UNIVERSITÉ DE LORRAINE | BIBLIOTHÈQUES UNIVERSITAIRES

AVERTISSEMENT
DROIT D'AUTEUR

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

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document. Toute contrefaçon, plagiat ou reproduction illicite encourt une poursuite pénale.

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

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

UNIVERSITÉ DE LORRAINE
34 COURS LÉOPOLD — CS 25233
54052 NANCY CEDEX
TÉL : 03 54 50 54 00
CONTACT@UNIV-LORRAINE.FR
WWW.UNIV-LORRAINE.FR

UNIVERSITÉ DE LORRAINE

École doctorale IAEM Lorraine

Apprentissage automatique et fusion de données pour la surveillance environnementale. Application à la prévention des risques d'exposition aux polluants dans les environnements de travail

THÈSE

présentée et soutenue publiquement le 17 décembre 2025

pour l'obtention du

Doctorat de l'Université de Lorraine

(Automatique, Traitement du signal et des images, Génie informatique)

par

Narech Hougbénou HOUESSOU

Composition du jury

Rapporteurs :	Valeriu VRABIE	Professeur des universités, Université de Reims Champagne-Ardenne
	Daniel DAUGERON	Maître de conférences, HDR, Université Clermont Auvergne
Examinateurs :	David BRIE	Professeur des universités, Université de Lorraine
	Evelyne GEHIN	Professeur des universités, Université Paris-Est Créteil
Directeur de thèse :	Sébastian MIRON	Maître de conférences, HDR, Université de Lorraine
Co-directeur de thèse :	Philippe DUQUENNE	Responsable du laboratoire ASTEC, INRS France
Président du jury:	David BRIE	

Centre de Recherche en Automatique de Nancy
CRAN UMR 7039, Université de Lorraine, CNRS
Faculté des Sciences et Technologies, BP 70239, 54506 Vandoeuvre Cedex

Mis en page avec la classe thesul.

Remerciements

Au terme de ce travail de thèse, je souhaite exprimer ma gratitude à toutes celles et ceux qui ont, de près ou de loin, accompagné ce parcours exigeant et fondateur.

Je tiens tout d'abord à adresser mes remerciements les plus sincères à mon directeur et co-directeur de thèse, Sébastian MIRON et Philippe DUQUENNE, pour leur confiance, leur bienveillance ainsi que leur exigence scientifique ; Votre accompagnement et conseils avisés ont guidé mes réflexions et mes analyses tout au long de ce travail. Je suis reconnaissante d'avoir pu faire ce chemin à vos côtés, Il n'a pas été facile mais sous votre direction, J'ai pu explorer la richesse d'un sujet à la croisée de plusieurs disciplines. Vos expertises ont été déterminantes dans la construction et l'aboutissement de ce projet de recherche. J'emporte avec moi le meilleur de chacune de vos approches.

Je souhaite également remercier Valeriu VRABIE, Daniel DAUGERON, David BRIE et Evelyne GEHIN d'avoir accepté de faire partie de mon jury de thèse et d'évaluer ce travail. Un remerciement particulier à Radu RANTA et David BRIE membres de mon comité de thèse pour leur suivi, leurs conseils tout au long de mon parcours doctoral.

Je remercie également Jean-Pierre LECLERC, responsable du département IP en début de thèse, pour son accueil au sein du département ainsi que le suivi du projet.

Au cours de ce projet, j'ai eu l'occasion et le privilège de collaborer avec de nombreuses personnes. Un grand merci aux membres de l'équipe du laboratoire ASTEC à l'INRS qui ont conduit l'étude à l'origine de ce projet : Karine GERARDIN, Bruno GALLAND, Esther MONNOYER et Nathalie MONTA, votre accompagnement sur les aspects méthodologiques en santé au travail ont grandement contribué à la mise en place du dispositif expérimental. Merci également aux collègues techniciens (Esther, Nathalie, Thomas) pour leur aide précieuse lors des campagnes de mesures. Leur disponibilité a largement facilité la réalisation de ce travail.

Dans le même sens, je souhaite remercier les collègues d'autres laboratoires à l'INRS, Denis BEMER pour les essais sur les particules, Sullivan LECHENE pour les simulations CFD, Michael SARREY et Johan HAYARD pour leur accompagnement à l'utilisation du Robot Sherpa.

Je remercie également Fabien GERARDIN, responsable du département IP, voisin de bureau d'alors. Ces conseils et diverses discussions enrichissantes m'ont encouragé tout au long de ce travail.

Aux membres de l'equipe Simul du département Biosis du CRAN, merci à chacun d'entre vous pour les échanges scientifiques, la convivialité quotidienne qui ont nourri mon engagement sur ce projet de thèse. Les moments passées au CRAN restent pour moi un temps d'apprentissage que je n'oublie pas. À Philippe, Saulo, Faustine, Laura, merci pour les moments partagés entre doctorants.

Je suis aussi reconnaissante pour toutes les discussions et échanges que cette thèse m'a permis d'avoir : conférences, séminaires.

Je n'oublie pas Dr Ratheil HOUNDJI et Dr Koubouratou IDJATON, Merci pour votre soutien des lors que j'ai émis l'idée de faire une thèse, vos conseils et encouragements ont accompagné mes premiers pas dans cette aventure scientifique.

A ma famille, pilier indéfectible de ce parcours, Mes parents Jean HOUESSOU et Elise GNACADJA, Sandrine HOUESSOU, Probus KIKI et Béryl HOUESSOU, merci pour votre soutien et vos encouragements.

A mes amis, Taloro, Likita, Luc, Urielle, Ricardo, Karen, Merci pour les moments de rires, de partages qui ont su agrémentés ces années de thèse.

Enfin, je souhaite remercier toutes celles et ceux qui, par un message inopiné, un mot ou un geste m'ont apporté un brin de réconfort. Merci à chacun d'entre vous d'avoir fait de cette aventure scientifique une expérience humaine.

Table des matières

Table des figures vii

Liste des tableaux xiii

Liste des sigles et abréviations xv

Notations xvii

Introduction générale 1

Chapitre 1 État de l'art de la cartographie des concentrations de polluants en milieu de travail 7

1.1 Introduction . . . 7
1.2 Utilisation de la cartographie spatio-temporelle en santé au travail . . . 9
1.2.1 Cartographie spatiale avec une acquisition itinérante des données . . . 9
1.2.2 Vers l'acquisition spatiale multi-points simultanée . . . 12
1.2.3 Cartographie spatio-temporelle avec une acquisition simultanée des données 12
1.3 Méthodes de cartographie spatio-temporelle des expositions aux polluants en santé au travail . . . 15
1.3.1 Historique des méthodes de cartographie spatio-temporelle . . . 15
1.4 Analyse de l'état de l'art sur les méthodes de cartographie spatio-temporelle . . . 18
1.5 Hypothèses de travail et démarche proposée pour les recherches . . . 21
1.5.1 Apports de l'état de l'art . . . 21
1.5.2 Hypothèse de travail . . . 21
1.5.3 Démarche envisagée . . . 21

Chapitre 2 Méthodes de cartographie spatio-temporelle 23

2.1 Introduction . . . 23
2.2 Formulation générale du problème . . . 24
2.3 Indicateurs de fiabilité et de robustesse de la cartographie . . . 25
2.4 Génération de données spatio-temporelles synthétiques . . . 26
2.5 Méthodes de cartographie spatiales . . . 29
2.5.1 Pondération inverse à la distance . . . 29

2.5.2 Complétion de matrice de rang faible avec contraintes de douceur spatiale 30
2.5.3 Krigeage . . . 30
2.6 Analyse comparative des méthodes à partir de données simulées . . . 33
2.7 Méthodes de cartographie spatio-temporelle . . . 37
2.7.1 Prise en compte de la dimension temporelle . . . 37
2.7.2 Régularisation tensorielle . . . 38
2.7.3 Réseaux de neurones informés par la physique . . . 43

Chapitre 3 Méthodes et stratégies de mesures **51**
3.1 Introduction . . . 52
3.2 Organisation générale des essais . . . 52
3.3 Description des espaces expérimentaux . . . 53
3.3.1 Chambre expérimentale . . . 54
3.3.2 Salle aménagée à l'INRS . . . 55
3.4 Stratégie de déploiement des instruments de mesures et des sources de pollution 56
3.4.1 Stratégie de mesure à points fixes . . . 56
3.4.2 Stratégie de mesure à points mobiles . . . 57
3.4.3 Stratégie de pollution . . . 57
3.4.4 Déroulement des essais . . . 57
3.5 Description des sources de pollution . . . 61
3.5.1 Source de pollution gaz . . . 61
3.5.2 Source de pollution particules . . . 61
3.6 Méthodes et instruments de mesures . . . 62
3.6.1 Réseau de capteurs sans fil gaz (RCSFg) . . . 62
3.6.2 Réseau filaire de capteurs particules (RFCp) . . . 64
3.6.3 Détecteurs de CO_2 portables = instruments de référence . . . 64
3.6.4 Thermo-hygromètre . . . 65
3.6.5 Scanner 3D . . . 65
3.6.6 Robot mobile . . . 66
3.6.7 Station météo . . . 66
3.7 Vérification et étalonnage des instruments de mesures . . . 67
3.7.1 Étalonnage des instruments de mesures gaz . . . 67
3.7.2 Étalonnage du réseau filaire de capteurs particules . . . 70
3.8 Pré-traitement des données . . . 72

Chapitre 4 Exploitation des données **75**
4.1 Introduction . . . 76
4.2 Application des méthodes de cartographie spatio-temporelle aux données réelles : choix des paramètres . . . 76
4.3 Étude des caractéristiques de la cartographie spatio-temporelle . . . 78

4.3.1 Influence de la méthode de cartographie spatio-temporelle 78
4.3.2 Influence de la grille de positionnement des points de mesure 94
4.3.3 Influence du nombre de points de mesure 101
4.3.4 Influence des caractéristiques de la source de pollution 109
4.3.5 Influence du débit d'émission 117
4.3.6 Influence du nombre de sources 123
4.3.7 Influence du type de polluant : essais avec un aérosol 127
4.4 Fusion de données : Représentation 3D de la chambre expérimentale avec les niveaux de concentrations 137
4.5 Fusion de données : Reconstruction des expositions individuelles 140
4.5.1 Objectifs et principes des essais 140
4.5.2 Méthodologie 141
4.5.3 Résultats et discussions 144
4.6 Applications en santé au travail : retombées indirectes des travaux de thèse . . . 162
4.6.1 Apport sur les stratégies de mesure 163
4.6.2 Périmètre de « validité » autour d'un point de mesure 165
4.6.3 Identification des zones de danger 169
4.6.4 Conclusion partielle 170

Conclusion et perspectives **171**

Annexes **177**

Publications **181**

Bibliographie **183**

Table des figures

1.1 Les composantes de l'analyse spatio-temporelle des expositions appliquée aux environnements professionnels . . . 8
1.2 Exemples de cartographie des concentrations d'aérosols en milieu professionnel obtenus à l'aide de la méthode Grid Map [O'Brien03] . . . 10
1.3 Exemple de cartographie de particules au sein de l'environnement professionnel(à gauche) avec une acquisition itinérante de données avec le capteur (+) déplacé d'une position a l'autre (à droite)[Evans08] . . . 11
1.4 Exemple de cartographie de concentrations de CO_2 générée par le logiciel Surfer avec la méthode de krigeage [Peters12] . . . 11
1.5 Exemple de cartographie multi-risques (Particules, CO, Bruit) réalisée avec la méthode IDW [Zuidema19b] . . . 14
1.6 Comparaison des moyennes de concentrations de particules estimés sur quatre intervalles de temps par les méthodes de krigeage et ANN [Li18] . . . 18
1.7 Démarche envisagée pour le développement et l'évaluation des méthodes de cartographie spatio-temporelle dans le cadre des travaux de la thèse . . . 22

2.1 Domaine D à cartographier avec Δ_x, Δ_y les résolutions spatiales et Δ_t la résolution temporelle, les points de mesures (en vert) et les points à estimer (en blanc). . . . 25
2.2 Exemple de cartes de pollution simulées représentant l'évolution temporelle (t = 1 à 12) de deux sources gaussiennes . . . 27
2.3 Série de 15 configurations de grilles d'échantillonnage (régulières, irrégulières, denses et éparses) de 10 points . . . 27
2.4 Série de 15 configurations de grilles d'échantillonnage (régulières, irrégulières, denses et éparses) de 15 points . . . 28
2.5 Série de 15 configurations de grilles d'échantillonnage (régulières, irrégulières, denses et éparses) de 20 points . . . 28
2.6 Série de 15 configurations de grilles d'échantillonnage (régulières, irrégulières, denses et éparses) de 25 points . . . 28
2.7 Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 10 points de la figure 2.3, obtenue par les méthodes IDW, krigeage et SLRMC . . . 34
2.8 Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 15 points de la figure 2.4, obtenue par les méthodes IDW, krigeage et SLRMC . . . 34
2.9 Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 20 points de la figure 2.5, obtenue par les méthodes IDW, krigeage et SLRMC . . . 35
2.10 Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 25 points de la figure 2.6, obtenue par les méthodes IDW, krigeage et SLRMC . . . 35

2.11 Cartes reconstruites à partir de 20 points des grilles d'échantillonnage présentées en figure 2.5 avec les méthodes IDW, krigeage et SLMRC. 36
2.12 Cartes reconstruites à partir de 20 points des grilles d'échantillonnage présentées en figure 2.5 avec les méthodes IDW, Krigeage, SLMRC et Régularisation. 43
2.13 Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 20 points de la figure 2.6, obtenue par les méthodes IDW, krigeage, SLRMC et Régularisation tensorielle 43
2.14 Exemple d'échantillonnage des points de mesures, de collocation et de frontières. 46
2.15 Architecture du réseau de neurones informé par la physique pour la résolution de PDE de type diffusion–réaction 47
2.16 Cartes reconstruites à partir de 20 points des grilles d'échantillonnage présentées en figure 2.5 avec les méthodes IDW, Krigeage, SLMRC, Régularisation et PINNs. 49
2.17 Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 20 points de la figure 2.6, obtenue par les méthodes IDW, krigeage, SLRMC, Régularisation tensorielle et PINNs 50

3.1 Protocole de mesure et de traitement des données pour la cartographie spatio-temporelle 52
3.2 Illustration de l'enceinte expérimentale 54
3.3 Photo de la salle aménagée dans les locaux de l'INRS 55
3.4 Exemple d'une grille d'échantillonnage indiquant les points de mesure (M1 à M20), les points de référence (R1 à R10) et le point de pollution (S1) 57
3.5 Représentation du plan d'expérience A avec les différentes phases chronologiques 58
3.6 Représentation des phases chronologiques du plan d'expérience B 58
3.7 Bouteille de CO2 (à gauche) et régulateur de débit massique [RDM] associé (à droite)) 61
3.8 Générateur de particules salines (NaCl [1%] (Palas AGK 2000)) 62
3.9 Illustration du réseau de capteurs sans fil gaz utilisé pour les essais (à gauche) et d'un module constituant le réseau de capteurs avec une batterie portable pour son alimentation(à droite) 63
3.10 Module capteur constituant le réseau Aranet 4 63
3.11 Capteur SPS30 Sensirion avec le dispositif de connexion filaire 64
3.12 Détecteur de CO_2 portable (ToxiRAE Pro CO_2) 64
3.13 Thermo-hygromètre (Testo 174H) 65
3.14 RTC360 LT Laser scanner 3D (Leica Geosystems) 65
3.15 Robot autonome mobile SHERPA 66
3.16 Station météorologique BWS200 67
3.17 Disposition des instruments de mesure dans l'enceinte de calibration 67
3.18 Analyse comparative des signaux enregistrés par les modules de capteurs et les détecteurs de CO_2 étalonnés suivant le plan de calibration A 69
3.19 Analyse comparative des signaux enregistrés par les modules de capteurs et les détecteurs de CO_2 étalonnés suivant le plan de calibration B 69
3.20 Analyse comparative des signaux enregistrés par les capteurs du réseau Aranet et les détecteurs de CO_2 étalonnés suivant le plan de calibration B 70
3.21 Enceinte de calibration équipée des capteurs SPS30, du générateur et de l'instrument de référence 71
3.22 Étapes du pré-traitement des données avant exploitation 72

4.1 RMSE en fonction des hyperparamètres de régularisation spatio-temporelle (λ, μ) 77
4.3 Température (°C) et Humidité relative de l'air (%) mesurées au cours de l'essai S1-E9 aux points de référence 79

4.4 Grille synthétique d'échantillonnage des points des essais S1-E8, S1-E9 (a) et S1-E12(c) et de la simulation CFD (b) . . . 81
4.5 Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) de l'essai S1-E8 . . . 82
4.6 Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) de l'essai S1-E9 . . . 82
4.7 Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) de l'essai S1-E12 . . . 83
4.8 Cartographies spatio-temporelles des concentrations de CO_2 et évaluation du RMSE aux instants $T = 25$ min, 40 min, 83 min, 95 min, 120 min de l'essai S1-E8 84
4.9 Cartographies spatio-temporelles des concentrations de CO_2 et évaluation du RMSE aux instants $T = 25$ min, 40 min, 83 min, 95 min, 120 min de l'essai S1-E9 85
4.10 Cartographies spatio-temporelles des concentrations de CO_2 et évaluation du RMSE aux instants $T = 25$ min, 40 min, 83 min, 95 min, 120 min de l'essai S1-E12 86
4.11 Analyse comparative des signaux enregistrés par les capteurs et les données issues de la simulation CFD . . . 88
4.12 Évolution temporelle du RMSE obtenues avec les méthodes de krigeage, de régularisation et de PINNs pour la reconstruction spatio-temporelle des concentrations de CO_2 des essais S1-E8 et S1-E9 . . . 89
4.13 Analyse des performances des méthodes (krigeage, Régularisation, PINNs) pour la reconstruction des concentrations de CO_2 au point de référence 1 | Essai S1-E9 90
4.14 Analyse des performances des méthodes (krigeage, Régularisation, PINNs) pour la reconstruction des concentrations de CO_2 au point de référence 5 | Essai S1-E9 91
4.15 Évolution temporelle du RMSE obtenues avec les méthodes de krigeage, de régularisation et de PINNs pour la reconstruction spatio-temporelle des concentrations de CO_2 de l'essai S1-E12 . . . 93
4.16 Quatres grilles d'échantillonnage avec 20 points de mesures . . . 94
4.17 Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) des essais S1-E9 et S1-E10 . . . 95
4.18 Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) des essais S1-E11 et S1-E13 . . . 96
4.19 Comparaison des performances (RMSE) des méthodes de krigeage, de régularisation tensorielle et de PINNs, calculées sur les 10 points de référence (R1–R10) et sur la durée totale des essais S1–E9, S1–E10, S1–E11 et S1–E13. . . . 99
4.21 Cartographies spatio-temporelles des concentrations de CO_2 et évaluation du RMSE à $T = 25$ min de l'essai S1-E9 avec 20, 15, 10 et 5 points de mesures . . . 103
4.22 Évolution temporelle du RMSE obtenues avec les méthodes de krigeage, de régularisation et de PINNs pour la reconstruction spatio-temporelle des concentrations de CO_2 de l'essai S1-E9 avec 20, 15, 10 et 5 points de mesures . . . 106
4.23 Grille d'échantillonnage régulière avec les modules capteurs du RCSFg en bleu (M1-M20), les détecteurs de référence en vert (R1-R10) et la source de pollution en rouge S1 . . . 110
4.24 Dispositif expérimental et source installés pour les essais S1-E1 à gauche (Source orientée vers le bas) et S1-E2 à droite (Source orientée vers le haut) dans la chambre expérimentale . . . 110

4.25 Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) au cours des essais S1-E1 et S1-E2 . . . 112
4.26 Comparaison des mesures de concentrations de CO_2 enregistrées au cours des essais S1-E1 et S1-E2 aux points de co-localisation des capteurs des réseaux RCSFg et les détecteurs de référence ; M6-R3 (a), M12-R6 (b) . . . 113
4.27 Cartographies spatio-temporelles des concentrations de CO_2 et évaluation du RMSE à $T = 35$ min de l'essai a)S1-E1 et b) S1-E2 . . . 113
4.28 Evolution des concentrations de CO_2 de référence et estimées par les méthodes de cartographie spatio-temporelle (Krigeage, Regularisation tensorielle, Pinns) au point de référence 2 pour l'essai S1-E1 (a) et l'essai S1-E2 (b) . . . 116
4.29 Grille d'échantillonnage régulière avec les modules capteurs du RCSFg en bleu (M1-M20), les détecteurs de référence en vert (R1-R10) et la source de pollution en rouge S1 . . . 117
4.30 Evolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) au cours des essais S1-E3, S1-E4 et S1-E5 . . . 119
4.31 Cartographies spatio-temporelles des concentrations de CO_2 à la phase de stabilisation pour les essais S1-E3 (a), S1-E4 (b), S1-E5 (c) . . . 121
4.32 Grilles d'échantillonnages régulières avec 2 sources de pollution (a) et 3 sources de pollution (b) . . . 123
4.33 Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) au cours des essais S1-E6 (2 sources), S1-E7 (3 sources) . . . 124
4.34 Cartographies spatio-temporelles des concentrations de CO_2 à la phase de stabilisation pour les esssais S1-E6 (a), S1-E7 (b) . . . 126
4.35 Installation du dispositif expérimental pour la réalisation de l'essai avec les microaérosols S2-E1 . . . 129
4.36 Grille de mesures de l'essai S2-E1 avec les microaérosols . . . 129
4.37 Évolution des concentrations numérique en particules de taille 0.5 µm (PM 0.5) et 1 µm (PM 1) enregistrées par les modules capteurs du RFCp (C1 à C20) au cours de l'essai S2-E1 . . . 131
4.38 Évolution des concentrations numérique et massique en particules de taille 10 µm (PM10) et 2,5 µm (PM 2,5) enregistrées par les modules capteurs du RFCp (C1 à C20) au cours de l'essai S2-E1 . . . 132
4.39 Cartographies spatio-temporelles des concentrations numérique des aérosols de taille 0.5 µm (a) et 1 µm (b) à $T = 21$ min de l'essai S2-E1 . . . 134
4.40 Cartographies spatio-temporelles des concentrations massiques de nanoaréosols de taille 10 µm (a) et 2.5 µm (b) à $T = 21$ min de l'essai S2-E1 . . . 134
4.41 Cartographies spatio-temporelles des concentrations en nombre en particules de taille 10 µm (a) et 2.5 µm (b) à $T = 21$ min de l'essai S2-E1 . . . 135
4.42 Représentation de l'intérieur de la chambre expérimentale en 3D avec les niveaux de concentrations en CO_2 les plus élevées aux instants $T = 45$ min (a) , 75 min (b) de l'essai S1-E2 . . . 139
4.43 Grille d'échantillonnage de l'environnement expérimental. Les points de mesure sont indiqués en bleu (M1 à M20), les points de référence en vert (R1 à R19), les points de sources de pollution en rouge (S1 à S3), en orange, la trajectoire à suivre par le robot équipé avec le capteur M21 et l'instrument de référence R20 . 141
4.44 Robot mobile SHERPA équipé des instruments de mesures au sein de l'environnement expérimental . . . 142

4.45 Température (°C) et Humidité relative de l'air (%) mesurées au point 10 sur la grille d'échantillonnage 4.43 pour les essais de la série 3. 145
4.46 Représentation des grilles de mesure suivant les deux scénarios de pollution . . . 146
4.47 Concentrations en CO_2 mesurées au cours des essais de la série 3 avec les capteurs du réseau Aranet (a), du RCSFg (b), détecteurs de référence (c) au point 10 de la grille d'échantillonnage 4.43. 147
4.48 Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) de l'essai S3-E3 148
4.49 Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) de l'essai S3-E3 148
4.50 Évolution des mesures de concentrations de CO_2 enregistrées par les capteurs du réseau Aranet (A1 à A20) de l'essai S3-E3 149
4.51 Évolution des mesures de concentrations de CO_2 enregistrées par les capteurs de référence (R1 à R19) de l'essai S3-E3 149
4.52 Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) de l'essai S3-E14 149
4.53 Évolution des mesures de concentrations de CO_2 enregistrées par les capteurs du réseau Aranet (A1 à A20) de l'essai S3-E14 150
4.54 Évolution des mesures de concentrations de CO_2 enregistrées par les détecteurs de référence (R2 à R19) de l'essai S3-E14 150
4.55 Comparaison des mesures de concentrations de CO_2 enregistrées au cours de l'essai S3-E3 aux points de co-localisation des capteurs des réseaux RCSFg et Aranet et les capteurs de référence ; M4-A4-R4 (a), M7-A7-R7 (b), M10-A10-R10 (c), M17-A17-R16 (d). 151
4.56 Comparaison des mesures de concentrations de CO_2 enregistrées au cours de l'essai S3-E14 aux points de co-localisation des capteurs des réseaux RCSFg et Aranet et les détecteurs de référence ; M4-A4-R4 (a), M7-A7-R7 (b), M10-A10-R10 (c), M17-A17-R16 (d). 152
4.57 Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase de pollution à $T = 25$ min de l'essai S3 E3 avec le réseau Aranet 153
4.58 Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase de stabilisation à $T = 90$ min de l'essai S3 E3 avec le réseau Aranet 153
4.59 Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase d'épuration à $T = 115$ min de l'essai S3 E3 avec le réseau Aranet 153
4.60 Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase de pollution à $T = 25$ min de l'essai S3 E3 avec le réseau RCSFg 154
4.61 Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase de stabilisation à $T = 90$ min de l'essai S3 E3 avec le réseau RCSFg 154
4.62 Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase d'épuration à $T = 115$ min de l'essai S3 E3 avec le réseau RCSFg 154
4.63 Évolution temporelle du RMSE obtenues avec les méthodes de krigeage, de régularisation et de PINNs pour la reconstruction spatio-temporelle des concentrations de CO_2 | Essai S3 E3 avec le réseau Aranet 156
4.64 Évolution temporelle du RMSE obtenues avec les méthodes de krigeage, de régularisation et de PINNs pour la reconstruction spatio-temporelle des concentrations de CO_2 | Essai S3 E3 avec le réseau RCSFg 156
4.65 Évolution temporelle de l'exposition individuelle au CO_2 reconstruite $\hat{\boldsymbol{X}}_\gamma(t)$ avec les méthodes de krigeage, de régularisation et de PINNs, comparée aux mesures de référence | Essai S3 E3 avec le réseau Aranet 158

