE, June 2020, pp. 4452–4461. URL: <https://ieeexplore.ieee.org/document/9156625/>.
- [57] Jun-Yan Zhu et al. “Toward Multimodal Image-to-Image Translation”. In: *Advances in Neural Information Processing Systems (NeurIPS)*. Vol. 30. Curran Associates, Inc., 2017. URL: <https://proceedings.neurips.cc/paper/2017/hash/819f46e52c25763a55cc642422644317-Abstract.html>.
- [58] Diederik P. Kingma et Max Welling. “An Introduction to Variational Autoencoders”. In: *FNT in Machine Learning* 12.4 (2019), pp. 307–392. arXiv: 1906.02691. URL: <http://arxiv.org/abs/1906.02691>.

- [59] Pan Zhang et al. *Semi-Supervised Image-to-Image Translation using Latent Space Mapping*. Mar. 29, 2022. arXiv: 2203.15241 [cs]. URL: <http://arxiv.org/abs/2203.15241>.
- [60] Roberto Confalonieri et al. "A historical perspective of explainable Artificial Intelligence". In: *WIREs Data Mining and Knowledge Discovery* 11.1 (2021), e1391. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/widm.1391>.
- [61] W. James Murdoch et al. "Definitions, methods, and applications in interpretable machine learning". In: *Proc. Natl. Acad. Sci. U.S.A.* 116.44 (Oct. 29, 2019), pp. 22071–22080. URL: <https://pnas.org/doi/full/10.1073/pnas.1900654116>.
- [62] Vaishak Belle et Ioannis Papantonis. "Principles and Practice of Explainable Machine Learning". In: *Front Big Data* 4 (July 1, 2021), p. 688969. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8281957/>.
- [63] Ribana Roscher et al. "Explainable Machine Learning for Scientific Insights and Discoveries". In: *IEEE Access* 8 (2020), pp. 42200–42216.
- [64] Qianli Liao et Tomaso Poggio. "Object-Oriented Deep Learning". In: (2017).
- [65] Hamidreza Eivazi et al. "Physics-informed neural networks for solving Reynolds-averaged Navier–Stokes equations". In: *Physics of Fluids* 34.7 (July 2022), p. 075117. URL: <https://aip.scitation.org/doi/full/10.1063/5.0095270>.
- [66] Pu Ren et al. "PhyCRNet: Physics-informed convolutional-recurrent network for solving spatiotemporal PDEs". In: *Computer Methods in Applied Mechanics and Engineering* 389 (Feb. 1, 2022), p. 114399. URL: <https://www.sciencedirect.com/science/article/pii/S0045782521006514>.
- [67] Han Gao, Luning Sun et Jian-Xun Wang. "PhyGeoNet: Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state PDEs on irregular domain". In: *Journal of Computational Physics* 428 (Mar. 2021), p. 110079. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0021999120308536>.
- [68] Xiaoyu Zhao et al. "Physics-informed convolutional neural networks for temperature field prediction of heat source layout without labeled data". In: *Engineering Applications of Artificial Intelligence* 117 (Jan. 2023), p. 105516. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0952197622005061>.
- [69] Lu Lu et al. "DeepXDE: A Deep Learning Library for Solving Differential Equations". In: *SIAM Rev.* 63.1 (Jan. 2021), pp. 208–228. URL: <https://epubs.siam.org/doi/10.1137/19M1274067>.
- [70] Pongpisit Thanasutives, Masayuki Numao et Ken-ichi Fukui. "Adversarial Multi-task Learning Enhanced Physics-informed Neural Networks for Solving Partial