4.66 Cartographie spatio-temporelle des concentrations de CO_2 avec le méthode de régularisation tensorielle pendant l'essai S3-E3 avec le réseau Aranet à $T = 50$ min, $T = 80$ min, $T = 100$ min représentant des instants où le vecteur mobile ▲ se trouve en bordure de l'espace étudié 158
4.67 Évolution temporelle de l'exposition individuelle au CO_2 reconstruite $\hat{\boldsymbol{X}}_\theta(t)$ avec les méthodes de krigeage, de régularisation et de PINNs, comparée aux mesures de référence | Essai S3 E3 avec le réseau RCSFg 159
4.68 Cartographie spatio-temporelle des concentrations de CO_2 avec le méthode de régularisation tensorielle pendant l'essai S3-E3 avec le réseau RCSFg à $t = 50$ min, $t = 80$ min, $t = 100$ min représentant des instants où le vecteur mobile ▲ se trouve en bordure de l'espace étudié 159
4.69 Corrélation entre exposition individuelle reconstruite $\hat{E}(T)$ et exposition de référence $E(T)$ | Essai S3 E3 avec le réseau Aranet 160
4.70 Corrélation entre exposition individuelle reconstruite $\hat{E}(T)$ et exposition de référence $E(T)$ | Essai S3 E3 avec le réseau RCSFg 161
4.71 Grille d'échantillonnage des points de mesures pour la cartographie spatio-temporelle avec M1-M20 les points de mesures du RCSFg, S1 la source de pollution, R1-R20 les points de référence 165
4.72 Représentation du périmètre de validité du point de référence R2 en fonction du niveau de corrélation spatiale (R^2) 166
4.73 Représentation du périmètre de validité du point de référence R5 en fonction du niveau de corrélation spatiale (R^2) 167

74 Synthèse des résultats obtenus et perspectives des travaux de recherche 175

Liste des tableaux

3.1 Description des séries d'essais réalisés . . . 53
3.2 Résumé des essais de la série 1 . . . 59
3.3 Résumé de l'essai de la série 2 . . . 60
3.4 Résumé de l'ensemble des essais de la série 3 . . . 60
3.5 Coefficients d'étalonnage des capteurs SPS30 . . . 71

4.1 Analyse des métriques (RMSE et R^2) pour les 10 points de référence (R1-R10) de l'essai S1-E8 . . . 90
4.2 Analyse des métriques (RMSE et R^2) pour les 10 points de référence (R1-R10) de l'essai S1-E9 . . . 91
4.3 Analyse des metriques (RMSE et R^2) pour les 10 points de référence (R1-R10) de l'essai S1-E12 . . . 92
4.4 Analyse des métriques (RMSE et R^2) pour les 10 points de référence (R1-R10) pour les essais S1-E9, S1-E10, S1-E11 et S1-E13 avec la méthode de krigeage . . . 97
4.5 Analyse des métriques (RMSE et R^2) pour les 10 points de référence (R1-R10) pour les essais S1-E9, S1-E10, S1-E11 et S1-E13 avec la méthode de régularisation 97
4.6 Analyse des metriques (RMSE et R^2) pour les 10 points de référence (R1-R10) pour les essais S1-E9, S1-E10, S1-E11 et S1-E13 avec la méthode de Pinns . . . 98
4.7 Comparaison des performances (RMSE et R^2) pour 10 points de référence – Krigeage . . . 104
4.8 Comparaison des performances (RMSE et R^2) pour 10 points de référence – Regularization . . . 104
4.9 Comparaison des performances (RMSE et R^2) pour 10 points de référence – PINNs 105
4.10 Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E1 . . . 114
4.11 Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E2 . . . 115
4.12 Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E3 . . . 119
4.13 Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E4 . . . 120
4.14 Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E5 . . . 120
4.15 Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E6 (2 sources) . . . 125
4.16 Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E7 (3 sources) . . . 126
4.17 Liste des essais de la série 3 réalisés . . . 144
4.18 Rayons de validité associés à chaque point de référence (R1-R10) en dm calculés avec les données de cartographie spatio-temporelle de l'essai S1-E8 . . . 167

19 Erreurs quadratiques moyennes (RMSE) et coefficients (R^2) entre les mesures des capteurs du réseau RCSFg acquises au cours de l'essai S1-E8 et les données CFD. 177
20 Erreurs quadratiques moyennes (RMSE) et coefficients (R^2) entre les mesures des modules capteurs du RCSFg acquises au cours de l'essai S1-E9 et les données CFD. 178
21 Erreurs quadratiques moyennes (RMSE) et coefficients (R^2) entre les mesures de référence et les données CFD. 178
22 Erreurs quadratiques moyennes (RMSE) et coefficients (R^2) entre les mesures des modules capteurs du réseau Aranet acquises au cours des essais de la série 3 et les données des détecteurs de référence 179
23 Erreurs quadratiques moyennes (RMSE) et coefficients (R^2) entre les mesures des modules capteurs du réseau RCSFg acquises au cours des essais de la série 3 et les données des détecteurs de référence 179

Liste des sigles et abréviations

RCSFg	Réseau de capteurs sans fil gaz
RFCp	Réseau filaire de capteurs particules
IMTR	Instrument de mesures temps réel
VLCT	Valeur limite d'exposition à court terme
VLEP	Valeur limite d'exposition professionnelle
PM	Particulate matter (Particules en suspension)
Ppm/v	Parties par million en volume (parts per million by volume)
QAI	Qualité de l'Air Intérieur
HR	Humidité Relative
RMSE	Root Mean Square Error (Erreur quadratique moyenne)
CO_2	Dioxyde de carbonne
CFD	Computational Fluid Dynamics (Dynamique des fluides numériques)
IDW	Inverse Distance Weighting
SLMRC	Smoothing Low-Rank Matrix Completion
ANN	Artificial Neural Networks
EDP	Equation aux dérivées partielles
PINNS	Physics Informed Neural Networks
FNN	Feedforward Neural Network
MLP	Multilayer Perceptron

Notations

$\boldsymbol{x}$	Vecteur
$\boldsymbol{X}$	Matrice
$\boldsymbol{\mathcal{X}}$	Tenseur
$\mathbf{1}_M$	Vecteur colonne de taille M dont toutes les composantes valent 1
$(\cdot)^{-1}$	Inverse d'une matrice
$(\cdot)^{\dagger}$	Pseudo-inverse de Moore-Penrose d'une matrice
$(\cdot)^{\mathsf{T}}$	Transposée d'une matrice/vecteur
$\boxdot$	Produit de Hadamard (produit élément par élément)
$\otimes$	Produit de Kronecker des matrices
$\underset{n}{\bullet}$	Produit n-mode d'un tenseur par une matrice
$\text{Diag}\{\cdot\}$	Opérateur construisant une matrice diagonale à partir des composantes du vecteur fourni
$\boldsymbol{I}_N$	La matrice identité de dimensions $N \times N$
$\|\cdot\|_F$	Norme de Frobenius d'une matrice ou d'un tenseur
$\|\cdot\|_2$	Norme ℓ_2 d'un vecteur
$\langle \cdot, \cdot \rangle$	Produit scalaire des vecteurs
$\text{vec}\{\cdot\}$	Opérateur de vectorisation d'une matrice ou d'un tenseur
$\text{unvec}\{\cdot\}$	Opérateur qui réalise l'opération inverse à l'opération $\text{vec}\{\cdot\}$
$\nabla f(\cdot)$	Gradient de f
$\nabla^2 f(\cdot)$	Laplacien de f
$\mathbb{E}[\cdot]$	Espérance mathématique
$\mathbb{R}_+$	Ensemble de vecteurs à valeurs réelles non-négatives
$\lfloor \cdot \rfloor$	Partie entière d'un nombre réel

Introduction générale

Le travail de recherche a été mené au laboratoire CRAN [anciennement Centre de Recherche en Automatique de Nancy - (CNRS UMR 7039)], dans l'équipe SiMul (Signaux Multidimensionnels) du département BioSiS et au laboratoire ASTEC (Analyse SaptioTemporelle des Expositions Chimiques], du département Ingénierie des Procédés (IP) de l'Institut National de Recherche et de Sécurité (INRS). Il s'inscrit dans le cadre d'une étude portant sur l'exploitation des données de mesures spatio-temporelle des polluants pour des applications en prévention à différents secteurs d'activités

L'objectif de cette collaboration était de développer une méthode automatique pour la cartographie spatio-temporelle des concentrations des polluants ainsi qu'une méthodologie pour l'évaluation des risques d'exposition des travailleurs, en utilisant les données acquises par différents types d'instruments de mesure déployés dans l'environnement de travail.

Contexte

L'exposition professionnelle aux polluants représente un enjeu majeur en santé publique. En fonction des environnements dans lesquels ils évoluent, les travailleurs peuvent être exposés à des nuisances incluant des substances chimiques, des agents biologiques et physiques. Ces expositions concernent divers secteurs d'activité tels que l'industrie (métallurgie, agroalimentaire, etc.), les bâtiments et travaux publiques (construction, réfection de voiries, etc.), l'agriculture (élevage, culture, traitement des grains, etc.), l'exploitation des ressources naturelles (minières, forestières, etc.) et bien d'autres. L'exposition peut se produire de différentes manières, à savoir l'inhalation, l'ingestion ou le contact cutané avec les substances et les agents.

Selon la nuisance considérée, la durée et l'intensité de l'exposition, celle-ci peut entraîner des effets néfastes sur la santé des travailleurs. Ainsi, l'exposition professionnelle peut provoquer divers symptômes respiratoires, généraux ou cutanés, se traduisant notamment par des maladies allergiques, des troubles neurologiques, des cancers ou encore des troubles mentaux. Les conséquences de ces maladies peuvent être sévères, parfois irréversibles, pouvant conduire à une invalidité permanente ou même au décès. D'après un rapport conjoint de l'OMS et de l'OIT, les maladies professionnelles liées à l'exposition aux polluants sont à l'origine de 7% des décès dans le monde [Pega22].

Il est donc important de mettre en place une stratégie visant à prévenir les risques professionnels liés à l'exposition aux polluants. Cela implique de surveiller les niveaux d'exposition aux substances chimiques, physiques et biologiques dans les environnements de travail, afin de respecter les seuils de sécurité établis par les autorités en santé au travail. Pour y parvenir, des mesures des polluants présents dans l'air des lieux de travail doivent être réalisées. Ces mesures permettent d'évaluer les expositions, d'identifier les sources de pollution et de vérifier l'efficacité des dispositifs de prévention.

Diverses méthodes et stratégies de mesure peuvent être mises en œuvre pour analyser les concentrations ambiantes et les niveaux d'exposition dans les environnements professionnels. Elles reposent sur plusieurs types d'acquisitions à savoir :

— L'acquisition ponctuelle de données, qui consiste à effectuer des mesures fixes d'un emplacement à un autre ;

— L'acquisition mobile de données qui permet de mesurer les concentrations des polluants dans l'air grâce à des équipements portés par des travailleurs ou installés sur des véhicules ou chariots mobiles ;

— L'acquisition continue de données qui utilise des réseaux de capteurs synchronisés pour mesurer en continu les niveaux des polluants à plusieurs emplacements et au cours du temps.

L'acquisition continue de mesures en temps réel a fait émerger des méthodes d'analyse et de représentation des données pour la surveillance environnementale et la prévention des risques professionnels, notamment à travers la cartographie spatio-temporelle.

L'analyse spatio-temporelle consiste à étudier les caractéristiques spatiales et les dynamiques temporelles des phénomènes en vue de mieux comprendre leur fonctionnement et mieux les gérer. L'étude des mécanismes régissant ces processus requiert des méthodes et des stratégies adaptées pour collecter les informations, ainsi que des méthodes pour la représentation cartographique et l'analyse statistique des données. Les données peuvent également être utilisées pour concevoir des modèles mathématiques visant à prédire les évolutions spatio-temporelles des phénomènes étudiés.

Les applications de l'analyse spatio-temporelle concernent historiquement l'étude du climat et des phénomènes climatiques ([Cantat05], [Mukwada18]) ainsi que de la navigation sur terre [Wang13], sur les océans [Chen17] et dans les airs [Buschmann16]. Aujourd'hui, divers domaines sont concernés. Il s'agit de la santé publique, l'épidémiologie ([Rainham10], [Perchoux13]), la médecine [Wang10], l'exposition des populations aux pollutions particulaires et gazeuses dans l'air extérieur et dans les habitations ([Steinle13], [Park17]) et aux rayonnements électromagnétiques [Aerts18], l'écologie des écosystèmes (évolution des populations animales, végétales et microbiennes) [Xue19], la gestion des risques industriels [Mandić-Rajčević16], etc.

Une application récente consiste à utiliser des instruments de mesure en temps réel et des

réseaux de capteurs pour surveiller la pollution de l'air en milieu urbain. Par exemple, la mise en place de capteurs innovants à faible coût permet de réaliser des mesures continues et simultanées en différents points des villes. Ces dispositifs offrent la possibilité d'analyser la variabilité spatiale et temporelle des particules atmosphériques, d'estimer l'impact relatif du trafic routier et des émissions liées à la combustion du bois sur les concentrations de particules extérieures, et d'identifier certains facteurs clés [Zikova17].

Dans le cadre de la prévention des risques chimiques et biologiques en milieu professionnel, la cartographie spatio-temporelle vise à décrire l'hétérogénéité des polluants dans l'espace et le temps. Elle est particulièrement utile dans des contextes où les émissions sont diffuses, mal localisées, ou lorsqu'il s'agit de gérer des expositions multiples à des polluants issus de sources et de modalités d'émission variées (comme les sites et sols pollués, les magasins, l'agroalimentaire, ou encore les centres de tri des déchets recyclables). Cette approche repose sur des mesures répétées dans le temps, effectuées en différents points de l'espace, une représentation graphique ou cartographique des niveaux de pollution et de leurs variations, ainsi qu'une analyse approfondie des données collectées pour une interprétation optimale.

La cartographie spatio-temporelle des expositions, des concentrations et des émissions représente une approche innovante en santé au travail. Elle offre la possibilité d'approfondir la compréhension des situations professionnelles à risque tout en fournissant des informations rapides sur les sources d'émission et les niveaux d'exposition, facilitant ainsi la conception et la mise en œuvre de mesures préventives adaptées. Cette méthode est particulièrement utile pour visualiser les concentrations ou les dangers, localiser les sources et les émissions, identifier les tâches exposantes, communiquer efficacement sur les risques, ou encore analyser la dispersion des polluants et l'efficacité des systèmes de ventilation dans les espaces de travail.

Dans ce contexte, l'émergence des réseaux de capteurs sans fil à faible coût a transformé les méthodes de collecte des données spatio-temporelles d'expositions aux polluants en santé au travail, ouvrant la voie à de nouvelles possibilités d'analyse. L'intégration de techniques d'interpolation a favorisé le développement de cartographies spatio-temporelles, permettant de visualiser les variations des concentrations de polluants dans l'espace et le temps. Ces avancées renforcent ainsi les capacités de surveillance et de prévention des risques d'exposition en milieu professionnel.

Cependant, les méthodes d'interpolation traditionnelles, comme l'interpolation linéaire ou le krigeage, reposent souvent sur des hypothèses simplificatrices concernant la distribution des données. Elles peuvent ainsi manquer de précision lorsque les phénomènes à modéliser sont complexes, non linéaires ou sujets à des variations rapides dans l'espace et le temps. De plus, ces méthodes sont particulièrement sensibles à la qualité et à la densité des données collectées. Dans des environnements où les capteurs sont distribués de manière inégale ou où les données

manquent de cohérence, les résultats produits par l'interpolation peuvent être biaisés, compromettant la fiabilité des cartographies obtenues.

Pour répondre à ces défis, nous proposons de tirer parti de méthodes récentes de traitement du signal, fondées sur des modèles tensoriels et les réseaux de neurones informés par la physique (PINNs, Physics-Informed Neural Networks) afin de capturer plus finement la dynamique complexe des données environnementales. Ces approches visent à renforcer la robustesse de la modélisation face aux phénomènes non stationnaires et aux variations rapides des concentrations mesurées, tout en améliorant la précision et la résolution des représentations spatio-temporelles obtenues. L'utilisation de ces nouvelles méthodes permet d'identifier des structures latentes et d'extraire des informations physiquement cohérentes, contribuant ainsi à une modélisation plus fidèle des concentrations de polluants en milieu de travail.
Il devient possible de créer un outil de suivi de la dynamique d'évolution de la pollution dans les environnements de travail avec ces méthodes combinées aux techniques de fusion de données visant à intégrer des informations provenant de diverses sources (capteurs, système de géolocalisation, modèles de numérisation 3D).

Objectifs des travaux envisagés

Les objectifs de cette thèse sont les suivants :

— Développer et évaluer une méthode automatique pour la cartographie spatio-temporelle des concentrations en polluants présents dans l'air des lieux de travail : Il s'agit de mettre en place une méthodologie pour la modélisation et l'analyse des données spatio-temporelles, avec une prise en compte explicite de la continuité temporelle lors de la construction des cartes de pollution.

— Proposer des indicateurs ou des métriques permettant d'évaluer la fiabilité des cartes produites

— Développer une méthodologie de fusion de la cartographie spatio-temporelle avec des données de sources hétérogènes (Modélisation 3D, Données de géolocalisation, etc.) pour des applications en santé au travail (Estimation des expositions individuelles des travailleurs, identification des zones de danger, etc.).

La suite de ce manuscrit de thèse est structurée en quatre chapitres. Le premier présente l'état de l'art sur les méthodes de cartographie spatio-temporelle existantes et leurs utilisations dans le domaine de la santé au travail. Ce chapitre fournit le cadre conceptuel sur lequel re-

posent les développements méthodologiques de cette thèse.

Le deuxième chapitre est consacré au cœur méthodologique de la thèse. Il y est question des limites des approches existantes en matière de cartographie spatio-temporelle, en particulier dans le contexte de la santé au travail. À travers une étude comparative approfondie, ce chapitre propose le développement de deux nouvelles méthodes permettant de répondre à ces limitations : une approche fondée sur une formulation tensorielle régularisée et une autre basée sur un réseaux de neurones informé par la physique (PINNs). Les performances de ces nouvelles méthodes sont analysées à l'aide de données simulées, afin d'évaluer leur robustesse et leur fiabilité dans divers scénarios.

Le troisième chapitre décrit les matériels et méthodes utilisés pour la collecte des données exploitées dans cette thèse. Cette section met en évidence les dispositifs expérimentaux utilisés au cours des campagnes de mesures et leurs particularités techniques pouvant assurer la reproductibilité des expériences menées.

Le dernier chapitre se concentre sur l'exploitation des données expérimentales recueillies. Il expose les résultats obtenus à partir de cas concrets d'exposition aux polluants en conditions expérimentales et explore les diverses applications pratiques des méthodes développées. L'objectif est de fournir aux experts en santé au travail des outils efficaces pour la prévention des risques d'exposition, facilitant ainsi l'identification des zones à risque et l'optimisation des stratégies de gestion des expositions.

Enfin, la conclusion fait la synthèse des travaux et présente les perspectives envisagées.

Chapitre 1

État de l'art de la cartographie des concentrations de polluants en milieu de travail

Sommaire

1.1 Introduction . . . 7

1.2 Utilisation de la cartographie spatio-temporelle en santé au travail 9

1.2.1 Cartographie spatiale avec une acquisition itinérante des données . . . 9

1.2.2 Vers l'acquisition spatiale multi-points simultanée . . . 12

1.2.3 Cartographie spatio-temporelle avec une acquisition simultanée des données . . . 12

1.3 Méthodes de cartographie spatio-temporelle des expositions aux polluants en santé au travail . . . 15

1.3.1 Historique des méthodes de cartographie spatio-temporelle . . . 15

1.4 Analyse de l'état de l'art sur les méthodes de cartographie spatio-temporelle . . . 18

1.5 Hypothèses de travail et démarche proposée pour les recherches . . 21

1.5.1 Apports de l'état de l'art . . . 21

1.5.2 Hypothèse de travail . . . 21

1.5.3 Démarche envisagée . . . 21

1.1 Introduction

Le cadre contextuel de ma thèse s'inscrit dans une démarche de modélisation spatio-temporelle des expositions aux polluants atmosphériques en milieu professionnel. Cette approche implique

une compréhension globale du processus de modélisation (Voir Figure 1.1), depuis la conception de stratégies de mesure adaptées jusqu'à l'exploitation des résultats et leur représentation cartographiques à des fins de prévention des risques.

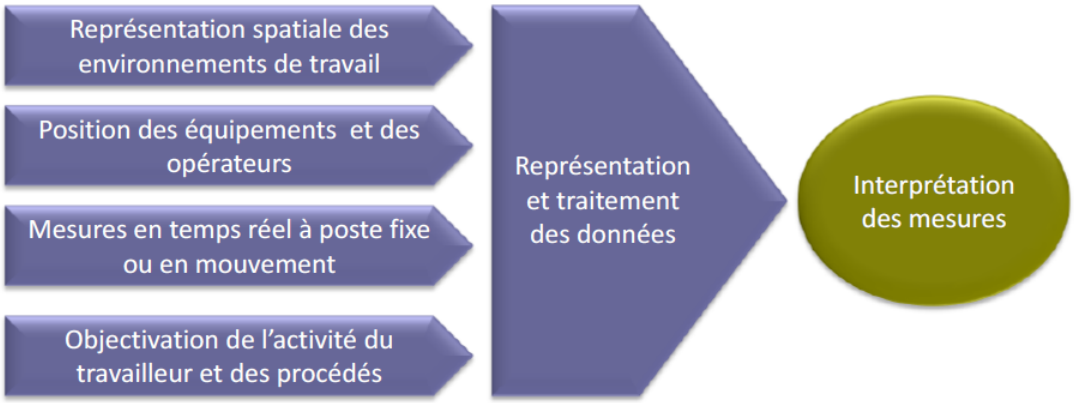


FIGURE 1.1 – Les composantes de l'analyse spatio-temporelle des expositions appliquée aux environnements professionnels

Plus précisément, il s'agit d'analyser :

- La définition des stratégies de collecte des niveaux de concentration, en fonction des contraintes opérationnelles et des objectifs de surveillance,
- La méthodologie d'acquisition et de traitement des données issue des réseaux de capteurs ou autres instruments de mesures,
- La modélisation et la représentation spatio-temporelle des concentrations, permettant une visualisation fine de la dynamique des polluants dans l'espace et dans le temps,
- La fusion de ces données avec d'autres sources d'informations (paramètres environnementaux, données de localisation spatiale des individus) en vue d'applications élargies, par exemple pour la reconstruction de l'exposition individuelle.

Pour ce faire, nous réaliserons dans ce chapitre un état de l'art des approches méthodologiques utilisées pour la cartographie spatio-temporelle des concentrations de polluants en milieu professionnel. La revue s'articule autour de trois axes complémentaires correspondant aux évolutions récentes dans le domaine : l'analyse des méthodes de collecte (acquisition itinérante, multi-points simultanée et acquisition continue via des capteurs low-cost), l'évaluation des approches de modélisation et de visualisation des expositions, puis l'identification des limites actuelles afin de définir les perspectives d'amélioration qui structureront la stratégie de recherche et les hypothèses de travail.

1.2 Utilisation de la cartographie spatio-temporelle en santé au travail

Initialement consacrée à l'étude des phénomènes climatiques, l'analyse spatio-temporelle a été appliquée progressivement à des domaines tels que la santé publique, l'épidémiologie et la surveillance des expositions environnementales. D'abord appliquée aux environnements urbains, la cartographie des expositions a été transposée aux environnements professionnels avec l'apparition de capteurs miniaturisés et à faible coût. Ainsi l'essor de l'Internet des objets (IoT) et des technologies embarquées a levé les contraintes de l'acquisition itinérante des données d'expositions et a permis de surmonter les limitations associées aux systèmes fixes permettant la conception de diverses stratégies de mesures plus flexibles et adaptatives pour une caractérisation spatio-temporelle fine et dynamique des expositions en situation réelle.

Les analyses bibliométriques confirment cette avancée avec une augmentation substantielle des publications scientifiques consacrées à la cartographie spatio-temporelle appliquée à la santé au travail. Cette progression soutenue depuis les années 2000 reflète à la fois un intérêt scientifique croissant et un élargissement des champs d'application.

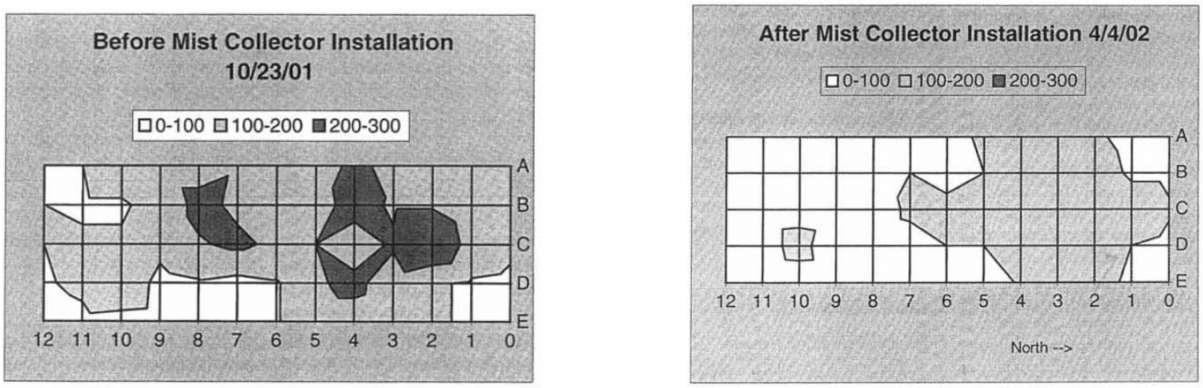
1.2.1 Cartographie spatiale avec une acquisition itinérante des données

L'acquisition itinérante des données d'expositions consiste à effectuer des relevés de concentrations de polluants en déplaçant un ou plusieurs appareils de mesures sur différents points d'un espace donné, selon un protocole de déplacement prédéfini. Cette méthode permet de cartographier la variabilité spatiale des concentrations dans un environnement sans recourir à un déploiement fixe et dense de capteurs.

Aux prémices de l'application de la cartographie spatio-temporelle en santé au travail, l'acquisition itinérante s'est imposée comme une approche privilégiée dans de nombreux contextes de surveillance environnementale et industrielle notamment en raison du coût élevé et de la complexité des appareils de mesure disponibles. Les instruments utilisés pour la quantification des polluants atmosphériques tels que le dioxyde de carbone (CO_2), les particules en suspension et les composés organiques volatils (COV), étaient caractérisés par leur encombrement, leur forte consommation énergétique et la nécessité d'une maintenance régulière, rendant leur déploiement en réseau fixe économiquement et opérationnellement peu viable.

Les premières applications en santé au travail étaient donc des cartographies spatiales avec une acquisition itinérante des données. Elles reposaient sur la méthode GridMap, utilisée pour visualiser les concentrations de polluants sur une grille régulière à partir de mesures ponctuelles. Cette approche a permis de cartographier les concentrations de gaz à l'aide de détecteurs à photoionisation [Andersson93] et d'évaluer la performance de systèmes de ventilation pour le contrôle du styrène dans des ateliers de laminage [Rosen90]. De manière similaire, les concen-

trations de mycobactéries ont été cartographiées dans une usine mécanique à partir de mesures réalisées sur une grille à hauteur des voies respiratoires, avec une représentation sous forme de cartes de contours via Microsoft Excel [O'Brien03](Voir figure 1.2). Ces cartes ont servi à la fois à localiser les sources de pollution et à vérifier l'efficacité des stratégies de prévention mises en place.



(a) Cartographie d'aérosols avant la ventilation

(b) Cartographie d'aérosols après la ventilation

FIGURE 1.2 – Exemples de cartographie des concentrations d'aérosols en milieu professionnel obtenus à l'aide de la méthode Grid Map [O'Brien03]

D'autres travaux ont suivi cette approche, notamment pour cartographier les concentrations massiques et numériques de particules dans une usine d'assemblage de moteurs à partir de 212 points mobiles, également visualisées par Excel [Liu10].

Toutefois, la méthode GridMap présente des limites inhérentes : une seule mesure par zone sans interpolation spatiale, entraînant une faible continuité des concentrations sur l'espace étudié. Pour améliorer la représentativité des données, certaines études ont eu recours à des analyses statistiques telles que l'ANOVA, utilisée pour évaluer les différences significatives entre les moyennes de concentrations mesurées dans une usine automobile à l'aide de chariots mobiles équipés de compteurs optiques et à condensation [Peters06].

L'utilisation du logiciel Surfer (Golden, CO) a marqué une avancée, en intégrant l'interpolation par krigeage, méthode géostatistique reposant sur la covariance spatiale [Wackernagel03]. Cette approche a permis de produire des cartes continues des concentrations en nombre et en masse d'aérosols dans plusieurs environnements industriels [Peters06], [Evans08] (Voir figure 1.3), ainsi que des concentrations de CO_2 et de particules dans des usines de confection [Peters12] (voir figure 1.4), des bâtiments d'élevage [Vosburgh11] et des restaurants [Park10].

Malgré ces améliorations, ces outils présentent des limitations, notamment un contrôle limité de la résolution spatiale et l'absence d'intégration de la dimension temporelle. Les dynamiques temporelles ont été approximées par la production manuelle de séries de cartes spatiales à différents instants t.

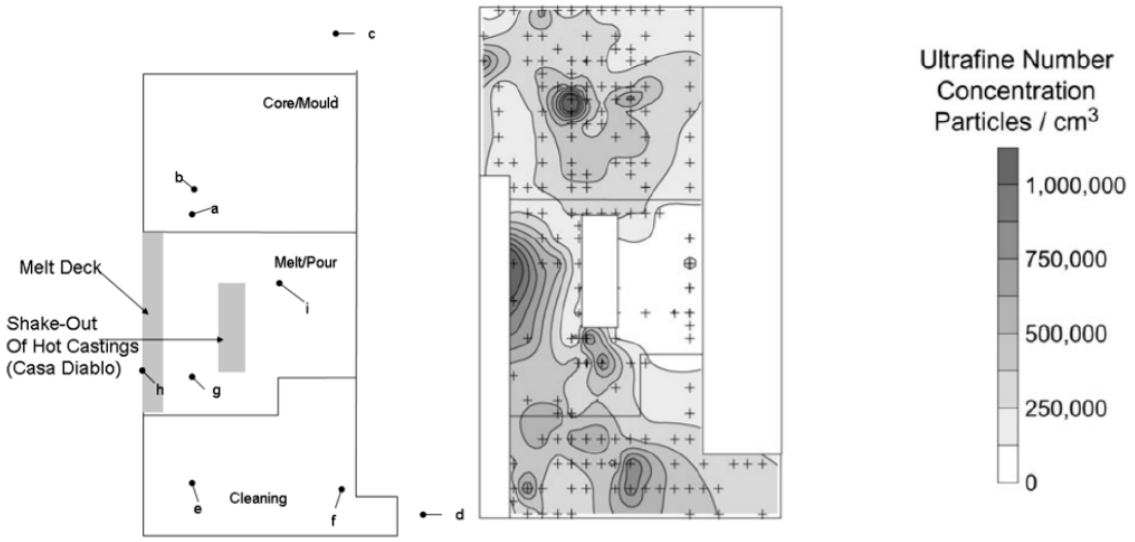


FIGURE 1.3 – Exemple de cartographie de particules au sein de l'environnement professionnel(à gauche) avec une acquisition itinérante de données avec le capteur (+) déplacé d'une position a l'autre (à droite)[Evans08]

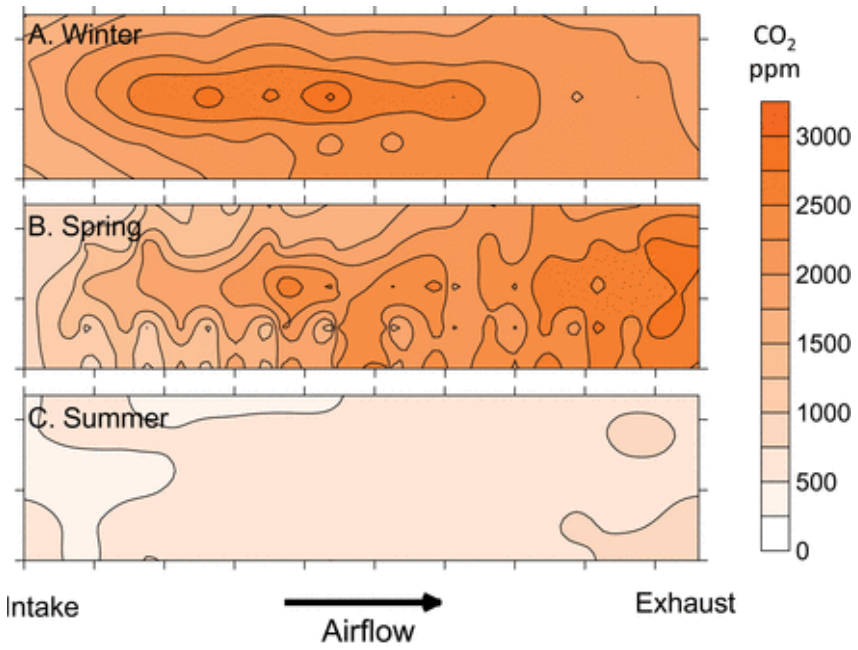


FIGURE 1.4 – Exemple de cartographie de concentrations de CO_2 générée par le logiciel Surfer avec la méthode de krigeage [Peters12]

Ces contraintes méthodologiques ont orienté les recherches récentes vers des outils de cartographie plus flexibles, intégrant pleinement les dimensions spatiales et temporelles des expositions professionnelles.

1.2.2 Vers l'acquisition spatiale multi-points simultanée

Lake et *et al.* ont conduit une étude visant à cartographier les expositions au bruit dans deux environnements industriels distincts : une usine de plasturgie et un site d'essai de moteurs. Une stratégie de déploiement hybride a été mise en œuvre, combinant des mesures itinérantes, réalisées en déplacement d'un point à un autre, et des mesures simultanées en poste fixe, à l'aide de dosimètres personnels et de moniteurs statiques [Lake15]. Cette approche a permis de capturer à la fois la variabilité spatiale et les niveaux d'exposition individuels. Pour assurer la cohérence spatiale des données, la méthode de krigeage a été appliquée, utilisant des variogrammes pour modéliser la dépendance spatiale en fonction de la distance entre points d'échantillonnage. Le traitement des données a été réalisé sous Matlab, permettant l'ajustement des modèles de variogramme et le calcul de poids optimaux pour l'estimation spatiale. Les analyses ont démontré que le modèle sphérique offrait la meilleure représentation de la structure spatiale des niveaux d'exposition acoustique dans les environnements étudiés.

Cette étude a constitué une contribution méthodologique importante, soulignant le rôle central de la résolution spatiale dans la cartographie des expositions professionnelles. Les cartes produites se sont révélées être des outils opérationnels à la fois pour la communication du risque, grâce à leur lisibilité, et pour l'aide à la décision dans la définition de stratégies de prévention ciblées. Toutefois, les auteurs rappellent que les cartographies basées uniquement sur des données itinérantes, bien que faciles à mettre en œuvre, demeurent limitées car elles négligent la variabilité temporelle souvent significative en milieu professionnel [Koehler11, Lake15].

Des travaux plus récents ont mis en évidence l'intérêt d'une cartographie spatio-temporelle intégrée, permettant d'interpoler les expositions à la fois dans l'espace et dans le temps, offrant ainsi une représentation plus fidèle de l'évolution des niveaux d'exposition [Ludwig17]. Cette approche est renforcée par l'utilisation combinée de capteurs itinérants et de capteurs statiques collectant des séries temporelles continues, permettant ainsi d'élargir la couverture spatio-temporelle des mesures [Koehler13].

1.2.3 Cartographie spatio-temporelle avec une acquisition simultanée des données

1.2.3.1 Avènement des réseaux de capteurs « low-cost » sans fil

La disponibilité croissante de réseaux de capteurs électrochimiques et optiques à coût modéré permet désormais d'appréhender avec finesse la dynamique temporelle des polluants dans les

environnements de travail. En adoptant des configurations multi sites, simultanées et pilotées en temps réel, il devient possible de caractériser les gradients de concentration, d'identifier les pics transitoires et de corréler ces derniers aux activités opératoires. À titre d'exemple, une étude menée en 2018 a exploré la répartition spatio-temporelle des concentrations particulaires dans un atelier de menuiserie à l'aide d'un réseau de huit capteurs stratégiquement positionnés à proximité des sources potentielles de pollution [Li18]. Des mesures continues des concentrations en nombre et en masse de particules ont été collectées sur une période de 4000 secondes.

Le traitement des données s'est appuyé sur trois méthodes à savoir un krigeage implémenté dans Statgraphics Centurion XVII (StatPoint, Inc., 2014) avec ajustement du variogramme selon six modèles théoriques (gaussien, exponentiel, circulaire, pentasphérique, fonction puissance et sphérique), une interpolation spline biharmonique sous MATLAB, ainsi qu'un réseau de neurones artificiels destiné à la reconstruction non linéaire des champs de concentration. Ces méthodes ont permis de produire des cartographies spatio-temporelles détaillées, capturant les fluctuations des concentrations massiques lors des opérations génératrices de poussières (sciage, perçage, ponçage), contribuant ainsi à l'élaboration de mesures préventives ciblées.

Dans la même perspective, un réseau de capteurs sans fil dédiés à la mesure des particules en suspension a été installé dans un environnement résidentiel. Les séries temporelles obtenues ont été traitées sous SPSS Statistics 24 (IBM Corp.) : la détection et la quantification des pics de concentration ont reposé sur la méthode de la ligne tangente, tandis qu'une régression multivariée a permis de modéliser la variabilité spatio-temporelle. L'analyse a corrélé les épisodes de concentration élevée aux événements sources de pollution, caractérisé la granulométrie associée et conduit à la recommandation d'un système de ventilation mécanique, afin de réduire durablement la charge particulaire [Shen21].

1.2.3.2 Réseaux de capteurs multi-risques

En 2019, s'appuyant sur les résultats prometteurs obtenus dans des études antérieures portant sur l'efficacité des réseaux de capteurs low-cost sans fil, Zuidema *et al.* ont initié le déploiement d'un réseau multirisques composé de capteurs low-cost conçu pour des environnements industriels. Ainsi, au sein d'une usine spécialisée dans la fabrication de véhicules lourds, quarante capteurs multirisques capables de mesurer simultanément des paramètres tels que les gaz (notamment le dioxyde de carbone), les particules en suspension et le bruit, ont été stratégiquement déployés selon une répartition régulière optimisée, conformément aux recommandations des études antérieures menées dans des contextes similaires [Sousan18, Berman18]. Les capteurs ont assuré une collecte continue des concentrations de dioxyde de carbone (CO_2), des concentrations particulaires (en nombre et en masse), des niveaux de pression acoustique, ainsi que des variables environnementales telles que la température et l'humidité, sur une durée ininterrompue de cinq mois [Zuidema19b]. Cette campagne de mesure a généré une base de données centralisée

comportant plus de 2,12 millions d'enregistrements, exploitée pour produire des cartographies spatio-temporelles via la méthode d'interpolation par pondération inverse à la distance (Inverse Distance Weighting, IDW). Bien que cette méthode ne prenne pas explicitement en compte les corrélations spatiales et temporelles intrinsèques aux données, les cartes produites (Voir figure 1.5) ont néanmoins permis de vérifier la cohérence interne des données issues du réseau et d'identifier avec précision les points critiques correspondant aux différentes sources d'exposition industrielle.

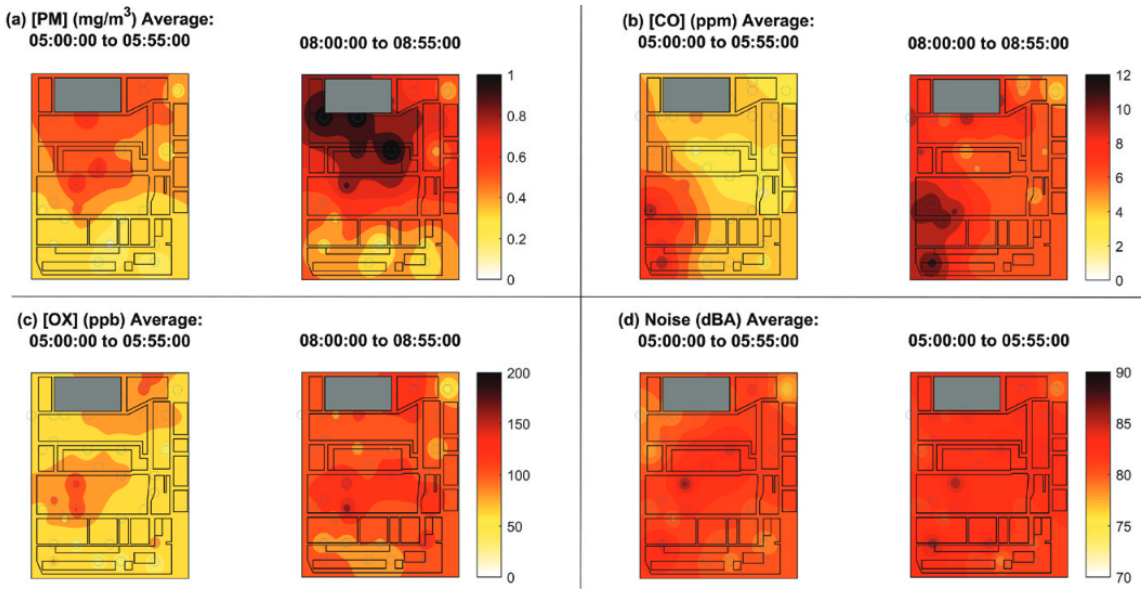


FIGURE 1.5 – Exemple de cartographie multi-risques (Particules, CO, Bruit) réalisée avec la méthode IDW [Zuidema19b]

La richesse et la volumétrie exceptionnelles des données obtenues par cette étude ont facilité, en 2021, les travaux de Patton *et al.*, qui ont entrepris une modélisation avancée des mesures de référence de CO_2, complétée par d'autres variables environnementales et de localisation [Patton22]. La modélisation prédictive a été réalisée à l'aide d'une approche par arbres de décision probabilistes à gradient renforcé (Gradient Boosted Decision Trees, GBDT), avec l'implémentation de la bibliothèque open-source NGBoost développée par le Stanford Machine Learning Group.

Les prédictions issues du modèle NGBoost ont été interpolées par pondération inverse de la distance (IDW) pour générer des cartographies prédictives. La performance des modèles a été évaluée à l'aide de deux métriques standards : l'erreur quadratique moyenne (RMSE) et le score probabiliste continu classé (CRPS), calculés sur un jeu de données de test indépendant, exclu de la phase d'apprentissage, assurant ainsi la validation des prédictions.

1.3 Méthodes de cartographie spatio-temporelle des expositions aux polluants en santé au travail

Dans l'état de l'art, les méthodes les plus utilisées pour la cartographie spatio-temporelle nécessitent l'estimation des concentrations aux emplacements non échantillonnés au sein de l'environnement étudié. Ces approches s'appuient sur des techniques d'interpolation, exploitant la variabilité spatiale et temporelle des valeurs mesurées pour représenter des champs de concentration.

1.3.1 Historique des méthodes de cartographie spatio-temporelle

Dans cette section nous ferons une synthèse des principales méthodes utilisées pour la cartographie spatio-temporelle des concentrations de polluants, présentées autant que possible selon un ordre chronologique. L'objectif est de mettre en évidence l'évolution des approches méthodologiques et les spécificités techniques associées à chaque période, en s'appuyant sur les contributions majeures identifiées dans la littérature scientifique.

— **La pondération inverse à la distance (Inverse Distance Weighting, IDW) :** L'interpolation par IDW constitue l'une des approches les plus répandues pour l'estimation des valeurs aux points non échantillonnés. Cette méthode repose sur le principe selon lequel la contribution d'un point de mesure à l'estimation d'une valeur diminue avec l'éloignement spatial. Concrètement, un poids est attribué à chaque point de données mesuré, calculé en fonction de la distance qui le sépare du point d'estimation selon une fonction décroissante. La valeur interpolée est ensuite obtenue par une moyenne pondérée des observations connues, pondération étant inversement proportionnelle à la distance. IDW est couramment employée dans l'analyse de données environnementales telles que la température, la qualité de l'air, la pollution de l'eau, ou encore la distribution spatiale de la faune. Son usage s'étend également aux contextes industriels, notamment dans le domaine de la santé au travail, pour la cartographie des expositions aux polluants atmosphériques. À titre d'exemple, cette méthode a été utilisée par [Zuidema19b] pour cartographier des expositions multiples incluant des gaz, des particules et le bruit dans une usine de fabrication de véhicules lourds, ainsi que par [Patton22] pour la cartographie des concentrations de CO_2 dans un environnement automobile.

Dans certaines applications spécifiques, notamment pour des expositions caractérisées par une atténuation suivant la loi du carré inverse (telles que le rayonnement ou le bruit), un facteur de pondération inversement proportionnel au carré de la distance est privilégié. Toutefois, cette pondération peut ne pas convenir à d'autres types de polluants, comme les aérosols, dont la dispersion spatiale obéit à des mécanismes diffusifs plus complexes. Il est également important de souligner que l'augmentation du nombre de points voisins

intégrés dans le calcul a pour effet de lisser la carte interpolée, atténuant l'influence des valeurs extrêmes. Enfin, il est à noter que les valeurs interpolées peuvent différer des mesures effectives aux points d'échantillonnage en raison de l'influence exercée par les observations environnantes [Koehler13].

— **Le krigeage :** C'est une méthode d'interpolation spatiale fondée sur la théorie des processus stochastiques stationnaires. Il repose sur l'exploitation de la corrélation spatiale des données, modélisée par le variogramme, pour estimer une variable à des emplacements non échantillonnés. Du point de vue statistique, il constitue l'estimateur linéaire non biaisé à variance minimale. La valeur à estimer en un point donné est calculée comme une combinaison linéaire des observations disponibles, pondérées selon leur position relative et leur interdépendance spatiale. Ces pondérations sont déterminées à partir d'un variogramme expérimental, fonction caractérisant la structure de dépendance spatiale du phénomène étudié. Plusieurs modèles de variogrammes peuvent être utilisés, en fonction de la nature des données :

  — Modèle sphérique : la covariance diminue avec la distance jusqu'à un certain seuil au-delà duquel elle devient nulle. Il est adapté aux phénomènes présentant une corrélation locale forte qui décroît rapidement.
  — Modèle exponentiel : la décroissance de la covariance est continue et asymptotique, sans seuil de coupure net, ce qui permet de modéliser une influence spatiale persistante mais décroissante.
  — Modèle gaussien : similaire au modèle exponentiel mais avec une décroissance plus douce et symétrique, souvent utilisé pour des phénomènes présentant une variabilité spatiale régulière.
  — Modèle linéaire : suppose une décroissance proportionnelle à la distance, adapté à des phénomènes dont la variabilité est particulièrement régulière.

  Le choix du modèle de variogramme repose généralement sur des connaissances a priori du phénomène observé.

  Le krigeage est largement utilisé pour la cartographie des expositions aux polluants dans les environnements de travail. On le retrouve notamment dans les études suivantes :

  — Cartographie des concentrations en particules (nombre et masse) dans un centre d'usinage de moteurs (2006–2007) [Peters06], [Heitbrink07] ;
  — Analyse spatio-temporelle des concentrations en CO_2 et particules dans une fonderie de fonte grise (2008) [Evans08] ;
  — Cartographie des concentrations particulaires dans un restaurant et une usine de moulage sous pression d'aluminium (2010) [Park10] ;
  — Étude d'une usine de confection (2011) [Vosburgh11] ;
  — Analyse des concentrations en CO_2 et particules dans une grange de gestation porcine

(2012) [Peters12] ;

— Cartographie du bruit dans une usine de plastique et un centre d'essais moteurs (2015) [Lake15] ;

— Étude d'un atelier de menuiserie (2018) [Li18].

Selon les hypothèses formulées sur le champ spatial, plusieurs variantes du krigeage peuvent être mobilisées :

— Krigeage simple : suppose la connaissance préalable de la moyenne et de la covariance du champ, stationnaire sur l'ensemble du domaine.

— Krigeage ordinaire : n'exige pas la connaissance de la moyenne, mais suppose qu'elle est constante localement.

— Krigeage universel : autorise l'intégration de tendances spatiales complexes, modélisées par des fonctions déterministes (ex. polynômes).

— Co-krigeage : permet l'estimation conjointe de plusieurs variables spatiales corrélées, en exploitant leur covariance croisée.

— **L'interpolation par spline biharmonique :** C'est une méthode d'interpolation fondée sur la théorie des équations aux dérivées partielles, spécifiquement adaptée aux problèmes de reconstitution de champs scalaires à partir de données discrètes. Elle repose sur l'ajustement de fonctions de Green associées à l'opérateur biharmonique (le laplacien appliqué deux fois), centré sur chaque point de mesure, afin d'estimer une surface lisse sur l'ensemble du domaine spatial. Contrairement aux approches classiques telles que l'interpolation linéaire ou polynomiale, la spline biharmonique ne requiert ni le choix d'un degré de polynôme ni l'ajustement de paramètres spécifiques. Cette absence de configuration a priori améliore sa robustesse et réduit le risque d'oscillations non physiques, en particulier dans les zones à forte variation locale. Sa formulation comme solution d'une équation différentielle partielle du quatrième ordre confère à l'interpolant une régularité élevée et un effet de lissage naturel, qui la rend particulièrement adaptée à la cartographie de phénomènes bruités ou présentant des gradients modérés.

Cette méthode a été appliquée dans l'étude de [Li18] pour visualiser l'évolution spatio-temporelle des concentrations massiques de particules en suspension dans un atelier de menuiserie. Elle a permis de générer des champs continus à partir de mesures ponctuelles, facilitant ainsi l'interprétation des dynamiques de dispersion dans l'environnement de travail.

— **Les réseaux de neurones artificiels (Artificial Neural Networks, ANN) :** Ce sont des modèles d'apprentissage supervisé capables de modéliser des relations non linéaires complexes entre des variables d'entrée et de sortie. Leur flexibilité architecturale les rend adaptés à un large éventail de tâches, incluant la classification, la régression, ou la modélisation de systèmes dynamiques. Dans l'étude présentée par [Li18], un réseau à

propagation avant a été utilisé pour la cartographie spatio-temporelle des concentrations en particules (nombre et masse) dans un atelier de menuiserie. Le modèle a été entraîné pour prédire les concentrations massiques à partir des coordonnées spatiales (x, y) et du temps t, en utilisant une architecture composée de deux couches cachées de 30 neurones, avec une activation de type tangente hyperbolique. Les résultats obtenus par cette approche sont comparés à ceux du krigeage sur différents intervalles temporels, comme illustré à la Figure 1.6, mettant en évidence des performances comparables entre les deux méthodes.

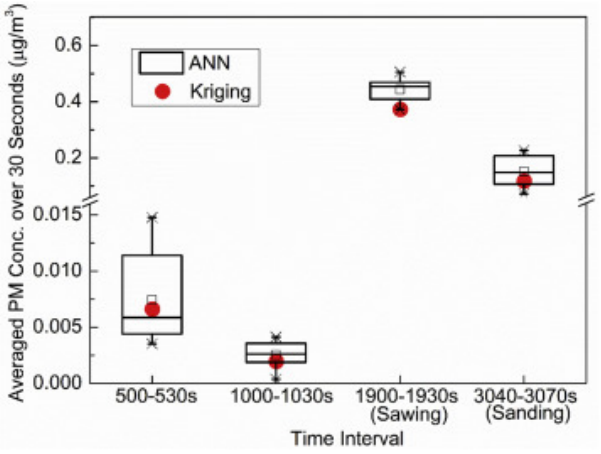


FIGURE 1.6 – Comparaison des moyennes de concentrations de particules estimés sur quatre intervalles de temps par les méthodes de krigeage et ANN [Li18]

Si les résultats obtenus montrent une bonne capacité d'ajustement, les ANN présentent certaines limites, notamment un manque d'interprètabilité et une intégration implicite, voire absente, des structures de dépendance spatiale et temporelle. Ces caractéristiques peuvent limiter leur performance en généralisation, en particulier par rapport aux méthodes géostatistiques comme le krigeage, qui exploitent explicitement la corrélation spatiale des observations.

1.4 Analyse de l'état de l'art sur les méthodes de cartographie spatio-temporelle

La cartographie spatio-temporelle des expositions aux polluants en milieu de travail a connu une évolution marquée par la montée en complexité des modèles, depuis des approches purement déterministes jusqu'aux techniques d'apprentissage profond. Cette progression reflète la nécessité croissante d'intégrer à la fois les structures de dépendance spatiale, la dynamique temporelle et les contraintes opérationnelles propres aux environnements de travail.

Approches déterministes fondées sur la distance

Les estimateurs qui pondèrent les observations en fonction de l'éloignement permettent une mise en œuvre rapide et un coût de calcul faible. Leur hypothèse isotrope, la décroissance uniforme de l'influence avec la distance, ne permet toutefois ni de saisir l'anisotropie (flux d'air préférentiels, obstacles) ni de caractériser l'incertitude. En outre, la pondération optimale varie selon les mécanismes physiques : une loi en d^{-2} est adaptée au bruit (propagation de type point source) mais surestime souvent la décroissance pour les polluants gazeux ou les aérosols, dont la diffusion est dominée par la turbulence et la convection.

Modèles géostatistiques

Le formalisme krigeage s'appuie sur la modélisation explicite de la covariance via le variogramme, fournissant l'estimateur linéaire non biaisé à variance minimale. Les analyses comparatives réalisées sur des réseaux de capteurs industriels, par exemple Xiao *et al.* [Xiao16] montrent que le krigeage *ordinaire* surpasse les autres variantes dès lors que le variogramme est rigoureusement calibré. Des conclusions similaires apparaissent dans [Koehler13], où l'influence des choix méthodologiques sur l'interprétation des cartes a été évaluée à partir des ensembles de données de [Peters06, Evans08, Peters12]. En effet, les performances du krigeage ordinaire appliqué sur des données de divers polluants, sont conditionnées par le choix et le paramétrage du modèle de variogramme, ainsi que la densité et la distribution spatiale des points de mesure ; lorsque ces exigences sont satisfaites, cette méthode demeure la plus apte à modéliser les variations spatiales pertinentes. Cependant ses limites résident dans l'hypothèse de stationnarité et ses mêmes paramètres. Un sous-échantillonnage ou erreurs systématiques de capteurs détériorent la précision et peuvent biaiser les cartes.