- Differential Equations". In: *2021 International Joint Conference on Neural Networks (IJCNN)*. July 2021, pp. 1–9.
- [71] Zhenguo Nie et al. "TopologyGAN: Topology Optimization Using Generative Adversarial Networks Based on Physical Fields Over the Initial Domain". In: *Journal of Mechanical Design* 143.3 (Mar. 1, 2021), p. 031715. URL: <https://asmedigitalcollection.asme.org/mechanicaldesign/article/doi/10.1115/1.4049533/1094063/TopologyGAN-Topology-Optimization-Using-Generative>.
- [72] Liu Yang et al. "Highly-scalable, Physics-Informed GANs for Learning Solutions of Stochastic PDEs". In: *2019 IEEE/ACM Third Workshop on Deep Learning on Supercomputers (DLS)*. Nov. 2019, pp. 1–11.
- [73] Liu Yang, Dongkun Zhang et George Em Karniadakis. "Physics-Informed Generative Adversarial Networks for Stochastic Differential Equations". In: *SIAM J. Sci. Comput.* 42.1 (Jan. 2020), A292–A317. URL: <https://epubs.siam.org/doi/abs/10.1137/18M1225409>.
- [74] Mohammed Mallik et al. "Towards Outdoor Electromagnetic Field Exposure Mapping Generation Using Conditional GANs". In: *Sensors* 22.24 (Jan. 2022), p. 9643. URL: <https://www.mdpi.com/1424-8220/22/24/9643>.
- [75] Weijie Wang et Yanmin Lu. "Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model". In: *IOP Conf. Ser.: Mater. Sci. Eng.* 324 (Mar. 2018), p. 012049. URL: <https://iopscience.iop.org/article/10.1088/1757-899X/324/1/012049>.
- [76] Suchuan Dong et Naxian Ni. "A method for representing periodic functions and enforcing exactly periodic boundary conditions with deep neural networks". In: *Journal of Computational Physics* 435 (June 2021), p. 110242. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0021999121001376>.
- [77] Utthara Gosa Mangai et al. "A Survey of Decision Fusion and Feature Fusion Strategies for Pattern Classification". In: *IETE Technical Review* 27.4 (July 2010), pp. 293–307. URL: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,url,uid&db=a9h&AN=51893554&lang=fr&site=ehost-live>.
- [78] Weiwei Song et al. "Hyperspectral Image Classification With Deep Feature Fusion Network". In: *IEEE Transactions on Geoscience and Remote Sensing* 56.6 (June 2018), pp. 3173–3184.
- [79] Quan-Sen Sun et al. "A new method of feature fusion and its application in image recognition". In: *Pattern Recognition* 38.12 (Dec. 2005), pp. 2437–2448. URL: <https://linkinghub.elsevier.com/retrieve/pii/S003132030500083X>.

- [80] Abhijit Sinha et al. "Estimation and decision fusion: A survey". In: *Neurocomputing* 71.13 (Aug. 2008), pp. 2650–2656. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0925231208002282>.
- [81] Jingling Li et al. "How does a Neural Network's Architecture Impact its Robustness to Noisy Labels?" In: *Advances in Neural Information Processing Systems (NeurIPS)*. Ed. by M. Ranzato et al. Vol. 34. Curran Associates, Inc., 2021, pp. 9788–9803. URL: [https://proceedings.neurips.cc/paper\\_files/paper/2021/file/51311013e51adebc3c34d2cc591fefee-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2021/file/51311013e51adebc3c34d2cc591fefee-Paper.pdf).
- [82] Lin Zhang et al. "FSIM: A Feature Similarity Index for Image Quality Assessment". In: *IEEE Transactions on Image Processing* 20.8 (Aug. 2011), pp. 2378–2386.
- [83] Rich Caruana. "Multitask Learning". In: *Machine Learning* 28.1 (July 1, 1997), pp. 41–75. URL: <https://doi.org/10.1023/A:1007379606734>.
- [84] Taco S. Cohen et al. "Spherical CNNs". In: Mar. 12, 2022. URL: <https://openreview.net/forum?id=Hkbd5xZRb>.
- [85] N. Perraudin et al. "DeepSphere: Efficient spherical convolutional neural network with HEALPix sampling for cosmological applications". In: *Astronomy and Computing* 27 (Apr. 1, 2019), pp. 130–146. URL: <https://www.sciencedirect.com/science/article/pii/S2213133718301392>.

# Publications

- Matthieu Dabrowski et al. "Semi-supervised GAN with sparse ground truth as Boundary Conditions". In: *2023 International Joint Conference on Neural Networks (IJCNN)*. June 2023, pp. 1–9. URL: <https://ieeexplore.ieee.org/document/10191911>.
- Matthieu Dabrowski et al. "Knowledge-inspired fusion strategies for the inference of PM<sub>2.5</sub> values with a Neural Network". In: *Geoscientific Model Development (GMD) (submitted)* (2024-08). This article was submitted and is awaiting review/publication. Preprint is available. URL: <https://egusphere.copernicus.org/preprints/2024/egusphere-2024-2676/>.
- Matthieu Dabrowski et al. "Physics-Informed Model for the prediction of PM<sub>2.5</sub> concentration". In: *Neurcomputing (submitted)* (2024-09). This article was submitted and is awaiting review/publication. Preprint is available. URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4963604](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4963604).