Interpolateurs régularisés par équation aux dérivées partielles

La spline biharmonique résout l'opérateur biharmonique ∇^4 et construit un champ lisse par superposition de fonctions de Green. Son absence de paramétrage de degré évite les oscillations de haute fréquence, avantage confirmé dans la cartographie massique de particules en menuiserie [Li18]. Néanmoins, l'extension à de grands jeux de données accroît la complexité matricielle (conditionnement défavorable) et complique l'intégration d'anisotropies fortes sans modification de l'opérateur différentiel.

Modèles neuronaux

La réalisation des cartographies spatio-temporelles avec les réseaux de neurones artificiels implique un leurs entraînements sur des quadruplets $(x, y, t, \text{concentration})$ et permet de capturer les non-linéarités en intégrant la dimension temporelle dans un cadre unifié. Dans [Li18], un réseau à deux couches cachées (30 neurones chacune avec une fonction d'activation tanh) reproduit les concentrations de PM avec un RMSE comparable aux variantes déterministes, mais demeure inférieur au krigeage sur la localisation des sources de pollution. Deux limitations majeures sont identifiées : l'opacité inhérente à la structure paramétrique des réseaux de neurones rend leur interprétation difficile ; leur forte capacité d'ajustement requiert des jeux de données denses et représentatifs : en l'absence de couverture suffisante, le modèle tend à sur-apprendre les variations locales bruitées, au détriment de sa généralisation.

Il en ressort de cette analyse plusieurs limites structurelles de ces méthodes. En effet, bien qu'elles aient permis des avancées notables pour la représentation de l'exposition environnementale, leur application dans des contextes opérationnels complexes demeure parfois limitée. Ces approches reposent généralement sur des hypothèses qui peuvent être mises en difficulté en présence d'obstacles physiques, de sources d'émission mobiles ou intermittentes, ou encore de variations rapides des conditions d'occupation et de ventilation. Par ailleurs, la prise en compte explicite de la cohérence temporelle dans la structuration des champs spatiaux reste un défi méthodologique, susceptible d'influencer l'interprétation des gradients de concentration.

Aussi, une problématique majeure en santé au travail reste la détermination de l'exposition individuelle des travailleurs aux polluants au sein de l'environnement de travail. Dans l'état de l'art, les données d'expositions individuelles aux polluants sont souvent confondues avec les données de concentrations issues de la cartographie spatio-temporelle. Cependant, ces dernières ne résolvent pas à elles seules cette problématique par manque d'informations sur les positions en temps réel des travailleurs au sein de l'environnement.

Les approches de modélisation spatio-temporelle intégrant des données provenant à la fois de capteurs fixes et de capteurs itinérants ont principalement été conçues dans l'optique d'améliorer la couverture spatiale du phénomène mesuré, en compensant la rareté des points fixes par des observations supplémentaires en mouvement. Ces travaux, tels que celui de [Ludwig17], n'ont pas été orientés vers une reconstitution de l'exposition individuelle, mais plutôt vers une amélioration de la résolution spatiale des champs de concentration. Si cette stratégie permet effectivement de mieux capturer la variabilité spatiale, elle reste limitée par une discontinuité temporelle inhérente aux mesures itinérantes, ce qui complique toute estimation fiable de l'exposition cumulée à l'échelle individuelle.

1.5 Hypothèses de travail et démarche proposée pour les recherches

1.5.1 Apports de l'état de l'art

Les constats mis en évidence par l'analyse de l'état de l'art ont permis d'identifier, dans le cadre de cette thèse, deux axes de recherche.

1. Développement d'un modèle automatique pour la cartographie spatio-temporelle des polluants avec une prise en compte conjointe de la cohérence spatiale et temporelle des données. Cette méthodologie devra être capable de s'adapter aux contraintes telles que la couverture spatiale, les changements de dynamique et la nature de pollution (gaz, particules), etc.
2. Développement d'une méthode de fusion de données issues de sources hétérogènes et de reconstruction de l'exposition individuelle des travailleurs à partir des cartes spatio-temporelles générées.

1.5.2 Hypothèse de travail

Le développement de ces deux axes repose sur un ensemble d'hypothèses pour orienter notre démarche scientifique.

— La conception du dispositif expérimental d'acquisition des données en identifiant les composantes à mettre en place (capteurs, configurations spatiales, protocoles de mesure), afin de garantir la représentativité et la qualité des données collectées.
— L'analyse des caractéristiques de la cartographie spatio-temporelle nécessite l'évaluation de l'influence de divers paramètres, tels que la méthode utilisée, la densité et la distribution spatiale des points de mesures, les propriétés de la source de pollution, le nombre et le positionnement des capteurs.
— L'intégration explicite de la variable temporelle dans la modélisation spatio-temporelle soulève la question des approches appropriées pour capter la dynamique temporelle des concentrations de polluants et générer des cartes spatio-temporelle pertinentes ? Développer une méthode de cartographie spatio-temporelle en intégrant ces approches.
— Le développement d'une méthode de fusion de données de localisation avec les données de la cartographie spatio-temporelle
— Peut-on évaluer l'exposition individuelle d'un travailleur avec cette fusion de données ?

1.5.3 Démarche envisagée

Pour atteindre ces objectifs, la démarche envisagée prévoit deux études consécutives illustrées dans la figure 1.7 :

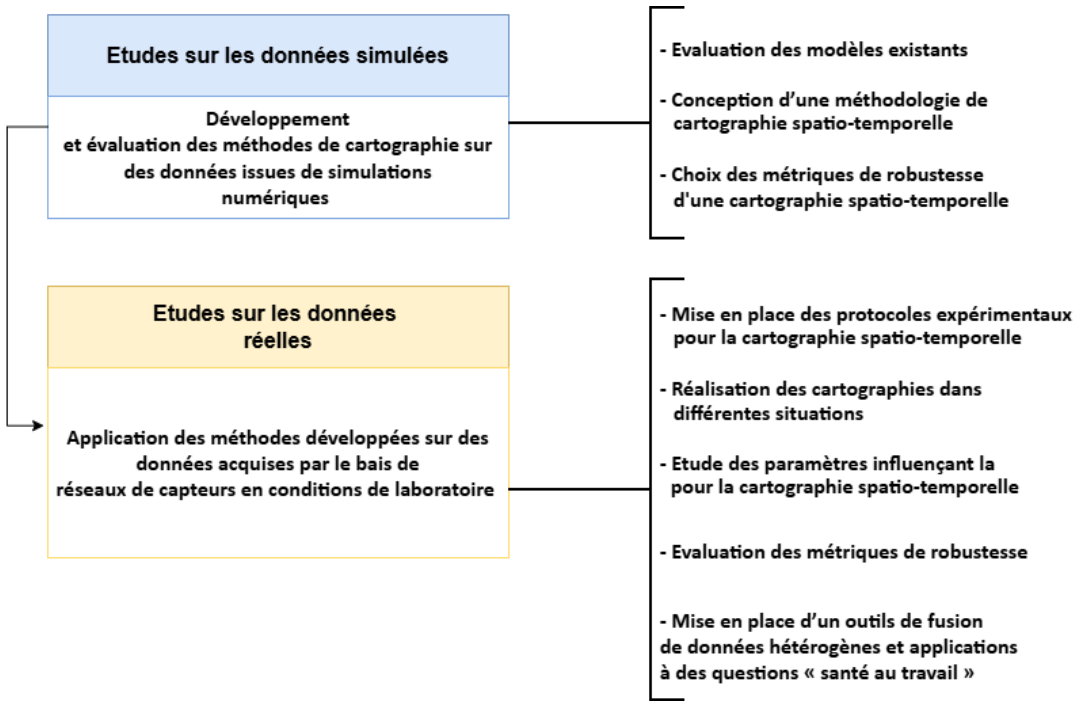


FIGURE 1.7 – Démarche envisagée pour le développement et l'évaluation des méthodes de cartographie spatio-temporelle dans le cadre des travaux de la thèse

— Une étude menée avec des données issues de simulations numériques et visant à développer et évaluer une méthode de cartographie spatio-temporelle. Cette étude intégrera la définition d'indicateurs numériques des cartes reproduites.

— Une étude menée avec des données réelles acquises par le bais de réseaux de capteurs en conditions de laboratoire. Cette étude vise à appliquer les méthodes développées ainsi que les indicateurs de fiabilité. Elle prévoit d'en étudier expérimentalement les principaux déterminants et aussi de concevoir une méthode de fusion des cartes avec des données hétérogènes. Ces développements seront mis à profit pour aborder des problématiques de santé au travail (estimation de l'exposition individuelle, intégration des cartes dans l'environnement cartographié, cartographie des dangers, etc.). Cette étude prévoit les actions suivantes :
  — Mise en place des dispositifs expérimentaux (espace expérimental, réseaux de capteurs et autres instruments de mesures).
  — Étude des paramètres influençant la cartographie spatio-temporelle et sa fiabilité.
  — Évaluation des indicateurs de fiabilité.
  — Fusion de la cartographie avec des données hétérogènes et application à des problématiques de santé aux travail.

NOTE : Les travaux de thèse prévoyaient une étude d'application dans une entreprise mais cette dernière n'a pas pu être engagée dans le temps imparti à la thèse.

Chapitre 2

Méthodes de cartographie spatio-temporelle

Sommaire

2.1 Introduction 23
2.2 Formulation générale du problème 24
2.3 Indicateurs de fiabilité et de robustesse de la cartographie 25
2.4 Génération de données spatio-temporelles synthétiques 26
2.5 Méthodes de cartographie spatiales 29
2.5.1 Pondération inverse à la distance 29
2.5.2 Complétion de matrice de rang faible avec contraintes de douceur spatiale 30
2.5.3 Krigeage 30
2.6 Analyse comparative des méthodes à partir de données simulées . . 33
2.7 Méthodes de cartographie spatio-temporelle 37
2.7.1 Prise en compte de la dimension temporelle 37
2.7.2 Régularisation tensorielle 38
2.7.3 Réseaux de neurones informés par la physique 43

2.1 Introduction

La cartographie spatio-temporelle constitue un cadre méthodologique destiné à représenter et à analyser la distribution ainsi que l'évolution d'un phénomène dynamique en fonction de l'espace et du temps, à partir de données quantitatives échantillonnées. Elle repose sur l'inférence de champs continus dérivés d'observations ponctuelles, par projection des valeurs mesurées sur une grille spatio-temporelle de résolution fine. Cette opération nécessite le recours à des méthodes d'estimation automatique capables d'extrapoler l'information issue d'un nombre restreint de

points de mesure, tout en tenant compte des corrélations spatiales et temporelles intrinsèques au phénomène étudié.

Dans ce chapitre, nous formalisons mathématiquement la problématique étudiée et nous évaluons, sur des données synthétiques, les méthodes de référence issues de l'état de l'art appliquées à la cartographie spatio-temporelle en santé au travail. Cette analyse met en évidence leurs limitations algorithmiques ainsi que les contraintes spécifiques inhérentes au contexte de la surveillance des polluants en milieu professionnel. Nous introduisons ensuite deux nouvelles approches, conçues pour pallier ces limitations et améliorer la représentativité de la dynamique des concentrations. Ces différentes méthodes sont également évaluées sur des jeux de données simulées, dans une perspective comparative, afin d'identifier la stratégie la plus adaptée aux exigences de notre problématique.

2.2 Formulation générale du problème

L'objectif est d'obtenir une cartographie spatio-temporelle en trois dimensions (deux dimensions spatiales x et y, et une dimension temporelle). Si Δt est la résolution temporelle souhaitée, il s'agit de générer toutes les Δt secondes, une carte des concentrations des polluants pour une région spatiale de forme rectangulaire, de dimensions $L \times l$.

En discrétisant cette région selon les deux dimensions spatiales, avec des pas respectifs Δx et Δy, on obtient une grille spatiale de dimensions $N_x \times N_y$, où : $N_x = \lfloor \frac{L}{\Delta x} \rfloor$, $N_y = \lfloor \frac{l}{\Delta y} \rfloor$ et où $\lfloor \cdot \rfloor$ désigne la partie entière d'un nombre. Nous disposons de K capteurs positionnés sur les points de la grille spatiale, mesurant des valeurs de concentration, de manière synchrone, toutes les Δt secondes (voir Fig. 2.1). A partir de ces mesures, l'objectif est d'estimer les valeurs de la concentration du polluant dans tous les autres points de la grille, au cours du temps. Pour atteindre cet objectif, il est nécessaire de recourir à des méthodes d'interpolation ou de modélisation qui exploitent les mesures fournies par les K capteurs. Ces méthodes permettent d'estimer les concentrations du polluant aux points non équipés de capteurs, en prenant en compte la corrélation spatiale et éventuellement, temporelle des données.

Si l'on considère la grille spatiale de dimensions $N_x \times N_y$ le problème peut être formulé comme l'estimation d'une séquence temporelle de matrices $\boldsymbol{X}_t \in \mathbb{R}_+^{N_x \times N_y}$, avec $t \in \{1, \ldots, N_t\}$, à partir d'un sous-ensemble de leurs entrées (spatiales) $\boldsymbol{X}_t(i,j)$, où $(i,j) \in \Omega \subseteq \{\{1, \ldots, N_x\} \times \{1, \ldots, N_y\}\}$. Dans ces expressions, N_t représente le nombre d'échantillons temporels pour un capteur, Ω regroupe l'ensemble des indices matriciels correspondant aux positions des capteurs (mesures) et $\mathbb{R}_+$ indique le fait que les valeurs des concentrations sont non-négatives. Les N_t matrices peuvent être regroupées sous la forme d'un tableau tridimensionnel (ou *tenseur*) de données $\boldsymbol{\mathcal{X}} \in \mathbb{R}_+^{N_x \times N_y \times N_t}$.

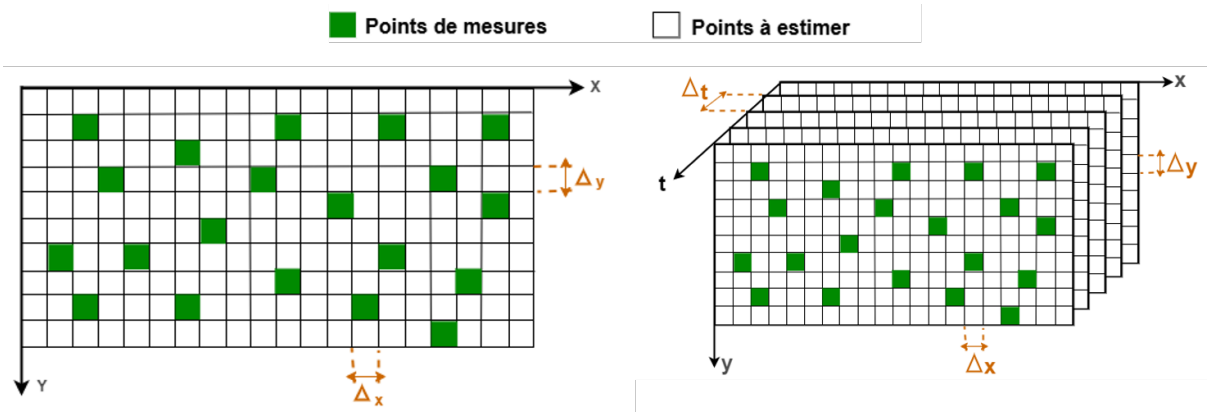


FIGURE 2.1 – Domaine D à cartographier avec Δ_x, Δ_y les résolutions spatiales et Δ_t la résolution temporelle, les points de mesures (en vert) et les points à estimer (en blanc).

2.3 Indicateurs de fiabilité et de robustesse de la cartographie

La validation des des cartes estimées nécessite l'introduction d'indicateurs statistiques quantifiant leur qualité, obtenus par comparaison entre les concentrations estimées $\hat{\boldsymbol{\mathcal{X}}}(i,j,t)$ par les méthodes et les mesures de référence $\boldsymbol{\mathcal{X}}(i,j,t)$ aux positions équipées de capteurs. Dans cette étude, les indicateurs retenus sont :

— **Erreur quadratique moyenne (RMSE)** : elle mesure l'écart quadratique moyen entre les valeurs estimées et les valeurs de référence. Elle s'exprime en ppm/v et est définie par :

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N_t}\sum_{(i,j)\in\Omega}\left(\hat{\boldsymbol{\mathcal{X}}}(i,j,t) - \boldsymbol{\mathcal{X}}(i,j,t)\right)^2} \tag{2.1}$$

— **Coefficient de corrélation ($\mathbf{R^2}$)** : il quantifie la proportion de la variance totale des mesures qui est expliquée par le modèle. Une valeur proche de 1 indique une bonne capacité de restitution de la dynamique spatiale / spatio-temporelle, tandis qu'une valeur proche de 0 traduit une mauvaise performance.

$$R^2 = 1 - \frac{\sum_{t=1}^{N_t}\sum_{(i,j)\in\Omega}\left(\boldsymbol{\mathcal{X}}(i,j,t) - \hat{\boldsymbol{\mathcal{X}}}(i,j,t)\right)^2}{\sum_{t=1}^{N_t}\sum_{(i,j)\in\Omega}\left(\boldsymbol{\mathcal{X}}(i,j,t) - \bar{\boldsymbol{\mathcal{X}}}\right)^2}, \tag{2.2}$$

où $\bar{\boldsymbol{\mathcal{X}}}$ désigne la moyenne des mesures de référence sur l'ensemble des positions et instants considérés.

2.4 Génération de données spatio-temporelles synthétiques

Afin d'évaluer l'efficacité des méthodes de référence dans le cadre de notre problématique, nous recourons à des simulations numériques fournissant une référence contrôlée. Le phénomène étudié correspond à l'évolution spatio-temporelle des concentrations de polluants gazeux, représentée par des cartes synthétiques intégrant une variabilité temporelle. Ces données sont générées à partir de deux sources modélisées par des gaussiennes bidimensionnelles non normalisées, dont les variances évoluent dans le temps, traduisant ainsi la diffusion progressive des gaz dans l'espace et le temps. Nous considérons par la suite que les dimensions spatiales et temporelles ont été discrétisées avec des pas Δx, Δy et Δt et nous nous intéresserons aux matrices obtenues : $\boldsymbol{X}_t(i,j)$, avec $i = 1,\ldots,N_x$, $j = 1,\ldots,N_y$ et $t = 1,\ldots,N_t$.

Ainsi, une source localisée aux coordonnées (i_0, j_0) sur la grille sera modélisée par une gaussienne 2D de moyenne $\boldsymbol{\mu}_0 = [i_0, j_0]^\mathsf{T}$ et de matrice de covariance $\boldsymbol{\Sigma}_t$. La valeur de la concentration de polluant en un point $\boldsymbol{d} = [i, j]^\mathsf{T}$ de la grille à l'instant t est

$$G_t(\boldsymbol{d}) = A \cdot \exp\left(-\frac{1}{2}(\boldsymbol{d}-\boldsymbol{\mu}_0)^\mathsf{T}\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{d}-\boldsymbol{\mu}_0)\right), \tag{2.3}$$

où A est l'amplitude maximale de la gaussienne, représentant l'intensité de la source de pollution. La matrice de covariance $\boldsymbol{\Sigma}_t^{-1}$ décrit la forme et l'étendue spatiale de la source et dépend de l'indice temporel t. Pour une source isotrope, on a $\boldsymbol{\Sigma}_t = \sigma_t^2\boldsymbol{I}$, où $\boldsymbol{I}$ est la matrice identité, et l'eq. (2.3) devient :

$$G_t(i,j) = A \cdot \exp\left(-\frac{(i-i_0)^2 + (j-j_0)^2}{2\sigma_t^2}\right), \tag{2.4}$$

où σ_t caractérise la dispersion spatiale de la source. Cette modélisation garantit que la concentration est maximale au centre de la source, en (i_0, j_0) et elle diminue de manière exponentielle en fonction de la distance. En augmentant la valeur de σ_t en fonction de l'indice temporel t on peut simuler un phénomène de dispersion spatio-temporelle de la concentration des polluants.

En effet, nous prenons en paramètres plusieurs facteurs déterminants que sont le choix de la taille des matrices représentant les cartes de pollution, les pas spatiaux et temporel, Δx, Δy et Δt et le nombre de sources de pollution K. De plus, pour chaque source, nous pouvons choisir sa position, son amplitude et sa portée (taille et forme de dispersion autour du centre de la source), qui est donnée par la matrice de covariance $\boldsymbol{\Sigma}$.

La Figure 2.2 illustre un exemple de données simulées, avec deux sources pour 12 instants d'échantillonnage.

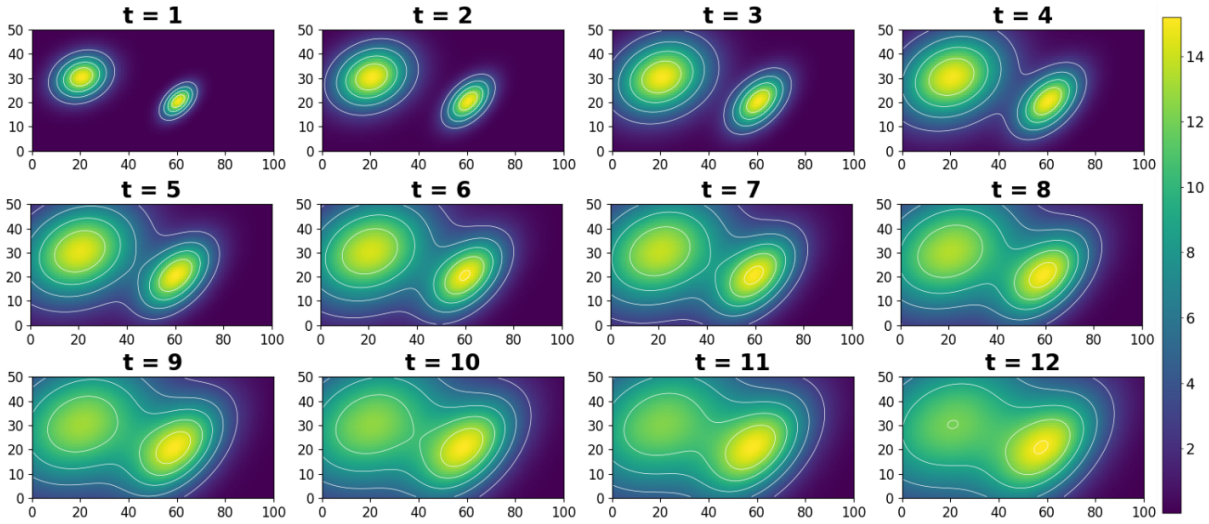


FIGURE 2.2 – Exemple de cartes de pollution simulées représentant l'évolution temporelle (t = 1 à 12) de deux sources gaussiennes

Afin d'évaluer la performance des méthodes de cartographie sur les cartes simulées, nous avons fait varier à la fois la position et le nombre des points de mesure. Les points ont été disposés selon des grilles régulières ou irrégulières, de manière à explorer différents scénarios d'échantillonnage spatial. Les figures 2.3, 2.4, 2.5 et 2.6 présentent un ensemble de quinze grilles de mesure, comportant respectivement 10, 15, 20 et 25 points de mesure.

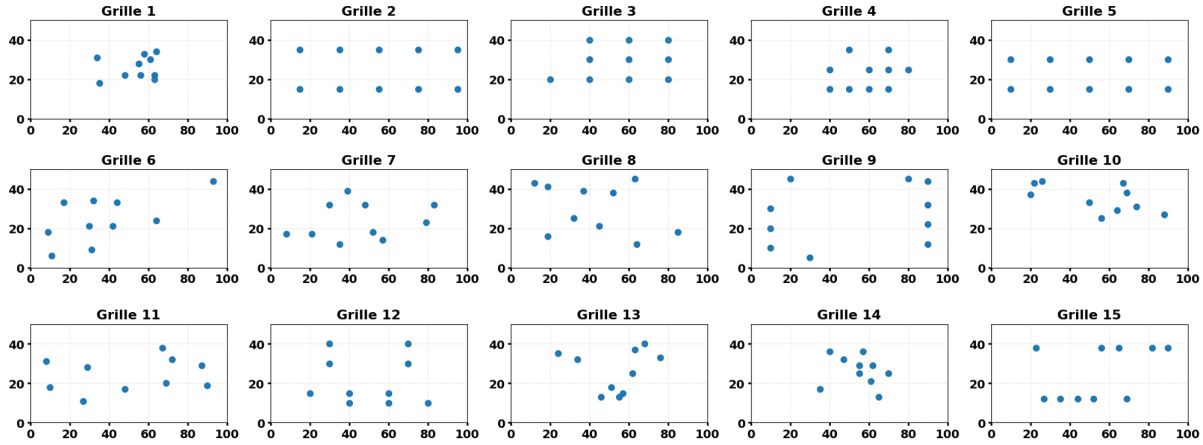


FIGURE 2.3 – Série de 15 configurations de grilles d'échantillonnage (régulières, irrégulières, denses et éparses) de 10 points

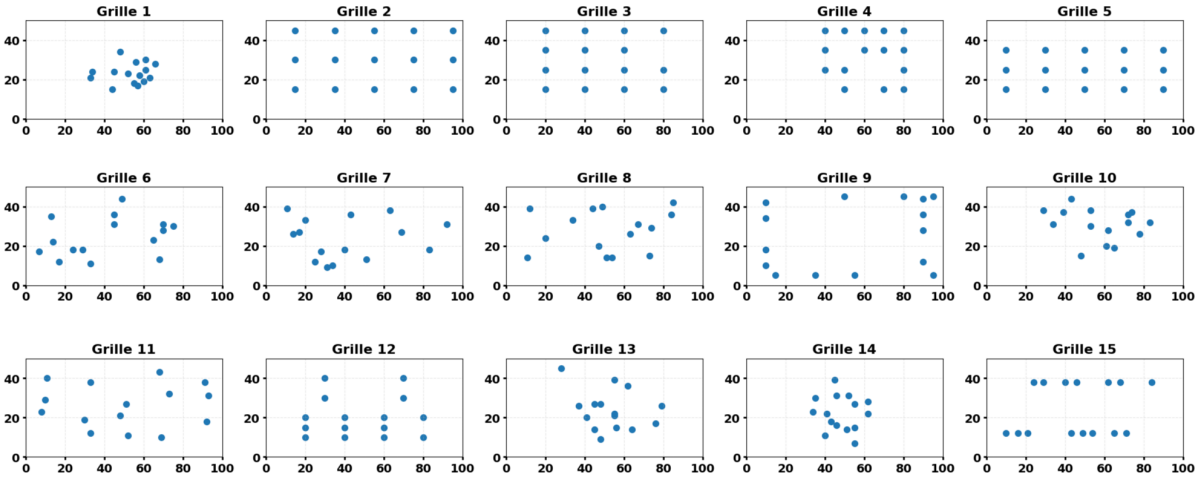


FIGURE 2.4 – Série de 15 configurations de grilles d'échantillonnage (régulières, irrégulières, denses et éparses) de 15 points

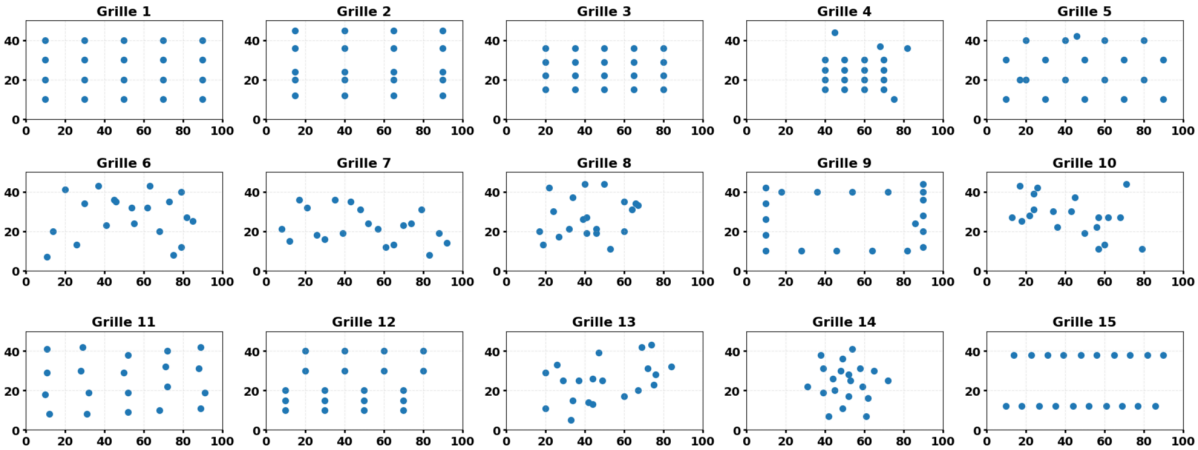


FIGURE 2.5 – Série de 15 configurations de grilles d'échantillonnage (régulières, irrégulières, denses et éparses) de 20 points

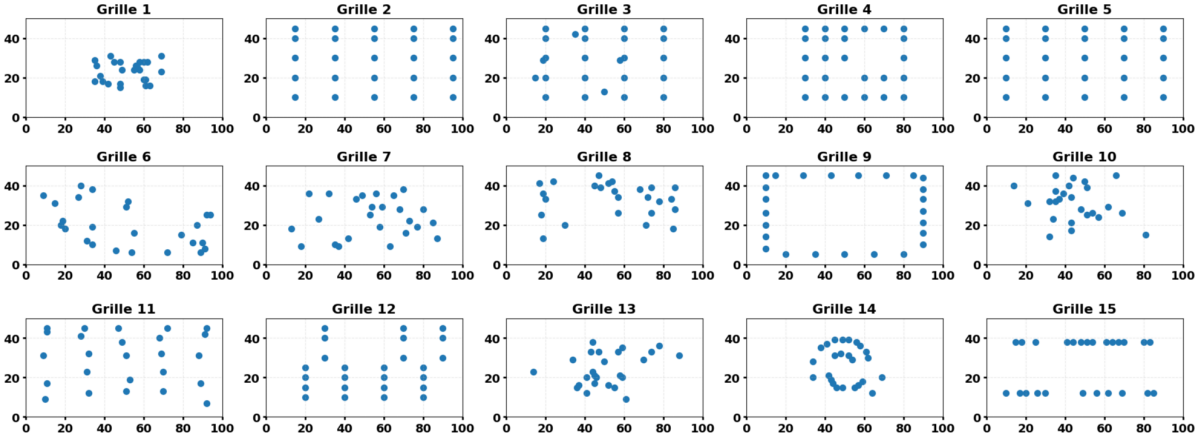


FIGURE 2.6 – Série de 15 configurations de grilles d'échantillonnage (régulières, irrégulières, denses et éparses) de 25 points

2.5 Méthodes de cartographie spatiales

Cette section présente les principales méthodes utilisées pour la cartographie spatiale des concentrations de polluants. Elle décrit les approches classiques d'interpolation et de modélisation spatiale, en exposant leurs fondements théoriques ainsi que les caractéristiques algorithmiques relatives à leurs performances et à leurs contraintes dans le contexte de l'analyse spatio-temporelle des concentrations de polluants. Les méthodes présentées dans cette section permettent d'estimer, pour chaque instant de temps t, des cartes spatiales de manière indépendante. Par conséquent, afin d'alléger la notation, l'indice t sera omis dans la suite.

2.5.1 Pondération inverse à la distance

La pondération inverse à la distance (Inverse Distance Weighting, IDW), introduite par Shepard [Shepard68], est une méthode déterministe d'interpolation spatiale. Elle estime la concentration en un point de la grille spatiale $(i, j) \in \{\{1, \ldots, N_x\} \times \{1, \ldots, N_y\}\}$ à partir des mesures disponibles aux positions équipées de capteurs. L'idée est de calculer une moyenne pondérée des observations, où les poids décroissent avec la distance spatiale :

$$\hat{\boldsymbol{X}}(i,j) = \frac{\displaystyle\sum_{m=1}^{M} \frac{1}{\| [i,j] - [i_m, j_m] \|_2^p} \boldsymbol{X}(i_m, j_m)}{\displaystyle\sum_{m=1}^{M} \frac{1}{\| [i,j] - [i_m, j_m] \|_2^p}}, \tag{2.5}$$

où :

M est le nombre de capteurs pris en compte,

$\boldsymbol{X}(i_m, j_m) = \boldsymbol{X}_t(i_m, j_m)$ est la concentration mesurée par le capteur en position (i_m, j_m) au temps t,

$\| [i,j] - [i_m, j_m] \|_2$ est la distance euclidienne entre la position (i, j) de la valeur à estimer et le capteur situé à (i_m, j_m),

$p > 0$ est le paramètre de puissance contrôlant la décroissance de l'influence avec la distance.

L'interprétation est intuitive : plus un capteur (i_m, j_m) est proche de la position (i, j), plus son influence est grande dans l'estimation. Le paramètre p module ce comportement :

— $p = 1$: décroissance linéaire, pondération relativement douce,

— $p = 2$: décroissance quadratique, forte influence des points proches,

— $p > 2$: les capteurs éloignés ont une contribution quasi négligeable.

Le choix de p est déterminant et généralement empirique, souvent fondé sur validation croisée [Li14]. Typiquement, $p \in [1, 3]$. Un p trop faible tend à lisser les gradients et à intégrer les points éloignés, tandis qu'un p trop élevé peut créer des discontinuités artificielles dans la cartographie.

2.5.2 Complétion de matrice de rang faible avec contraintes de douceur spatiale

Cette méthode vise à reconstruire une matrice spatiale à partir des observations partielles, en exploitant une hypothèse de structure de rang faible, pertinente pour modéliser la continuité spatiale des concentrations [Candes12]. En notant $\boldsymbol{X}(i,j)$ les concentrations mesurées et $\hat{\boldsymbol{X}}(i,j)$ leur reconstruction, le problème peut être formulé comme :

$$\min_{\hat{\boldsymbol{X}}} \ \text{rang}(\hat{\boldsymbol{X}}) \quad \text{s.c.} \quad \hat{\boldsymbol{X}}(i,j) = \boldsymbol{X}(i,j),\ \forall (i,j) \in \Omega, \tag{2.6}$$

où Ω désigne l'ensemble des positions équipées de capteurs. Ce problème étant NP-difficile, il est couramment relaxé en remplaçant la contrainte de rang par la norme nucléaire :

$$\|\hat{\boldsymbol{X}}\|_* = \sum_r \lambda_r,$$

où λ_r sont les valeurs singulières de $\hat{\boldsymbol{X}}$ [Candes12]. Plusieurs approches ont été proposées pour résoudre ce problème, incluant le seuillage des valeurs singulières (SVT) [Cai10], les modèles gaussiens hiérarchiques [Léger11] ou des variantes localisées [Shabat12].

Afin de renforcer la cohérence spatiale de la reconstruction, Hu et al. [Hu13] introduisent la complétion de matrice de rang faible régularisée par douceur spatiale (*Smoothing Low-Rank Matrix Completion*, SLRMC), formulée comme :

$$\min_{\hat{\boldsymbol{X}}} \ \|\hat{\boldsymbol{X}}\|_* + \lambda S(\hat{\boldsymbol{X}}) \quad \text{s.c.} \quad \hat{\boldsymbol{X}}(i,j) = \boldsymbol{X}(i,j),\ \forall (i,j) \in \Omega, \tag{2.7}$$

avec

$$S(\hat{\boldsymbol{X}}) = \|\mathcal{D}_x(\hat{\boldsymbol{X}})\|_F^2 + \|\mathcal{D}_y(\hat{\boldsymbol{X}})\|_F^2, \tag{2.8}$$

où $\mathcal{D}_x$ et $\mathcal{D}_y$ désignent les opérateurs de dérivation spatiale selon les directions x et y, et $\lambda > 0$ un hyperparamètre contrôlant l'importance de la contrainte de douceur.

2.5.3 Krigeage

Le krigeage est un estimateur non biaisé à variance minimale, largement employé en géostatistique pour la cartographie spatiale [Wackernagel03]. Il permet d'estimer les valeurs non échantillonnées $\hat{\boldsymbol{X}}(i,j)$ comme une combinaison linéaire des M valeurs mesurées $\boldsymbol{X}(i_m, j_m)$ en des positions $(i_m, j_m) \in \Omega$. L'objectif de cette méthode est donc de déterminer les coefficients $(\lambda_m)_{m=1:M}$ tels que :

$$\hat{\boldsymbol{X}}(i,j) = \sum_{m=1}^{M} \lambda_m \, \boldsymbol{X}(i_m, j_m). \tag{2.9}$$

Dans le cadre du krigeage, il est courant de formuler l'hypothèse que les valeurs des concentrations résultent d'un processus aléatoire stationnaire d'ordre 2. Cette hypothèse implique que la moyenne des variables reste constante dans l'espace et que la covariance entre deux points dépend uniquement de la distance qui les sépare, et non de leur position absolue.

L'absence de biais se traduit par le fait que l'espérance de la différence entre la valeur réelle $\boldsymbol{X}(i,j)$ et l'estimation $\hat{\boldsymbol{X}}(i,j)$ est nulle. L'hypothèse de stationnarité d'ordre 2 implique une moyenne spatiale constante et égale à μ. Ainsi, si $\hat{\boldsymbol{X}}(i,j)$ désigne la valeur estimée au point d'intérêt et $\boldsymbol{X}(i,j)$ la valeur réelle correspondante, on peut formuler la condition de non-biais de la manière suivante :

$$\begin{aligned}
\mathbb{E}[\hat{\boldsymbol{X}}(i,j) - \boldsymbol{X}(i,j)] = 0 &\iff \mathbb{E}\left[\sum_{m=1}^{M} \lambda_m \boldsymbol{X}(i_m, j_m)\right] - \mathbb{E}[\boldsymbol{X}(i,j)] = 0 \\
&\iff \sum_{m=1}^{M} \lambda_m \mathbb{E}[\boldsymbol{X}(i_m, j_m)] = \mathbb{E}[\boldsymbol{X}(i,j)] \\
&\iff \mu \sum_{m=1}^{M} \lambda_m = \mu \iff \sum_{m=1}^{M} \lambda_m = 1
\end{aligned} \tag{2.10}$$

Pour que l'estimateur soit non biaisé, il faut donc que : $\sum_{m=1}^{M} \lambda_m = 1$.

Quant à la minimisation de la variance de l'estimateur, celle-ci repose sur la recherche des coefficients λ_m qui réduisent au maximum l'incertitude de l'estimation. En utilisant les notations introduites précédemment, la variance de l'erreur d'estimation $\hat{\boldsymbol{X}}(i,j) - \boldsymbol{X}(i,j)$, peut s'écrire sous la forme suivante :

$$\begin{aligned}
\mathbb{E}[(\hat{\boldsymbol{X}}(i,j) - \boldsymbol{X}(i,j))^2] &= \mathbb{E}\left[\left(\sum_{m=1}^{M} \lambda_m \boldsymbol{X}(i_m, j_m) - \boldsymbol{X}(i,j)\right)^2\right] \\
&= \sum_{m=1}^{M} \lambda_m^2 \, \mathbb{E}[\boldsymbol{X}(i_m, j_m)^2] \\
&\quad + 2 \sum_{m=1}^{M} \sum_{m'>m} \lambda_m \lambda_{m'} \operatorname{Cov}(\boldsymbol{X}(i_m, j_m), \boldsymbol{X}(i_{m'}, j_{m'})) \\
&\quad - 2 \sum_{m=1}^{M} \lambda_m \operatorname{Cov}(\boldsymbol{X}(i_m, j_m), \boldsymbol{X}(i,j)) + \mathbb{E}[\boldsymbol{X}(i,j)^2].
\end{aligned} \tag{2.11}$$

Dans l'équation (2.11), $\mathbb{E}[\boldsymbol{X}(i,j)^2]$ représente la variance intrinsèque de la concentration, elle reste constante par rapport aux coefficients λ_m mais inconnue. Le terme $\sum_{m=1}^{M} \sum_{m'=1}^{M} \lambda_m \lambda_{m'}$ $\operatorname{Cov}(\boldsymbol{X}(i_m, j_m), \boldsymbol{X}(i_{m'}, j_{m'}))$ représente la mesure de la redondance des différents points de mesure et le dernier terme, quant à lui, représente la covariance entre le point dont on cherche la valeur et les points de mesure. Le problème que l'on cherche à résoudre est donc le suivant :

$$\min \sum_{m=1}^{M} \sum_{m'=1}^{M} \lambda_m \lambda_{m'} \mathrm{Cov}(\boldsymbol{X}(i_m, j_m), \boldsymbol{X}(i_{m'}, j_{m'})) - 2 \sum_{m=1}^{M} \lambda_m \mathrm{Cov}(\boldsymbol{X}(i_m, j_m), \boldsymbol{X}(i, j)) + \mathbb{E}[\boldsymbol{X}(i, j)^2], \tag{2.12}$$

sous contrainte : $\sum_{m=1}^{M} \lambda_m = 1$.

Posons $\boldsymbol{d}$ le vecteur contenant les coefficients de krigeage : $\boldsymbol{d} = (\lambda_1, \ldots, \lambda_M)^\mathsf{T}$. Soit A la matrice de covariance des mesures et b le vecteur contenant les covariances entre les points de mesure et le point à estimer :

$$\begin{aligned} \boldsymbol{A}(m, m') &= \big(\mathrm{Cov}(\boldsymbol{X}(i_m, j_m), \boldsymbol{X}(i_{m'}, j_{m'}))\big), \\ \boldsymbol{b} &= \big(\mathrm{Cov}(\boldsymbol{X}(i_1, j_1), \boldsymbol{X}(i, j)), \ldots, \mathrm{Cov}(\boldsymbol{X}(i_M, j_M), \boldsymbol{X}(i, j))\big)^\mathsf{T}. \end{aligned} \tag{2.13}$$

Si l'on note $\tilde{\boldsymbol{A}} = \begin{pmatrix} \boldsymbol{A} & \mathbf{1}_M \\ \mathbf{1}_M^\mathsf{T} & 0 \end{pmatrix}$, $\tilde{\boldsymbol{b}} = \begin{pmatrix} \boldsymbol{b} \\ 1 \end{pmatrix}$, alors, en notant $\tilde{\boldsymbol{A}}^\dagger$ la pseudo-inverse au sens de Moore-Penrose de $\tilde{\boldsymbol{A}}$, la solution de 2.12 est donnée par l'équation suivante :

$$\boldsymbol{d} = \tilde{\boldsymbol{A}}^\dagger \tilde{\boldsymbol{b}}. \tag{2.14}$$

Les valeurs dans $\tilde{\boldsymbol{A}}$ dépendent uniquement des covariances des données connues. En revanche, les valeurs de $\boldsymbol{b}$ sont liées à la covariance avec la valeur recherchée, ce qui nécessite leur estimation. Pour y parvenir, nous utiliserons le *variogramme*, un outil fondamental pour caractériser la dépendance spatiale d'un processus aléatoire. Il est défini comme la demi-variance moyenne des différences entre les valeurs de concentration mesurées en deux positions (i_1, j_1) et (i_2, j_2). Sous l'hypothèse de stationnarité d'ordre 2, le variogramme ne dépend que de la distance entre les deux points, $h = \| [i_1, j_1] - [i_2, j_2] \|_2$, et non de leur position absolue :

$$\gamma(h) = \frac{1}{2} \mathbb{E}\Big[\big(\boldsymbol{X}(i_1, j_1) - \boldsymbol{X}(i_2, j_2)\big)^2\Big]. \tag{2.15}$$

La covariance entre deux positions peut alors être exprimée en fonction du variogramme :

$$\mathrm{Cov}(\boldsymbol{X}(i_1, j_1), \boldsymbol{X}(i_2, j_2)) = \sigma^2 - \gamma(h), \tag{2.16}$$

où σ^2 désigne la variance du processus aléatoire stationnaire $\boldsymbol{X}(i, j)$. Ainsi, pour chaque point de mesure (i_m, j_m), la composante b_m du vecteur $\boldsymbol{b}$ s'écrit :

$$b_m = \mathrm{Cov}(\boldsymbol{X}(i_m, j_m), \boldsymbol{X}(i, j)) = \sigma^2 - \gamma(h_m), \tag{2.17}$$

avec $h_m = \| [i_m, j_m] - [i, j] \|_2$. En pratique, le variogramme $\gamma(h)$ est estimé empiriquement à partir des données disponibles sous la forme du variogramme expérimental :

$$\gamma_{\text{exp}}(h) = \frac{1}{2N(h)} \sum_{(m,m') \in P(h)} \Big(\boldsymbol{X}(i_m, j_m) - \boldsymbol{X}(i_{m'}, j_{m'}) \Big)^2, \tag{2.18}$$

où $P(h)$ désigne l'ensemble des paires de capteurs séparés d'une distance comprise dans un intervalle autour de h, et $N(h)$ le nombre de paires considérées. Un modèle théorique continu $\gamma(h)$ est ensuite ajusté à $\gamma_{\text{exp}}(h)$ pour garantir une estimation lisse et positive des covariances. Les modèles les plus couramment utilisés sont :

— le modèle *sphérique*, adapté aux structures présentant une portée finie a :

$$\gamma(h) = \begin{cases} C_0 + C \left[\dfrac{3h}{2a} - \dfrac{h^3}{2a^3} \right], & \text{si } h \leq a, \\ C_0 + C, & \text{si } h > a, \end{cases} \tag{2.19}$$

— le modèle *exponentiel* :

$$\gamma(h) = C_0 + C \left(1 - e^{-h/a} \right), \tag{2.20}$$

— le modèle *gaussien* :

$$\gamma(h) = C_0 + C \left(1 - e^{-(h/a)^2} \right), \tag{2.21}$$

où C_0 correspond à *l'effet pépite*, représentant la discontinuité à l'origine liée à l'erreur de mesure, C à *la contribution structurale*, c'est-à-dire la part de variance expliquée par la structure spatiale du phénomène, et a à *la portée*, distance au-delà de laquelle les observations ne sont plus spatialement corrélées. L'ajustement de ces modèles permet de décrire la structure de corrélation spatiale des données et d'évaluer les covariances nécessaires au calcul des coefficients de krigeage. Ainsi, le variogramme assure la cohérence entre la géométrie des mesures, la structure spatiale du phénomène et la précision de l'estimation spatiale obtenue par le krigeage. Parmi ces modèles de variogramme (sphérique, exponentiel, gaussien), nous retenons le modèle gaussien pour notre contexte d'étude, car il correspond naturellement à la diffusion continue des concentrations de polluants dans l'air. Dans l'état de l'art, ce modèle a souvent montré des performances robustes pour la modélisation de la pollution atmosphérique [Rivera-Gonzalez15], et il s'inscrit de façon cohérente dans l'approche du krigeage fondée sur les processus gaussiens [Gelfand16].

2.6 Analyse comparative des méthodes à partir de données simulées

Nous évaluons la reconstruction spatio–temporelle d'un champ synthétique (Voir Figure 2.2) à partir de $T = 12$ cartes et de 15 grilles d'échantillonnage spatiales. Pour chaque grille,

nous considérons un nombre de $n = 10, 15, 20$ et 25 points de mesure (Voir Figures 2.3-2.6). Les trois méthodes présentées dans la section précédente : IDW (puissance $p = 2$), krigeage ordinaire (variogramme gaussien, retenu après comparaison des modèles sphérique, gaussien et exponentiel) et SLRMC sont évaluées.

Les performances sont rapportées via la *RMSE* agrégée spatio-temporellement et représentées, pour chaque valeur de n, en fonction des 15 grilles (Figs 2.7–2.10).

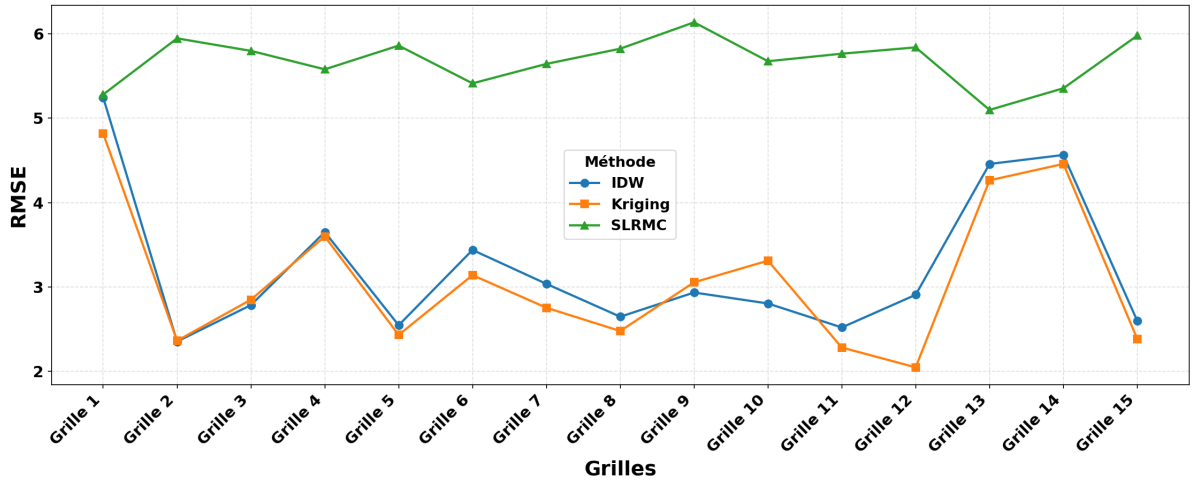


FIGURE 2.7 – Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 10 points de la figure 2.3, obtenue par les méthodes IDW, krigeage et SLRMC

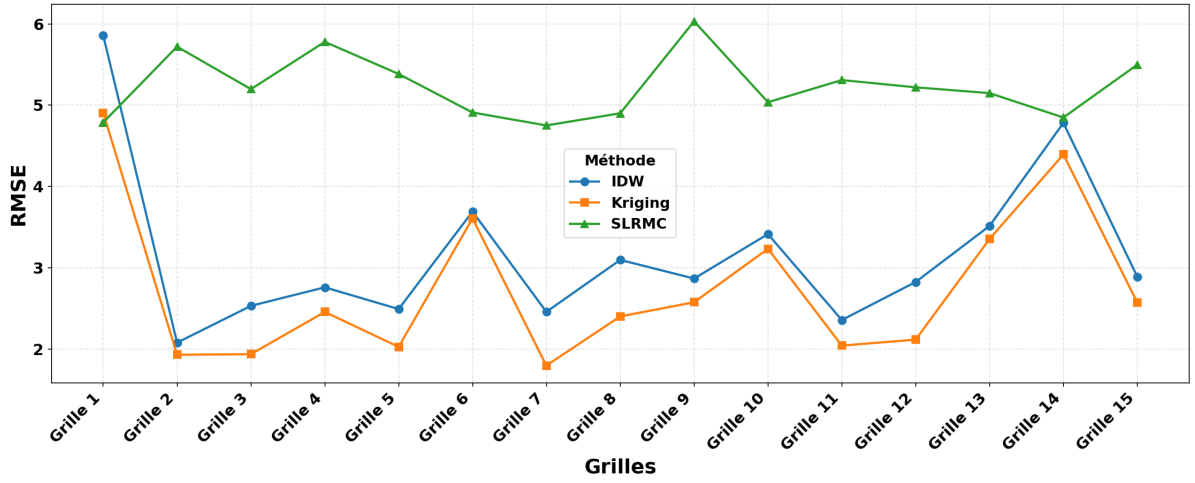


FIGURE 2.8 – Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 15 points de la figure 2.4, obtenue par les méthodes IDW, krigeage et SLRMC

On observe que, sur l'ensemble des grilles et pour les différents nombre de points de mesure n, le krigeage présente le plus souvent les RMSE les plus faibles (Voir figures 2.7-2.11). Il tire parti d'un modèle de dépendance spatiale explicite via un variogramme gaussien ajusté aux valeurs des covariances empiriques calculées et des propriétés de l'estimateur (non biaisé, à variance

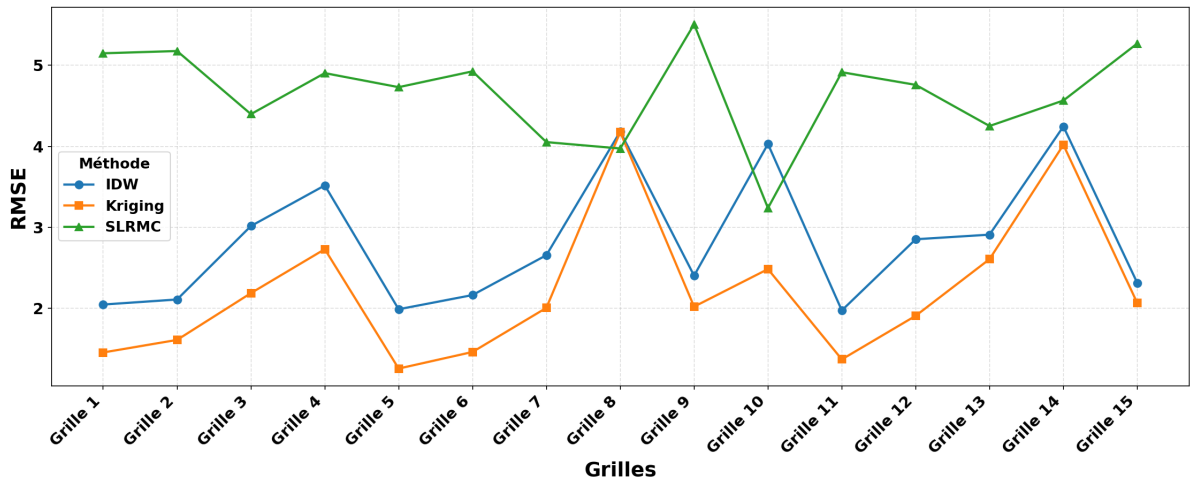


FIGURE 2.9 – Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 20 points de la figure 2.5, obtenue par les méthodes IDW, krigeage et SLRMC

minimale). L'IDW constitue une base efficace et peu coûteuse, avec des erreurs proches de celles du krigeage lorsque la géométrie d'échantillonnage couvre correctement les gradients locaux et/ou l'ajustement du variogramme est mauvais (faible n ou configuration très irrégulière). À l'inverse, la SLRMC présente des RMSE sensiblement plus élevées : la combinaison « rang faible + régulation » induit un biais de lissage qui atténue les extrêmes et les structures localisées, notamment lorsque le champ présente des pics anisotropes (Voir figure 2.11) .

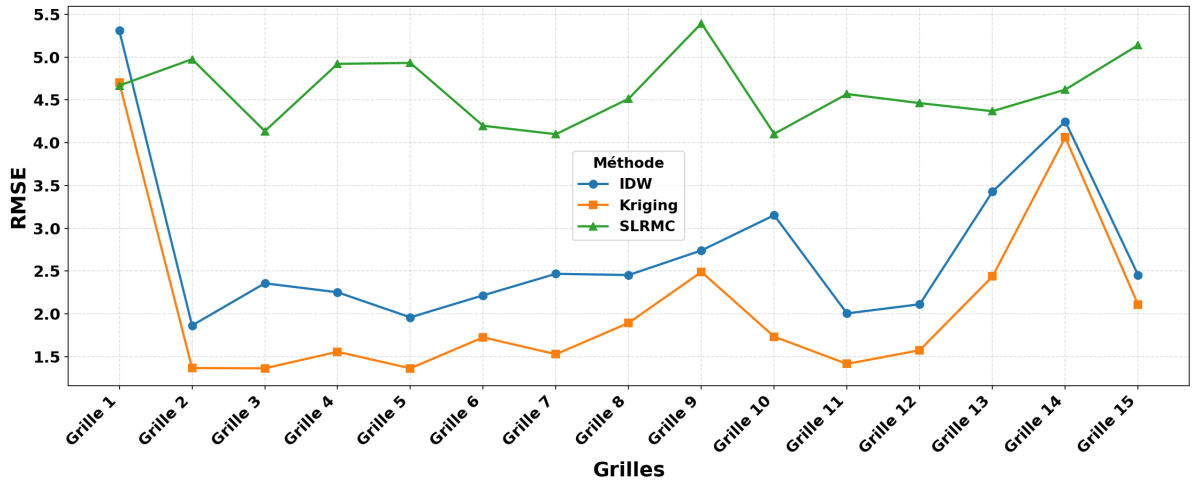


FIGURE 2.10 – Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 25 points de la figure 2.6, obtenue par les méthodes IDW, krigeage et SLRMC

Les Figures 2.7 à 2.10 mettent en évidence une décroissance du RMSE lorsque le nombre de points d'échantillonnage n augmente. Cette tendance est particulièrement marquée pour le krigeage et reflète une amélioration de la précision de reconstruction induite par l'accroissement de la densité de capteurs, qui optimise la capacité des modèles à capturer les gradients spatiaux.

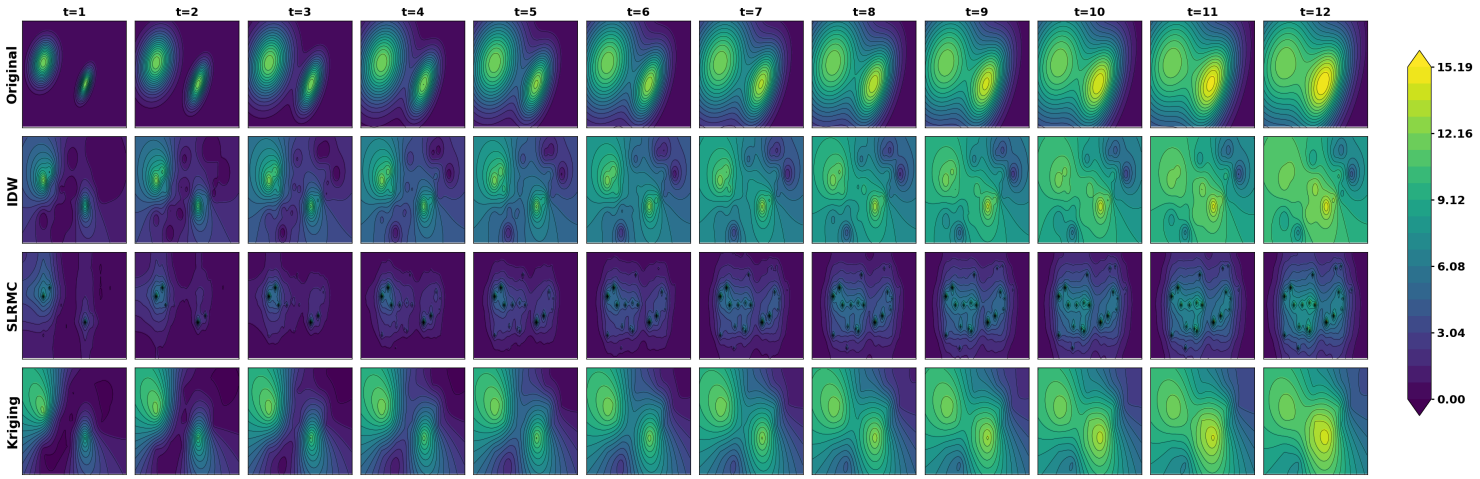


FIGURE 2.11 – Cartes reconstruites à partir de 20 points des grilles d'échantillonnage présentées en figure 2.5 avec les méthodes IDW, krigeage et SLMRC.

Plus précisément, la diminution la plus significative du RMSE est observée entre $n = 10$ et $n = 20$, période au cours de laquelle la couverture spatiale devient suffisante pour représenter les principales structures du champ synthétique (voir Figures 2.7 et 2.9). Au-delà de $n = 20$, la diminution de l'erreur se stabilise, indiquant une saturation liée à la redondance des points dans les zones déjà bien couvertes. La méthode SLRMC montre une amélioration plus limitée avec l'augmentation du nombre de points, la régularisation de douceur imposant un lissage quadratique qui atténue les structures fines et réduit la sensibilité du modèle à la densification du réseau, en particulier pour les champs à gradients anisotropes ou à pics localisés (voir Figure 2.11).

L'analyse des performances selon le positionnement des grilles d'échantillonnage révèle des comportements contrastés. Les grilles dont les points sont concentrés sur les bords (par exemple la Grille 1 des Figures 2.7 et 2.8) induisent une dégradation notable des performances pour l'ensemble des méthodes.

Dans ce cas, le krigeage est pénalisé par la nécessité d'extrapoler en dehors de la zone effectivement couverte, ce qui accroît les incertitudes associées aux estimations. L'IDW, plus robuste vis-à-vis des lacunes spatiales, conserve des RMSE stables mais génère localement des discontinuités marquées dans les zones faiblement échantillonnées. Les configurations quasi régulières ou condensées (par exemple les Grilles 8 et 12 des Figures 2.9 et 2.10) favorisent le krigeage, qui bénéficie alors d'un ajustement plus précis du variogramme gaussien. Dans ces situations, l'IDW présente également une bonne cohérence avec la structure du champ, mais au prix d'une continuité spatiale plus faible entre zones de densité variable. La SLRMC se montre peu sensible aux variations de l'échantillonnage grâce à sa régularisation spatiale qui impose une cohérence globale entre les points. Cependant, cette contrainte induit un lissage excessif, surtout dans les

zones présentant des gradients marqués ou des pics de concentration localisés (voir Figure 2.11).

De manière générale, le krigeage surpasse l'IDW dès lors que le variogramme est bien identifié, grâce à son exploitation explicite de la covariance spatiale. Cependant, lorsque le variogramme est mal ajusté (cas de faible n ou d'échantillonnage irrégulier, comme illustré par les Figures 2.7 et 2.8), les deux méthodes tendent à produire des erreurs comparables, soulignant la sensibilité du krigeage à la qualité de la modélisation de la dépendance spatiale.

Quant au SLMRC, en l'absence de terme de fidélité quadratique (contrainte stricte) et de toute forme de régularisation temporelle, le modèle favorise la reconstruction de structures globales lisses et de faible rang, au détriment de la restitution des variations locales. Ce biais structurel explique les erreurs quadratiques moyennes (RMSE) plus élevées observées pour des champs présentant une forte hétérogénéité spatiale (Figures 2.11 et 2.10). L'absence d'intégration de la dimension temporelle aggrave ce phénomène, car les dynamiques locales et transitoires, potentiellement informatives, ne sont pas exploitées.

En conclusion, dans le cadre de cette analyse et comme le confirme l'état de l'art, le krigeage ordinaire s'impose comme la méthode de référence dès qu'il y a une bonne couverture du nombre de points disponibles. L'IDW demeure une alternative simple offrant des performances compétitives dans les cas de forte incertitude sur le variogramme. En revanche, la SLRMC apparaît mal adaptée aux champs caractérisés par des hétérogénéités locales prononcées et par des évolutions temporelles significatives. Son efficacité pourrait être améliorée par l'introduction de régularisations spatio–temporelles et/ou de pénalités favorisant des discontinuités contrôlées, ouvrant ainsi la voie à une meilleure prise en compte conjointe des dimensions spatiales et temporelles dans la cartographie des concentrations.

2.7 Méthodes de cartographie spatio-temporelle

2.7.1 Prise en compte de la dimension temporelle

Les méthodes présentées dans la section précédente sont appliquées de manière indépendante à chaque instant temporel t. Cette approche ne garantit pas la cohérence temporelle des cartes de pollution, rendant difficile l'imposition d'une continuité temporelle dans la dynamique des polluants. Pour prendre en compte de manière explicite la continuité temporelle, plusieurs solutions sont envisageables.

Une solution serait d'intégrer la dimension temporelle dans la méthode de krigeage. Le krigeage spatio-temporel est une extension naturelle du krigeage classique, qui repose sur la modélisation conjointe de la covariance spatiale et temporelle au moyen de fonctions de variogrammes spatio-temporels ([De Iaco01], [Gneiting02]). Cette approche permet d'estimer les concentrations en tout point de l'espace-temps en tenant compte à la fois de la dépendance

spatiale entre capteurs et de la persistance temporelle des phénomènes de diffusion. Toutefois, plusieurs travaux ont mis en évidence des limites majeures : la difficulté de spécifier et d'ajuster un modèle de covariance pertinent dans des environnements hétérogènes tels que les espaces de travail [Stein05], la forte sensibilité aux données bruitées et aux réseaux peu denses [Cameletti13] (la « malédiction de la dimensionnalité »), ainsi que le coût computationnel qui devient rapidement prohibitif lorsque le nombre de points spatio-temporels croît, en raison de l'inversion des matrices de covariance de très grande taille. Cela restreint fortement l'usage du krigeage spatio-temporel [Sun12] dans les applications pratiques.

La contrainte computationnelle s'avère déterminante dans notre contexte. En effet, nos données présentent une résolution spatiale et temporelle fine, avec des séries temporelles à pas de temps d'une minute, afin de permettre une détection rapide des sources de pollution en milieu professionnel. Cette granularité, indispensable pour les applications en santé au travail, rend le recours au krigeage spatio-temporel difficilement envisageable pour une résolution spatio-temporelle fine. Pour les cas d'utilisation en santé au travail la résolution est très souvent aggrandie avec très peu de données. L'exemple rapporté par [Ruiter24] met également en lumière ces contraintes. Dans leur étude, bien que les capteurs enregistraient une mesure toutes les dix secondes, les données ont été agrégées pour l'interpolation spatio-temporelle, avec des résolutions temporelles de 16 minutes en laboratoire et de 45 minutes en entreprise, ce qui réduit drastiquement le nombre d'occurrences temporelles pour rendre les calculs réalisables, au prix d'une perte d'information locale. Nous avons par conséquent choisi d'écarter cette solution.

La méthode que nous avons décidé d'explorer est d'intégrer explicitement la variable temporelle dans le processus de reconstruction, en formulant le problème sous la forme d'une optimisation contrainte par des considérations de régularité spatio-temporelle. Elle sera développée dans la Section 2.7.2. Dans la continuité de cette démarche, une deuxième méthode vise à aller au-delà de la simple régularisation spatio-temporelle en intégrant explicitement dans le processus de reconstruction les lois physiques qui régissent la diffusion des polluants. Pour ce faire, nous proposons de recourir aux « Physics-Informed Neural Networks (PINNs) », qui permettent de combiner les données issues des points de mesures avec les équations différentielles partielles gouvernant la dynamique du phénomène étudié. Cette approche offre ainsi la possibilité d'améliorer la cohérence physique des reconstructions tout en renforçant leur capacité de généralisation en présence de données lacunaires ou bruitées. Elle sera développée dans la Section 2.7.3.

2.7.2 Régularisation tensorielle

Nous proposons dans cette section une nouvelle approche pour la cartographie spatio-temporelle utilisant une modélisation tensorielle des données. La méthode repose sur l'hypothèse que le processus de diffusion des polluants présente une variation lisse dans le temps et l'espace. Il est donc pertinent d'imposer des contraintes de douceur le long des dimensions spatiales x

et y ainsi que de la dimension temporelle t. Soit $\boldsymbol{\mathcal{X}}$ le tenseur de dimensions $N_y \times N_x \times N_t$ à reconstruire, représentant la distribution spatio-temporelle de la concentration des polluants. On note $\boldsymbol{\mathcal{Y}}$ le tenseur de mêmes dimensions que $\boldsymbol{\mathcal{X}}$ contenant les mesures temporelles de concentration de polluants sur ses fibres $\boldsymbol{\mathcal{Y}}(j,i,:)$, pour $(i,j) \in \Omega$, et nulles ailleurs. Soit également $\boldsymbol{\mathcal{H}}$ un tenseur de sélection, valant 1 aux positions $\boldsymbol{\mathcal{H}}(j,i,:)$ pour $(i,j) \in \Omega$, et 0 ailleurs.

Le problème de complétion tensorielle peut alors être formulé comme un problème d'optimisation :

$$\min_{\boldsymbol{\mathcal{X}}} \frac{1}{2}\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{H}} \boxdot \boldsymbol{\mathcal{X}}\|_F^2 + \frac{\lambda}{2}\|\boldsymbol{\mathcal{X}} \underset{1}{\bullet} \boldsymbol{D}_y\|_F^2 + \frac{\lambda}{2}\|\boldsymbol{\mathcal{X}} \underset{2}{\bullet} \boldsymbol{D}_x\|_F^2 + \frac{\mu}{2}\|\boldsymbol{\mathcal{X}} \underset{3}{\bullet} \boldsymbol{D}_t\|_F^2, \tag{2.22}$$

où "$\boxdot$" désigne le produit de Hadamard (produit élément par élément), et "$\underset{n}{\bullet}$" le produit n-mode d'un tenseur par une matrice. Les hyperparamètres λ et μ contrôlent respectivement la régularité spatiale et temporelle de la solution.

Le produit n-mode $\boldsymbol{\mathcal{X}} \underset{n}{\bullet} \boldsymbol{D}$ correspond à l'application de l'opérateur de régularisation le long de la dimension associée : le mode 1 (spatial) pour la direction y, le mode 2 (spatial) pour la direction x, et le mode 3 (temporelle) pour la dimension temporelle.

Les matrices $\boldsymbol{D}_\alpha$, avec $\alpha \in \{x, y, t\}$, sont des matrices de différentiation suivant la dimension α, de taille $(N_\alpha - 2) \times N_\alpha$, définies par :

$$\boldsymbol{D}_\alpha = \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & -2 & 1 \end{bmatrix}. \tag{2.23}$$

Soient $\boldsymbol{x} = \mathrm{vec}\{\boldsymbol{\mathcal{X}}\}$, $\boldsymbol{y} = \mathrm{vec}\{\boldsymbol{\mathcal{Y}}\}$ et $\boldsymbol{h} = \mathrm{vec}\{\boldsymbol{\mathcal{H}}\}$, où $\mathrm{vec}\{\cdot\}$ désigne l'opération de vectorisation d'une matrice ou d'un tenseur, telle que définie dans [Kolda09]. En utilisant les propriétés des opérateurs multilinéaires [Kolda06], on montre que :

$$\mathrm{vec}\{\boldsymbol{\mathcal{H}} \boxdot \boldsymbol{\mathcal{X}}\} = \mathrm{Diag}\{\boldsymbol{h}\}\boldsymbol{x},$$

$$\mathrm{vec}\{\boldsymbol{\mathcal{X}} \underset{1}{\bullet} \boldsymbol{D}_y\} = (\boldsymbol{I}_{N_t} \otimes \boldsymbol{I}_{N_x} \otimes \boldsymbol{D}_y)\boldsymbol{x},$$

$$\mathrm{vec}\{\boldsymbol{\mathcal{X}} \underset{2}{\bullet} \boldsymbol{D}_x\} = (\boldsymbol{I}_{N_t} \otimes \boldsymbol{D}_x \otimes \boldsymbol{I}_{N_y})\boldsymbol{x},$$

$$\mathrm{vec}\{\boldsymbol{\mathcal{X}} \underset{3}{\bullet} \boldsymbol{D}_t\} = (\boldsymbol{D}_t \otimes \boldsymbol{I}_{N_x} \otimes \boldsymbol{I}_{N_y})\boldsymbol{x},$$

où $\mathrm{Diag}\{\cdot\}$ désigne l'opérateur construisant une matrice diagonale à partir des composantes du

vecteur fourni, $\boldsymbol{I}_N$ est la matrice identité $N \times N$, et "$\otimes$" désigne le produit de Kronecker.

Ainsi, le problème (2.22) peut être reformulé sous forme vectorisée, comme un problème d'optimisation sans contrainte :

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}), \tag{2.24}$$

avec

$$\begin{aligned} f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y} - \mathrm{Diag}\{\boldsymbol{h}\}\boldsymbol{x}\|_2^2 + \frac{\lambda}{2}\|(\boldsymbol{I}_{N_t} \otimes \boldsymbol{I}_{N_x} \otimes \boldsymbol{D}_y)\boldsymbol{x}\|_2^2 \\ + \frac{\lambda}{2}\|(\boldsymbol{I}_{N_t} \otimes \boldsymbol{D}_x \otimes \boldsymbol{I}_{N_y})\boldsymbol{x}\|_2^2 + \frac{\mu}{2}\|(\boldsymbol{D}_t \otimes \boldsymbol{I}_{N_x} \otimes \boldsymbol{I}_{N_y})\boldsymbol{x}\|_2^2. \end{aligned}$$

Le problème (2.24) admet une solution explicite :

$$\boldsymbol{x} = \left(\mathrm{Diag}\{\boldsymbol{h} \boxdot \boldsymbol{h}\} + \lambda\, \boldsymbol{I}_{N_t N_x} \otimes \boldsymbol{D}_y^{\mathsf{T}}\boldsymbol{D}_y + \lambda\, \boldsymbol{I}_{N_t} \otimes \boldsymbol{D}_x^{\mathsf{T}}\boldsymbol{D}_x \otimes \boldsymbol{I}_{N_y} + \mu\, \boldsymbol{D}_t^{\mathsf{T}}\boldsymbol{D}_t \otimes \boldsymbol{I}_{N_x N_y}\right)^{-1} \mathrm{Diag}\{\boldsymbol{h}\}\, \boldsymbol{y}. \tag{2.25}$$

La cartographie spatio-temporelle est obtenue en re-organisant les valeurs du vecteur estimé $\boldsymbol{x}$ en un tenseur $N_y \times N_x \times N_t$. Le calcul de (2.25) nécessite l'inversion d'une matrice de taille $N_x N_y N_t \times N_x N_y N_t$, potentiellement très grande, ce qui constitue le principal « goulot d'étranglement » de la méthode. Cependant, cette matrice étant creuse et définie positive, des méthodes itératives efficaces permettent son inversion. Ici nous employons une méthode d'inversion basée sur le *gradient conjugué préconditionné* (PGC [1]), permettant de résoudre le système linéaire associé en moins d'une minute dans le cadre de l'application étudiée. Chaque point de la carte de pollution à estimer correspond à une concentration de polluant, une grandeur qui est par nature non négative. Cependant, dans la formulation (2.24), la fonction coût ne garantit pas la nonnégativité des cartes de concentration obtenues. Afin d'imposer cette contrainte de façon explicite, nous ajoutons un terme supplémentaire à la fonction coût et reformulons le problème comme suit :

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) + I_{\mathbb{R}_+}(\boldsymbol{x}), \tag{2.26}$$

où $I_{\mathbb{R}_+}(\boldsymbol{x})$ est la fonction indicatrice de l'orthant positif $\mathbb{R}_+$. Le minimiseur de (2.26) est calculé au moyen d'un algorithme de *gradient projeté*, mettant à jour $\boldsymbol{x}$ selon :

$$\boldsymbol{x} \longleftarrow \Pi_{\mathbb{R}_+}(\boldsymbol{x} - \eta \nabla_{\boldsymbol{x}} f(\boldsymbol{x})), \tag{2.27}$$

1. *Preconditioned Conjugate Gradient*, en anglais

Dans (2.27), $\Pi_{\mathbb{R}_+}(\cdot)$ désigne la projection sur l'orthant positif, et le gradient est donné par :

$$\begin{aligned}\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) = \big(&\operatorname{Diag}\{\boldsymbol{h} \boxdot \boldsymbol{h}\} + \lambda\, \boldsymbol{I}_{N_t N_x} \otimes \boldsymbol{D}_y^\top \boldsymbol{D}_y + \lambda\, \boldsymbol{I}_{N_t} \otimes \boldsymbol{D}_x^\top \boldsymbol{D}_x \otimes \boldsymbol{I}_{N_y} \\ &+ \mu\, \boldsymbol{D}_t^\top \boldsymbol{D}_t \otimes \boldsymbol{I}_{N_x N_y}\big)\boldsymbol{x} - \operatorname{Diag}\{\boldsymbol{h}\}\boldsymbol{y}. \end{aligned} \tag{2.28}$$

La cartographie spatio-temporelle est ensuite reconstruite en reformant le tenseur à partir de la solution vectorielle $\boldsymbol{x}^\star$ du problème (2.25), *i.e.,* $\boldsymbol{\mathcal{X}}^\star = \operatorname{unvec}\{\boldsymbol{x}^\star\}$.

L'algorithme du gradient projeté (2.27) alterne entre un pas de descente de gradient suivi par une projection sur $\mathbb{R}_+$ et est garanti de converger vers un point stationnaire sous des conditions non-restrictives. Néanmoins, une initialisation aléatoire peut entraîner une convergence lente de la procédure. Pour accélérer l'algorithme, nous utilisons comme point de départ la solution non contrainte (2.25) et adoptons une technique de « backtracking » afin de déterminer le pas optimal η ([Beck17][Section 10.4.2]). L'algorithme global : descente de gradient conjugué avec préconditionnement et gradient projeté avec *backtracking*, est résumée dans l'Algorithme 1. Une illustration des résultats obtenus par cette méthode de *régularisation tensorielle* est disponible sur la figure 2.12. La comparaison entre le krigeage et la régularisation tensorielle révèle des écarts notables tant visuels que quantitatifs. Comme illustré à la Figure 2.12, le krigeage produit des reconstructions globalement lisses et cohérentes, mais au prix d'un lissage excessif des gradients marqués et d'une atténuation des structures locales, limitant ainsi la précision dans les zones à forte variabilité spatiale (proche des sources de pollution).

Quant à la régularisation tensorielle, fondée sur l'exploitation explicite de la variable spatio-temporelle, elle permet de mieux préserver la dynamique des structures tout en maintenant la continuité temporelle. Les champs reconstruits présentent ainsi une meilleure fidélité aux données de référence, avec des contours plus définis et une réduction des artefacts résiduels.

L'évaluation quantitative reportée à la Figure 2.13 confirme ces observations. Le krigeage apparaît comme une méthode fiable, mais la régularisation tensorielle conduit, dans nos expériences, à des valeurs de RMSE plus faibles ou comparables, avec en outre une variabilité réduite selon les grilles testées. Cette stabilité suggère une meilleure capacité de généralisation, probablement liée à l'intégration conjointe de la régularité spatiale et temporelle.

Algorithm 1: Régularisation tensorielle par gradient projeté

Entrées : mesures $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{N_y \times N_x \times N_t}$ ($\boldsymbol{y} = \mathrm{vec}\{\boldsymbol{\mathcal{Y}}\}$), masque $\boldsymbol{\mathcal{H}} \in \{0,1\}^{N_y \times N_x \times N_t}$ ($\boldsymbol{h} = \mathrm{vec}\{\boldsymbol{\mathcal{H}}\}$) ; les opérateurs de dérivation $\boldsymbol{D}_y \in \mathbb{R}^{(N_y-2)\times N_y}$, $\boldsymbol{D}_x \in \mathbb{R}^{(N_x-2)\times N_x}$, $\boldsymbol{D}_t \in \mathbb{R}^{(N_t-2)\times N_t}$; (hyper-)paramètres $\lambda > 0$, $\mu > 0$, $\eta_0 > 0$, $\beta \in (0,1)$; tolérance $\varepsilon > 0$; itérations max $K_{\max}$.

Sortie : cartographie spatio-temporelle $\boldsymbol{\mathcal{X}}^\star \in \mathbb{R}_+^{N_y \times N_x \times N_t}$.

Définitions (opérateurs)

$f(\boldsymbol{x}) \leftarrow \frac{1}{2}\|\boldsymbol{y} - \mathrm{Diag}\{\boldsymbol{h}\}\boldsymbol{x}\|_2^2 + \frac{\lambda}{2}\|(\boldsymbol{I}_{N_t} \otimes \boldsymbol{I}_{N_x} \otimes \boldsymbol{D}_y)\boldsymbol{x}\|_2^2 + \frac{\lambda}{2}\|(\boldsymbol{I}_{N_t} \otimes \boldsymbol{D}_x \otimes \boldsymbol{I}_{N_y})\boldsymbol{x}\|_2^2 + \frac{\mu}{2}\|(\boldsymbol{D}_t \otimes \boldsymbol{I}_{N_x} \otimes \boldsymbol{I}_{N_y})\boldsymbol{x}\|_2^2$ `// fonction coût`

$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) \leftarrow \big(\mathrm{Diag}\{\boldsymbol{h} \boxdot \boldsymbol{h}\} + \lambda\, \boldsymbol{I}_{N_t N_x} \otimes \boldsymbol{D}_y^\mathsf{T}\boldsymbol{D}_y + \lambda\, \boldsymbol{I}_{N_t} \otimes \boldsymbol{D}_x^\mathsf{T}\boldsymbol{D}_x \otimes \boldsymbol{I}_{N_y} + \mu\, \boldsymbol{D}_t^\mathsf{T}\boldsymbol{D}_t \otimes \boldsymbol{I}_{N_x N_y}\big)\boldsymbol{x} - \mathrm{Diag}\{\boldsymbol{h}\}\boldsymbol{y}$ `// gradient de la fonction coût`

$\Pi_{\mathbb{R}^+}(\cdot) \leftarrow \max(\cdot, 0)$ `// projection sur` $\mathbb{R}_+$ `élément par élément`

Warm start (solution non contrainte)

Résoudre pour $\boldsymbol{x}$: $\big[\big(\mathrm{Diag}\{\boldsymbol{h} \boxdot \boldsymbol{h}\} + \lambda\, \boldsymbol{I}_{N_t N_x} \otimes \boldsymbol{D}_y^\mathsf{T}\boldsymbol{D}_y + \lambda\, \boldsymbol{I}_{N_t} \otimes \boldsymbol{D}_x^\mathsf{T}\boldsymbol{D}_x \otimes \boldsymbol{I}_{N_y} + \mu\, \boldsymbol{D}_t^\mathsf{T}\boldsymbol{D}_t \otimes \boldsymbol{I}_{N_x N_y}\big)\,\mathrm{Diag}\{\boldsymbol{h}\}\big]\boldsymbol{x} = \boldsymbol{y}$ `// PCG`

$\boldsymbol{x}^{(0)} \leftarrow \Pi_{\mathbb{R}^+}(\boldsymbol{x})$, $k \leftarrow 0$

while $k < K_{\max}$ **do**

  $\eta \leftarrow \eta_0$

  while $f\big(\Pi_{\mathbb{R}^+}(\boldsymbol{x}^{(k)} - \eta\, \nabla_{\boldsymbol{x}} f(\boldsymbol{x}^{(k)}))\big) > f(\boldsymbol{x}^{(k)}) + \langle \nabla_{\boldsymbol{x}} f(\boldsymbol{x}^{(k)}), \Pi_{\mathbb{R}^+}(\boldsymbol{x}^{(k)} - \eta\, \nabla_{\boldsymbol{x}} f(\boldsymbol{x}^{(k)}) - \boldsymbol{x}^{(k)}\rangle + \eta \Pi_{\mathbb{R}^+} \|\boldsymbol{x}^{(k)} - \eta\, \nabla_{\boldsymbol{x}} f(\boldsymbol{x}^{(k)}) - \boldsymbol{x}^{(k)}\|_2^2$ **do**

    $\eta \leftarrow \beta\, \eta$ `// Backtracking`

  $\boldsymbol{x}^{(k+1)} \leftarrow \Pi_{\mathbb{R}^+}(\boldsymbol{x}^{(k)} - \eta\, \nabla_{\boldsymbol{x}} f(\boldsymbol{x}^{(k)}))$

  if $\dfrac{\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\|_2}{\|\boldsymbol{x}^{(k)}\|_2} \leq \varepsilon$ **then**

    return $\boldsymbol{x}^{(k+1)}$

  $k \leftarrow k + 1$

$\boldsymbol{\mathcal{X}}^\star = \mathrm{unvec}\{\boldsymbol{x}^{(k+1)}\}$

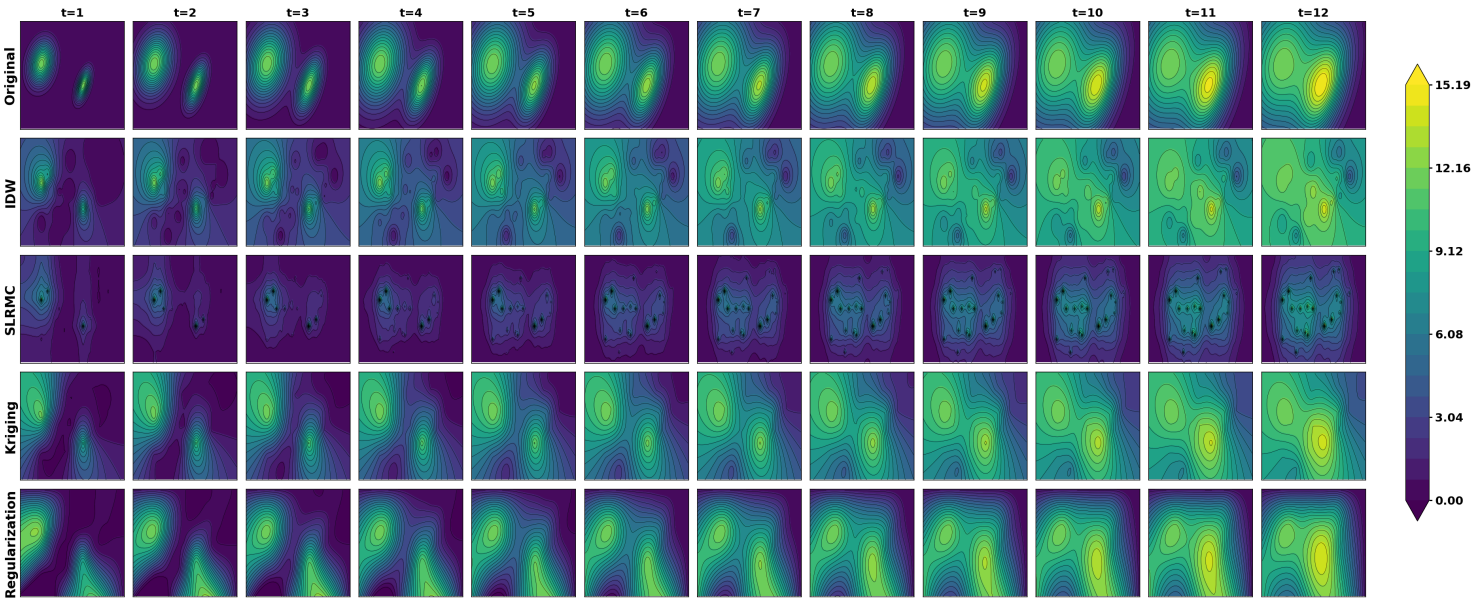


FIGURE 2.12 – Cartes reconstruites à partir de 20 points des grilles d'échantillonnage présentées en figure 2.5 avec les méthodes IDW, Krigeage, SLMRC et Régularisation.

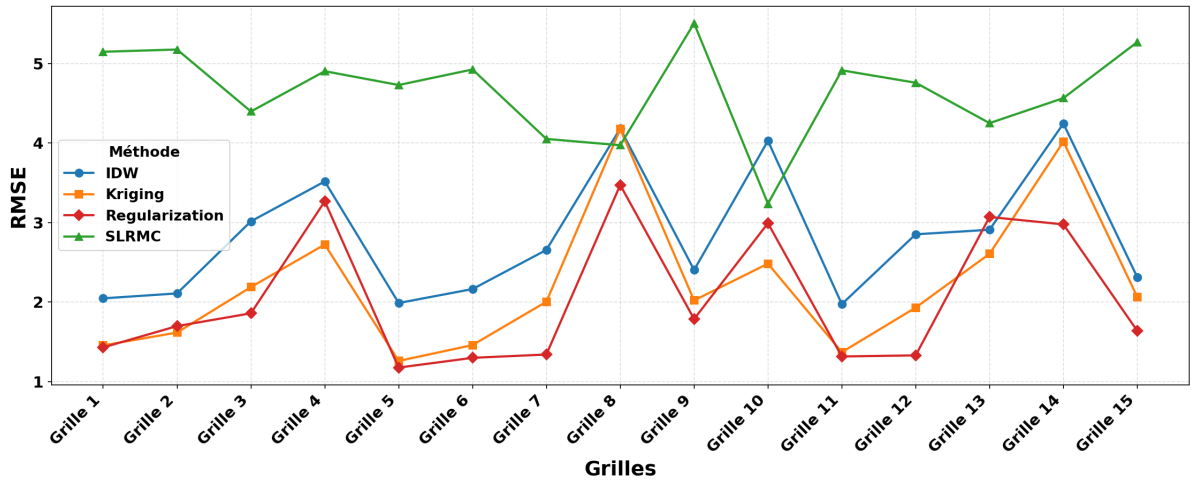


FIGURE 2.13 – Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 20 points de la figure 2.6, obtenue par les méthodes IDW, krigeage, SLRMC et Régularisation tensorielle

2.7.3 Réseaux de neurones informés par la physique

La méthode de régularisation tensorielle repose sur l'hypothèse d'une diffusion lente et uniforme du polluant étudié. Nous proposons d'aller au-delà de cette hypothèse en intégrant explicitement des connaissances physiques et structurelles relatives à la dynamique des polluants atmosphériques (par exemple des équations de diffusion-reaction).

Les réseaux de neurones artificiels (« Artificial Neural Networks », ANN) classiques sont

largement utilisés pour l'approximation de fonctions complexes et la régression non linéaire à partir de jeux de données denses et représentatifs. Toutefois, leur application à la cartographie spatio-temporelle des concentrations de polluants atmosphériques présente plusieurs limitations majeures. D'une part, ces modèles sont essentiellement *data-driven* : ils nécessitent un échantillonnage dense des variables d'entrée pour produire des reconstructions fiables. Or, dans un environnement de travail instrumenté par des capteurs, les mesures disponibles sont souvent sous-échantillonnées spatialement et bruitées, ce qui compromet la capacité prédictive des ANN standards [Goodfellow16].

D'autre part, les ANN classiques ne tiennent généralement pas compte des lois physiques gouvernant la dispersion des polluants (diffusion). Leur apprentissage se limite donc à capturer des corrélations statistiques locales sans garantir la cohérence avec les équations différentielles partielles (EDP) régissant le phénomène. Enfin, ces modèles ont tendance à sur-apprendre (*overfitting*) les données d'entraînement en l'absence de régularisation physique, ce qui réduit leur capacité de généralisation sur des points non instrumentés [Brunton20].

Ces limitations justifient le recours aux réseaux de neurones informés par la physique (« Physics-Informed Neural Networks », PINNs), introduits par Raissi et al. [Raissi19]. Contrairement aux ANN classiques, les PINNs intègrent explicitement les équations de conservation dans leur fonction de coût, de sorte que la solution apprise respecte à la fois les données de capteurs et les contraintes physiques du système étudié. Cette hybridation entre modèles statistiques et lois de la physique a montré son efficacité pour résoudre des problèmes en mécanique des fluides, en dynamique des gaz, ou encore en transport de polluants [Karniadakis21, Cuomo22].

2.7.3.1 Formulation du problème

La dynamique des polluants en milieu intérieur peut être modélisée par une équation aux dérivées partielles (EDP) de type diffusion–réaction. Considérons à cet effet un champ scalaire $\boldsymbol{X}(x, y, t)$, solution de cette EDP pouvant s'écrire sous la forme générale :

$$\frac{\partial \boldsymbol{X}}{\partial t}(x, y, t) - k(t)\, \nabla^2_{xy} \boldsymbol{X}(x, y, t) - f(x, y, t) = 0 \tag{2.29}$$

où $k(t) > 0$ désigne un coefficient de diffusion dépendant du temps, $f(x, y, t)$ un terme source représentant les phénomènes réactionnels ou les émissions spatio-temporelles, et $\nabla^2_{xy} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ est l'opérateur laplacien 2D. Dans ce modèle PINN, le coefficient de diffusion $k(t)$ et le terme source $f(x, y, t)$ ne sont pas imposés a priori mais appris directement par le réseau. Cette formulation permet de capturer la variabilité temporelle de la diffusion et d'inférer des sources locales ou diffuses en cohérence avec les données et les contraintes physiques, à l'instar de certains travaux récents en PINNs qui ont montré la possibilité d'identifier des coefficients de diffusion et/ou des termes sources à partir de données éparses ou incomplètes [Raissi19, Karniadakis21,

Chen21].

Le principe du PINN consiste à approximer la solution $\boldsymbol{X}(x, y, t)$ par une fonction $\hat{\boldsymbol{X}}(x, y, t)$ paramétrée par les poids et biais $(\boldsymbol{W}, \boldsymbol{b})$ du réseau de neurones. La structure du réseau est représentée sur la Figure 2.15. Les variables spatio-temporelles (x, y, t) constituent les entrées du modèle, tandis que la sortie correspond à l'approximation $\hat{\boldsymbol{X}}(x, y, t)$. Cela nécessite l'échantillonnage de différents ensembles de points dans le domaine spatio-temporel considéré, chaque catégorie contribuant spécifiquement à la définition de la fonction de perte et à l'optimisation du modèle. On distingue :

— **Points de mesures** : ce sont les positions spatio-temporelles où des mesures de concentrations $\boldsymbol{X}(x, y, t)$ sont effectuées (Voir Figure 2.14a). Ils permettent de définir la *data loss* (Voir l'équation 2.30), formulée sous la forme d'une erreur quadratique moyenne entre les données mesurées et les prédictions du réseau $\hat{\boldsymbol{X}}(x, y, t)$. Ces points constituent l'information empirique nécessaire à l'ancrage du modèle dans la réalité expérimentale, bien que leur nombre soit souvent limité.

$$\mathcal{L}_{\text{data}} = \frac{1}{N_{\text{data}}} \sum_{i=1}^{N_{\text{data}}} \left(\hat{\boldsymbol{\mathcal{X}}}(x_i, y_i, t_i) - \boldsymbol{\mathcal{X}}(x_i, y_i, t_i) \right)^2 \tag{2.30}$$

— **Points de collocation** : ils sont échantillonnés à l'intérieur du domaine spatio-temporel indépendamment des données expérimentales. Leur rôle est d'imposer le respect de l'équation aux dérivées partielles gouvernante, en évaluant le résidu de l'EDP par différentiation automatique des sorties du réseau. Les points de collocation interviennent donc dans la *physical loss* (Voir equation 2.31) et permettent d'exploiter la connaissance *a priori* du système physique afin de régulariser l'apprentissage dans des zones dépourvues de mesures.

$$\mathcal{L}_{\text{physical}} = \frac{1}{N_{\text{phys}}} \sum_{j=1}^{N_{\text{phys}}} \left(\frac{\partial \hat{\boldsymbol{X}}}{\partial t}(x_j, y_j, t_j) - k(t_j)\, \nabla^2_{xy} \hat{\boldsymbol{X}}(x_j, y_j, t_j) - f(x_j, y_j, t_j) \right)^2, \tag{2.31}$$

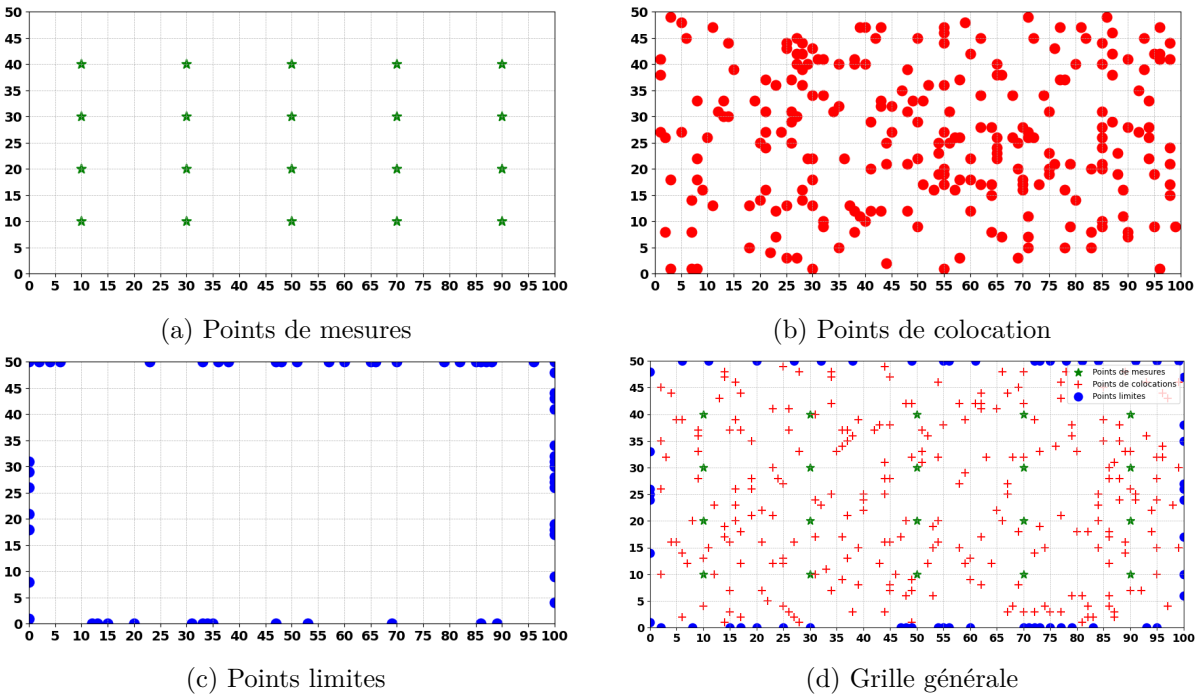
avec N_{phys} le nombre de points de collocation choisi. Le nombre et la distribution des points de collocation constituent des paramètres déterminants pour la stabilité et la précision des PINNs. Un échantillonnage insuffisant compromet l'approximation des contraintes physiques, tandis qu'un excès de points accroît le coût computationnel sans garantir une meilleure généralisation [Raissi19]. Pour surmonter cette difficulté, différentes stratégies ont été proposées : l'échantillonnage uniforme ou quasi-aléatoire, par séquences de Sobol ou de Halton, assurant une couverture homogène du domaine [Raissi19, Yang21] ; l'adaptation dynamique, qui concentre les points dans les zones de forts résidus ou gradients [Lu21, Wang21] ; et les approches hybrides, combinant un maillage initial grossier à un raffinement adaptatif fondé sur l'évolution de l'erreur ou de

la variance prédite [Karniadakis21, Daw22]. Dans ce travail, nous retenons une sélection aléatoire uniforme dans l'espace étudié (voir Figure 2.14b).

— **Points limites (boundary points)** : Situés sur les frontières spatiales du domaine, ils assurent l'application des conditions aux limites (Voir Figure 2.14c). Les contraintes de Dirichlet (valeur définie) et de Neumann (flux normal) peuvent être intégrées à travers ces points, qui participent à la définition de la *boundary loss*. Ils sont essentiels pour garantir la cohérence physique de la solution approchée avec les spécifications imposées au système. Dans notre cas nous avons utilisé une condition aux limites imposées par les contraintes de Neumann correspondant à l'hypothèse d'imperméabilité des parois (Voir équation 2.32).

$$\mathcal{L}_{\text{boundary}} = \frac{1}{N_{\text{bound}}} \sum_{l=1}^{N_{\text{bound}}} \left(\langle \nabla_{xy} \hat{\boldsymbol{X}}(x_l, y_l, t_l), \vec{n} \rangle - \phi(x_l, y_l, t_l) \right)^2, \tag{2.32}$$

où N_{bound} représente le nombre de point limites choisis, $\vec{n}$ le vecteur normal unitaire à la frontière et $\phi(x_l, y_l, t_l)$ le flux imposé à la frontière. L'hypothèse d'imperméabilité des parois se traduit par l'absence de transfert de polluant à travers les frontières du domaine étudié ($\phi(x_l, y_l, t_l) = 0$), ce qui correspond à un flux nul normal à la frontière, assimilable à une condition de type « mur étanche ».



(a) Points de mesures

(b) Points de colocation

(c) Points limites

(d) Grille générale

FIGURE 2.14 – Exemple d'échantillonnage des points de mesures, de collocation et de frontières.

En se basant sur la grille d'échantillonnage générale de l'espace étudié (voir Figure 2.14d), l'entraînement du réseau repose sur la minimisation d'une fonction de perte composite intégrant

ces trois fonctions coût :

$$\begin{aligned}\mathcal{L}(x,y,t) = \lambda_{\text{data}} \quad & \frac{1}{N_x N_y N_t} \sum_{i=1}^{N_x N_y N_t} \left(\hat{\boldsymbol{X}}(x_i,y_i,t_i) - \boldsymbol{X}(x_i,y_i,t_i)\right)^2 \\ + \lambda_{\text{phys}} \quad & \frac{1}{N_{\text{phys}}} \sum_{j=1}^{N_{\text{phys}}} \left(\frac{\partial \hat{\boldsymbol{X}}}{\partial t}(x_j,y_j,t_j) - k(t_j)\nabla^2_{xy}\hat{\boldsymbol{X}}(x_j,y_j,t_j) - f(x_j,y_j,t_j)\right)^2 \\ + \lambda_{\text{bound}} \quad & \frac{1}{N_{\text{bound}}} \sum_{l=1}^{N_{\text{bound}}} \left(\langle \nabla_{xy}\hat{\boldsymbol{X}}(x_l,y_l,t_l), \vec{n}\rangle - \phi(x_l,y_l,t_l)\right)^2, \end{aligned} \tag{2.33}$$

où $\lambda_{\text{data}}, \lambda_{\text{phys}}, \lambda_{\text{bound}}$ sont des hyperparamètres permettant d'ajuster la contribution des trois fonctions coûts dans la fonction de perte. La minimisation est faite par rapport aux paramètres du réseau de neurones, $\boldsymbol{W}$ et $\boldsymbol{b}$. Le coefficient de diffusion $k(t)$ ainsi que le terme source $f(x,y,t)$ sont appris en même temps à l'aide de deux sous-réseaux.

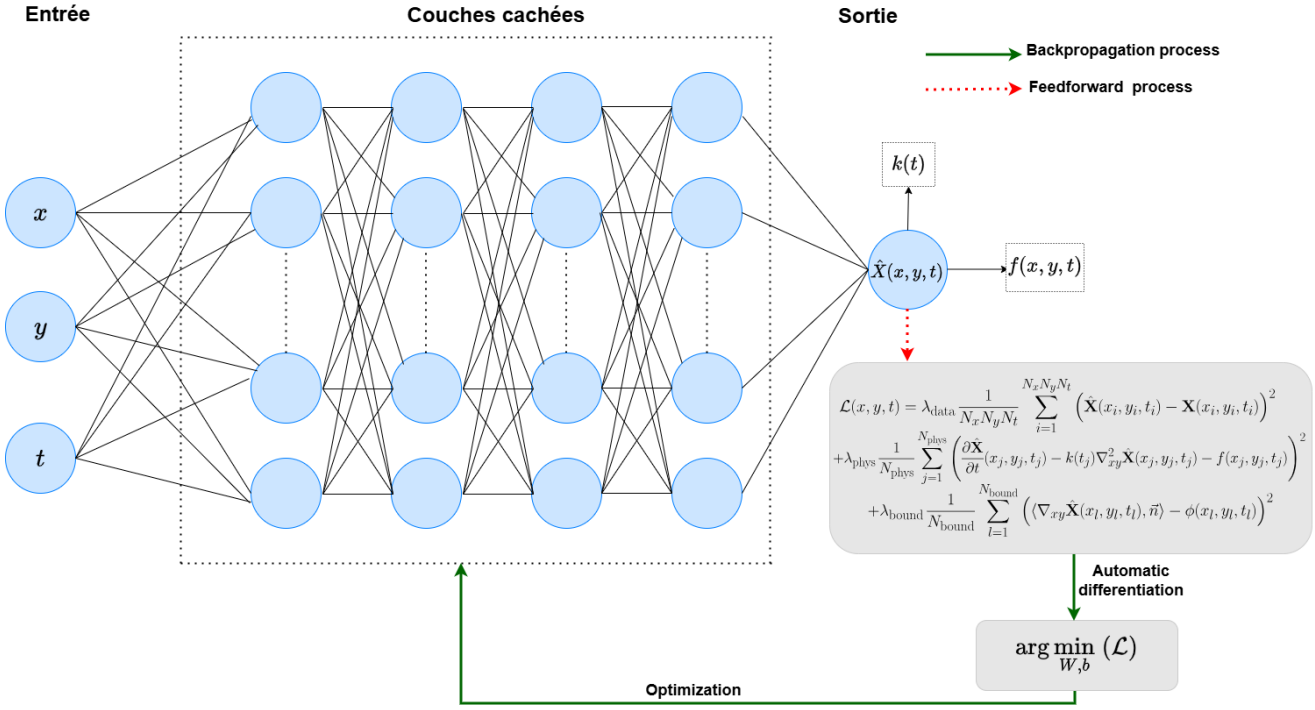


FIGURE 2.15 – Architecture du réseau de neurones informé par la physique pour la résolution de PDE de type diffusion–réaction

En effet, la solution $\boldsymbol{X}(x,y,t)$ est approchée par une fonction $\hat{\boldsymbol{X}}(x,y,t;\boldsymbol{W},\boldsymbol{b})$ dont les paramètres $(\boldsymbol{W},\boldsymbol{b})$ représentent les poids et biais du réseau principal. Deux sous-réseaux distincts sont entraînés conjointement afin d'approximer les fonctions inconnues $k(t)$ et $f(x,y,t)$, notées respectivement $\hat{k}(t;\boldsymbol{W}_k,\boldsymbol{b}_k)$ et $\hat{f}(x,y,t;\boldsymbol{W}_f,\boldsymbol{b}_f)$. Ainsi, l'ensemble du modèle est constitué de trois composantes apprises simultanément :

— le réseau principal $\hat{\boldsymbol{X}}$, chargé d'approximer la concentration du polluant ;

— le sous-réseau temporel $\hat{k}(t)$, dédié à l'identification du coefficient de diffusion dépendant

du temps ;

— le sous-réseau $\hat{f}(x, y, t)$, destiné à la reconstruction du terme source spatio-temporel.

Le réseau est de type « Fully Connected Feedforward Neural Network » (FNN) entièrement connecté, dont les entrées sont les variables spatio-temporelles (x, y, t) et la sortie correspond à la concentration prédite $\hat{\boldsymbol{X}}(x, y, t)$. Les dérivées spatiales et temporelles nécessaires à l'évaluation du résidu de l'équation différentielle sont obtenues par différentiation automatique. L'architecture retenue s'inspire des travaux fondateurs de [Raissi19], où les auteurs démontrent qu'un FNN de profondeur modérée (typiquement 4 à 8 couches cachées de 20 à 100 neurones chacune, avec fonction d'activation *tanh*) permet une approximation efficace des solutions de PDEs sous contrainte physique. *tanh* produit des surfaces et des gradients lisses, plus conformes aux comportements physiques que des activations non différentiables comme *ReLU*. Même si d'autres activations comme *sine*, *Swish* ou *GELU* ont été explorées récemment pour les PINNs, *tanh* reste le standard de référence car elle offre un compromis robuste entre stabilité, précision et régularité. Ce choix d'activation assure donc une continuité C^∞ favorable à la dérivation automatique. Des architectures comparables ont été appliquées avec succès à la résolution d'équations de diffusion–réaction [Giampaolo22].

Les questions de stabilité de l'apprentissage et d'équilibrage entre pertes physique et empirique ont par ailleurs été approfondies par [Wang21], qui proposent des stratégies d'atténuation des erreurs de gradient afin d'améliorer la convergence et la robustesse des modèles. Ces travaux confirment que des architectures FNN de complexité modérée, associées à des fonctions d'activation différentiables telles que *tanh*, offrent un compromis optimal entre stabilité numérique et cohérence physique pour la résolution de PDEs informées par la physique.

L'entraînement s'effectue en deux phases successives : une optimisation stochastique par l'algorithme Adam, suivie d'un affinage par un optimiseur quasi-Newton de type L-BFGS, assurant une convergence stable et un meilleur respect des contraintes physiques. La fonction de perte totale est donnée par l'équation (2.33), qui combine les termes de pertes expérimentales, physique et aux limites. Un équilibrage des pondérations associées à ces trois termes (λ_{data}, λ_{phys}, λ_{bound}) dans l'équation 2.33 est effectué afin d'éviter qu'une composante de la perte ne domine les autres et ne perturbe la convergence de l'optimisation globale, conformément aux recommandations récentes sur la stabilisation de l'entraînement des PINNs [Wang21].

Concernant les conditions initiales, le champ de concentration au temps initial $t = 0$ est défini par :

$$\boldsymbol{X}(x, y, 0) = \boldsymbol{X}_0(x, y), \tag{2.34}$$

où $\boldsymbol{X}_0(x, y)$ représente la distribution initiale du polluant sur le domaine spatial. Lorsque les mesures expérimentales sont disponibles à $t = 0$, ces valeurs sont directement intégrées dans la fonction de perte.

Cette condition initiale assure la cohérence temporelle dès $t = 0$ et constitue le point d'ancrage du processus d'apprentissage. Par ailleurs, les conditions aux limites imposées par l'équation (2.32), correspondant à des parois imperméables ($\phi(x_l, y_l, t_l) = 0$), garantissent l'absence de flux à la frontière du domaine, ce qui est compatible avec la condition initiale et permet d'éviter toute discontinuité du flux normal au temps initial.

En appliquant cette approche aux données synthétiques, nous obtenons les cartes présentées à la figure 2.16.

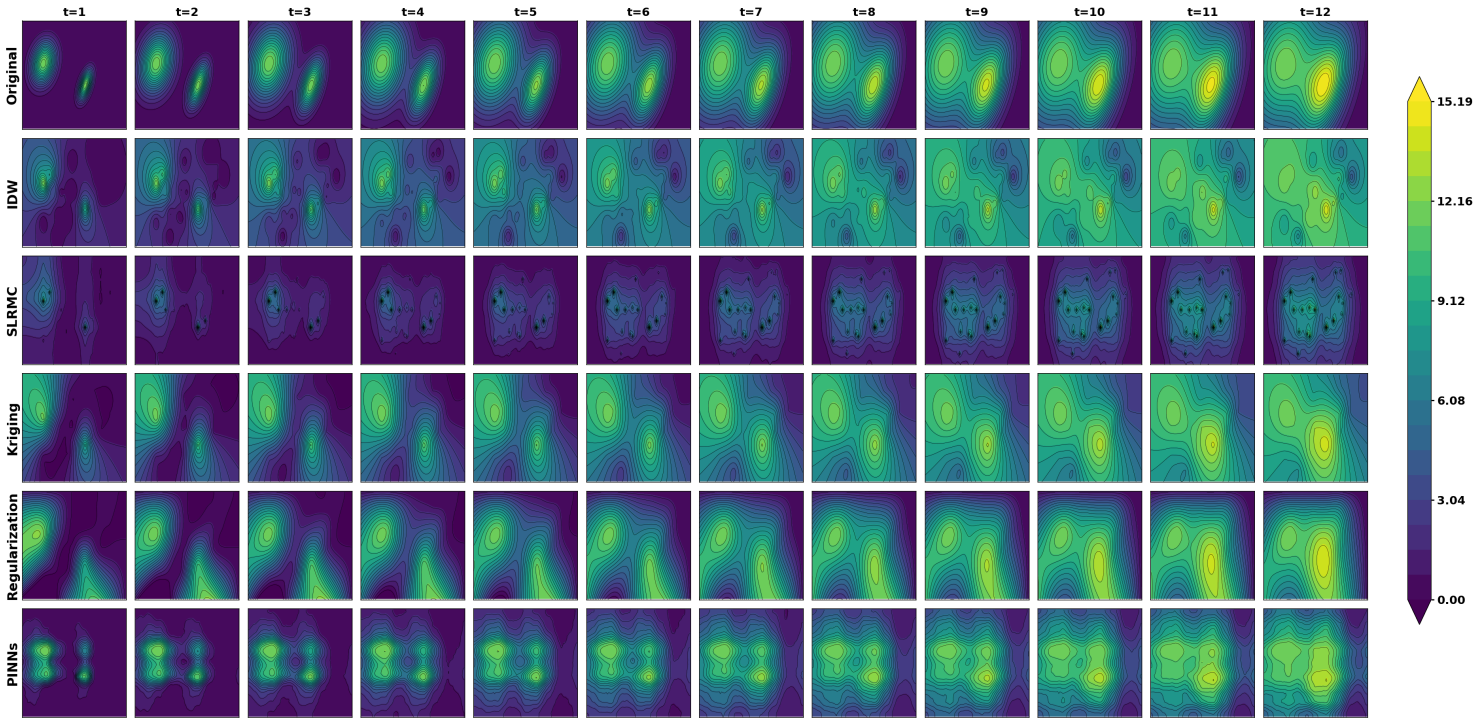


FIGURE 2.16 – Cartes reconstruites à partir de 20 points des grilles d'échantillonnage présentées en figure 2.5 avec les méthodes IDW, Krigeage, SLMRC, Régularisation et PINNs.

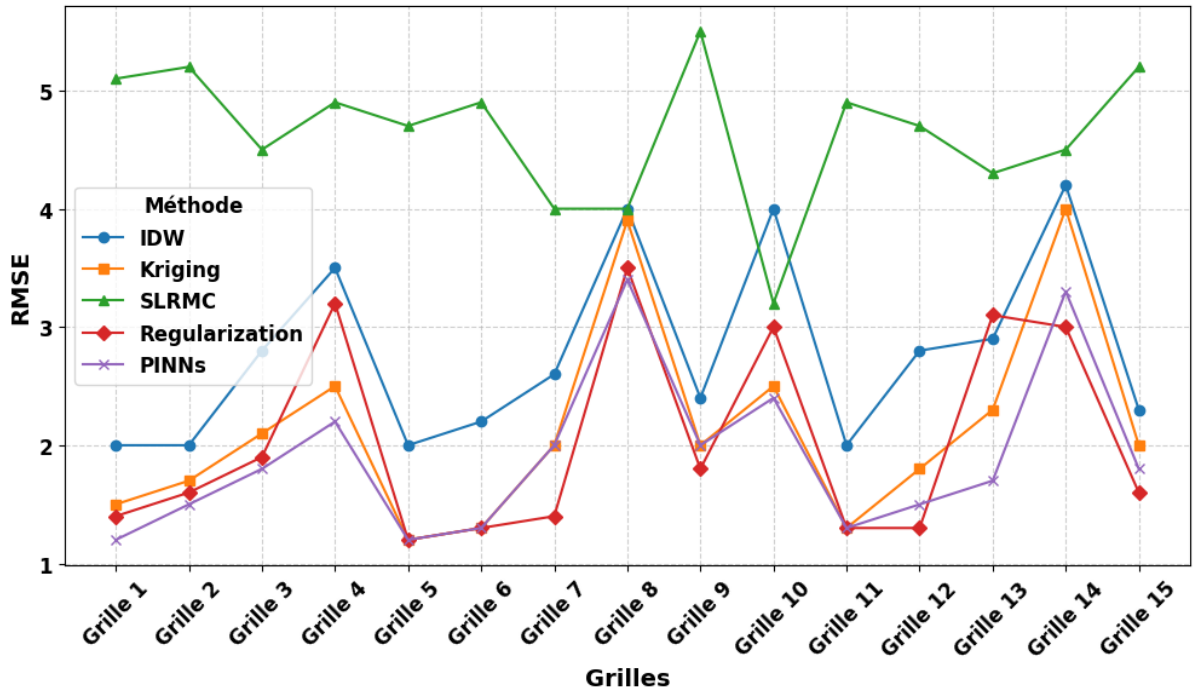


FIGURE 2.17 – Erreur de reconstruction des cartes présentées en figure 2.2 à partir des grilles d'échantillonnage à 20 points de la figure 2.6, obtenue par les méthodes IDW, krigeage, SLRMC, Régularisation tensorielle et PINNs

Le krigeage produit des reconstructions globalement lisses, mais présente une tendance marquée à atténuer les gradients spatiaux. L'introduction d'un modèle tensoriel avec régularisation spatio-temporelle améliore la fidélité des estimations en préservant davantage les structures locales ainsi que la continuité temporelle. Quant à la méthode PINNs, elle permet d'obtenir des reconstructions cohérentes et conformes aux principes physiques, y compris dans des situations de sous-échantillonnage. Les résultats illustrés par les courbes de RMSE (figure 2.17) indiquent que la régularisation tensorielle et les PINNs se distinguent du krigeage, en offrant une meilleure stabilité et une réduction de l'erreur moyenne.

Chapitre 3

Méthodes et stratégies de mesures

Sommaire

3.1 Introduction 52
3.2 Organisation générale des essais 52
3.3 Description des espaces expérimentaux 53
3.3.1 Chambre expérimentale 54
3.3.2 Salle aménagée à l'INRS 55
3.4 Stratégie de déploiement des instruments de mesures et des sources de pollution 56
3.4.1 Stratégie de mesure à points fixes 56
3.4.2 Stratégie de mesure à points mobiles 57
3.4.3 Stratégie de pollution 57
3.4.4 Déroulement des essais 57
3.5 Description des sources de pollution 61
3.5.1 Source de pollution gaz 61
3.5.2 Source de pollution particules 61
3.6 Méthodes et instruments de mesures 62
3.6.1 Réseau de capteurs sans fil gaz (RCSFg) 62
3.6.2 Réseau filaire de capteurs particules (RFCp) 64
3.6.3 Détecteurs de CO_2 portables = instruments de référence 64
3.6.4 Thermo-hygromètre 65
3.6.5 Scanner 3D 65
3.6.6 Robot mobile 66
3.6.7 Station météo 66
3.7 Vérification et étalonnage des instruments de mesures 67
3.7.1 Étalonnage des instruments de mesures gaz 67
3.7.2 Étalonnage du réseau filaire de capteurs particules 70
3.8 Pré-traitement des données 72

3.1 Introduction

Ce chapitre présente l'ensemble des méthodes et choix techniques adoptés pour la mesure des concentrations de polluants gazeux et particulaires au sein d'environnements expérimentaux maîtrisés. Il vise à établir un cadre structuré décrivant l'organisation des essais, le déploiement des dispositifs de mesure ainsi que le traitement des données exploitées dans ce manuscrit.

Dans un premier temps, nous décrivons l'organisation générale des campagnes de mesure, ainsi que les deux espaces expérimentaux sélectionnés à l'INRS ;

Nous exposons ensuite la stratégie de positionnement des capteurs et des sources de pollution, ainsi que le déroulement pratique des essais.

Les instruments de mesure sont ensuite présentés en détail : capteurs de gaz (réseaux sans fil), capteurs de particules (réseau filaire), détecteurs de CO_2, mais aussi des dispositifs complémentaires tels que le thermo-hygromètre, le scanner 3D et la station météo.

Enfin les méthodes de vérification, d'étalonnage et de traitement des données collectées sont décrites en vue de leur exploitation dans le cadre de la réalisation des cartographies spatio-temporelles des concentrations de polluants et de la reconstruction de l'exposition individuelle aux polluants.

3.2 Organisation générale des essais

Les essais ont été réalisés selon un protocole expérimental permettant d'assurer des conditions de mesure contrôlées et le plus reproductible possible (Voir Figure 3.1).

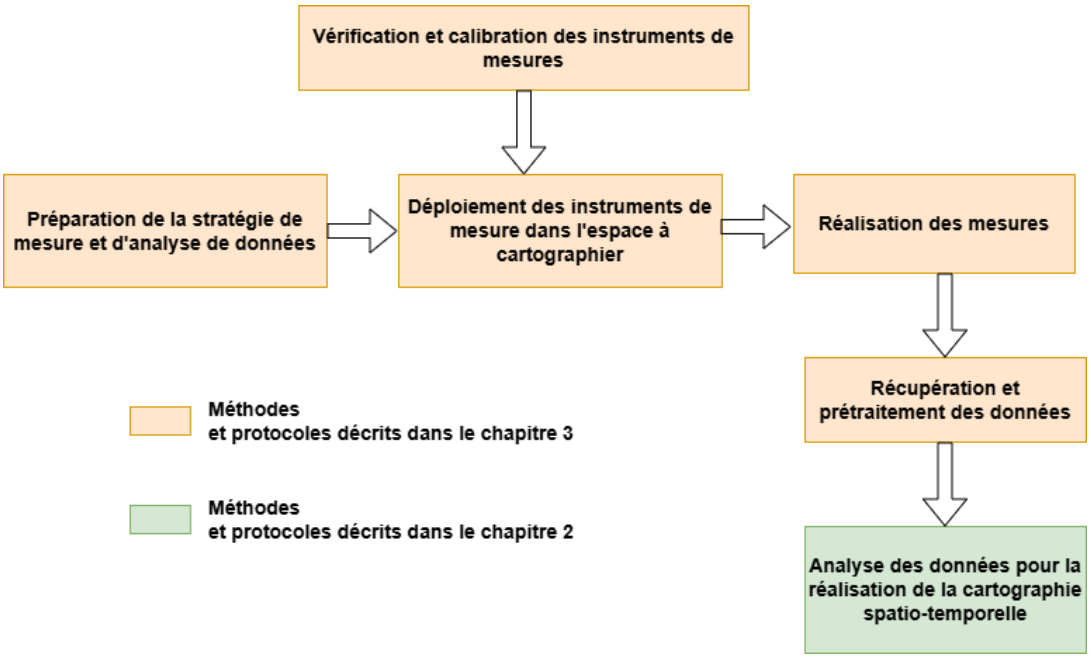


FIGURE 3.1 – Protocole de mesure et de traitement des données pour la cartographie spatio-temporelle

L'objectif de ce dernier est de collecter des données spatio-temporelles représentatives des concentrations de polluants chimiques dans un environnement intérieur maîtrisé.

Les essais reposent sur plusieurs composantes essentielles :

— La configuration de l'environnement expérimental dans lequel les essais sont réalisés,

— La nature des polluants ciblés (gazeux ou particulaires),

— Le choix des instruments de mesure,

— Les procédures de calibration associées,

— Le positionnement des instruments de mesures au sein de l'espace étudié.

Ainsi, trois séries d'essais ont été menées ; elles sont décrites dans le tableau 3.1. Ces séries ont été conçues pour répondre à des objectifs complémentaires. Les deux premières, menées dans la chambre expérimentale, visent à établir des cartographies spatio-temporelles de la concentration des polluants, respectivement gazeux et particulaires et à en étudier les caractéristiques et performances. La troisième série, réalisée dans une salle de plus grand volume aménagée à l'INRS, dans le but d'étudier la faisabilité d'estimer l'exposition individuelle à un polluant gazeux dans un environnement intérieur assimilable à un espace de travail.

La stratégie de déploiement des instruments de mesure et de la source de pollution a été adaptée pour chacun des essais, en fonction des paramètres ciblés pour l'analyse spatio-temporelle des données. Ainsi, selon les configurations retenues pour les essais, différents phénomènes de dispersion ou d'exposition ont pu être étudiés.

Série d'essais	**Type de polluant**	**Environnement expérimental**	**Nombre d'essais**	**Objectif**
Série 1	Gaz	Chambre expérimentale	25	Cartographie spatio-temporelle
Série 2	Particules	Chambre expérimentale	1	Cartographie spatio-temporelle
Série 3	Gaz	Salle aménagée	15	Reconstruction de l'exposition individuelle

Tableau 3.1 – Description des séries d'essais réalisés

3.3 Description des espaces expérimentaux

Les essais ont été menés dans deux espaces expérimentaux distincts situés dans les locaux de l'INRS à Vandoeuvre les Nancy. Le premier, appelé « Chambre expérimentale » dans la suite du manuscrit, permet de faire des essais en conditions contrôlées. Le seconde, appelé « salle aménagée » dans la suite du manuscrit, permet des faire des essais en conditions non contrôlées dans un espace de type « grand bureau ».

3.3.1 Chambre expérimentale

La chambre expérimentale utilisée pour les essais de la série 1 et 2 est une enceinte semi-hermétique d'un volume de 175 m^3 (10 m × 5 m × 3,5 m), installée à l'INRS à Vandœuvre-lès-Nancy [Chambre QAI ; Bâtiment 10]. Elle est conçue pour y mener des expériences de pollution en conditions contrôlées, tout en assurant une régulation de la température, de l'humidité relative et de la ventilation intérieure. Le dispositif de ventilation intégré permet d'ajuster divers paramètres, notamment le débit d'air entrant et sortant, ainsi que la position et le nombre des orifices d'aération. Dans ce contexte, la chambre est équipée de séries de trois orifices dédiés à l'admission et l'extraction de l'air placées en haut et en bas sur les murs latéraux (voir figure 3.2).

Pour les besoins des essais de la série 1, les orifices supérieurs ont été utilisées entièrement pour l'admission de l'air tandis que les trois autres orifices, situés en position inférieure sur le mur opposé, étaient utilisés pour extraire l'air de la chambre (orifices de ventilation en sortie) avec une ouverture à moitié de la troisième orifice. Pour les expériences, le débit de ventilation était uniformément fixé à 450 m^3/h tant pour l'admission que pour l'extraction. Quant à l'essai de la série 2, le système de ventilation de l'enceinte n'a pas été utilisé.

Le positionnement des instruments de mesure à l'intérieur de l'enceinte expérimentale a été réalisé à l'aide de trépieds photos réglables, permettant de fixer les dispositifs aux hauteurs souhaitées.

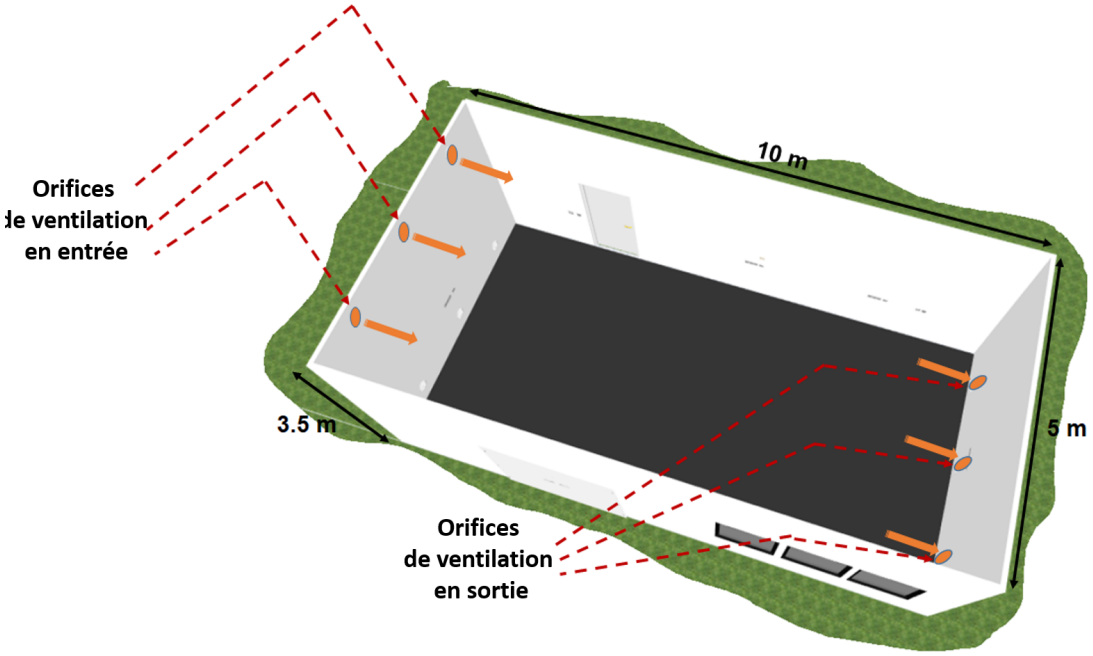


FIGURE 3.2 – Illustration de l'enceinte expérimentale

3.3.2 Salle aménagée à l'INRS

Le deuxième environnement expérimental utilisé pour les essais de la série 3 est une salle située dans les locaux de l'INRS à Vandœuvre-lès-Nancy [salle « Cabine QAI » ; Bâtiment N°10], mesurant 12,5 m × 10 m × 2,7 m de haut. La zone réellement exploitable pour les expérimentations est d'environ 12 m × 10 m × 2,7 m, une partie de l'espace étant occupée par un meuble de rangement de 0,5 m × 10 m de long, placé le long d'un mur. La salle est équipée de grandes baies vitrées latérales, situées aux extrémités gauche et droite, assurant une ouverture directe vers l'extérieur. Ces ouvertures offrent des conditions favorables à une ventilation naturelle, permettant l'évacuation efficace du polluant gazeux à l'issue de chaque essai, sans recours à un système mécanique (Voir Image 3.3). Dans un souci de sécurité, les concentrations générées sont en dessous de la VLEP du polluant utilisé, permettant une épuration sans risques. Le déploiement des instruments de mesure a été réalisé à l'aide de trépieds photos réglables, permettant un positionnement des capteurs aux hauteurs spécifiées par la stratégie de mesures



FIGURE 3.3 – Photo de la salle aménagée dans les locaux de l'INRS

3.4 Stratégie de déploiement des instruments de mesures et des sources de pollution

Comme évoqué précédemment, le CO_2 et un aérosol de particules fines ont été utilisés comme modèles de polluants atmosphérique. Les mesures ont porté soir sur le CO_2, soit sur les aérosols ainsi que sur la température et l'humidité relative de l'air. Les mesures ont pu être effectuées à points fixes

3.4.1 Stratégie de mesure à points fixes

La stratégie de mesure repose sur le positionnement, au sein de l'environnement expérimental, d'un ensemble de points de mesurage et de pollution répartis selon trois types (Voir Figure 3.4) :

— **Points de mesure avec des réseaux de capteurs :** Il s'agit des emplacements dédiés à aux modules constituant le réseau de capteurs. Pour le besoin des essais deux réseaux de capteurs ont été utilisés (un réseau de capteurs sans fil gaz (RCSFg) et un réseau de capteurs sans fil particules (RCSFp)). Ces réseaux sont décrits dans la section 3.6. La densité et la distribution des points de mesure varient selon les configurations de chaque essai. L'objectif est d'assurer une couverture adéquate de l'espace expérimental et de permettre de capter les variations spatio-temporelles de la concentration en polluant. Les données mesurées en continu à ces emplacements, permettent d'estimer les niveaux de concentrations en temps réel dans tout l'espace expérimental à l'aide de méthodes d'analyse spatio-temporelle.

— **Points de mesure avec des instruments de référence :** A ces emplacements sont positionnés des instruments de mesure possédant une meilleure précision et utilisés comme instruments de référence. Pour le besoin des essais, ils ont pu être positionnés au même endroit que les modules des réseaux de capteurs ou entre les modules (voir figure 3.4). Certains de ces appareils ont été associés à des thermo-hygromètres, mesurant simultanément la température de l'air (T, en °C) et l'humidité relative (HR, en %). Les données collectées aux points de référence sont considérées comme des valeurs de référence et ont servis à valider les méthodes d'analyse spatio-temporelle en comparant les concentrations estimées aux mêmes emplacements à partir des données du réseau de capteur. Les instruments de références sont décrits dans la section 3.6.3.

De manière générale, des mesures en continu ont été effectuées à l'ensemble de ces emplacements, portant sur les concentrations en polluants (gaz ou particules), la température de l'air (T, en °C) et l'humidité relative (HR, en %). Afin de garantir la cohérence temporelle des données collectées, les horloges de l'ensemble des instruments ont été systématiquement synchronisées en amont de chaque essai. L'espace est donc discrétisé suivant $\Delta_x = \Delta_y = 1dm$ (voir figure 3.4)

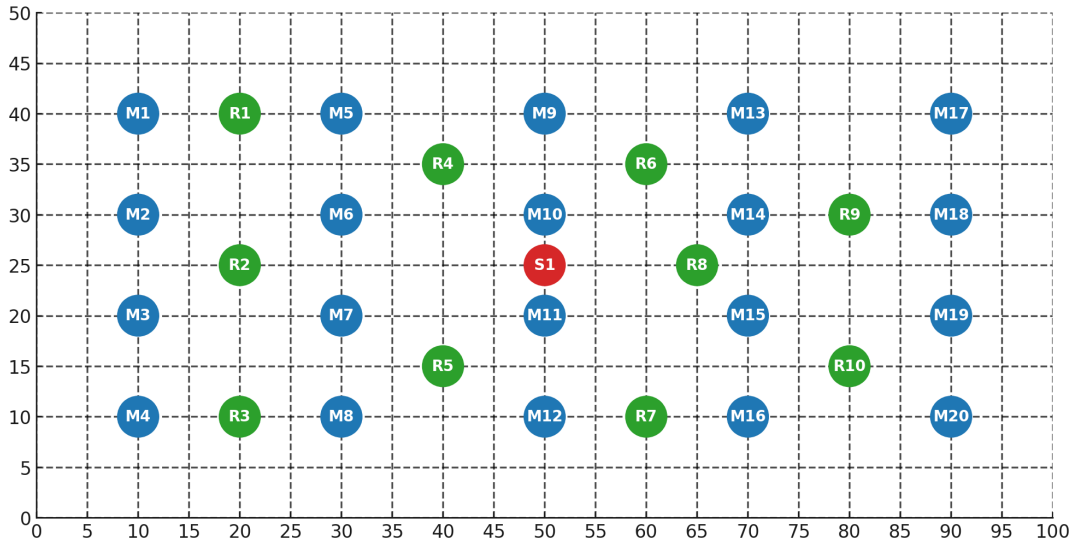


FIGURE 3.4 – Exemple d'une grille d'échantillonnage indiquant les points de mesure (M1 à M20), les points de référence (R1 à R10) et le point de pollution (S1)

et $\Delta_t = 1minute$

3.4.2 Stratégie de mesure à points mobiles

Pour les besoins d'une série d'essais, des mesures ont été effectuées à points mobiles. Dans ce cas, les dispositifs de mesures ont été positionnés sur le robot mobile SHERPA décrit dans la Section 3.6.6.

3.4.3 Stratégie de pollution

Les points de pollution correspondent aux emplacements des sources de pollution particulaires ou gazeuses dans l'environnement expérimental. Ces sources sont identifiées « S » dans la figure 3.4 et sont décrites dans la section 3.4.1. Pour les besoins des essais de une à trois sources ont pu être déployées à points fixes dans l'environnement expérimental.

3.4.4 Déroulement des essais

Chaque essai a nécessité l'utilisation d'une grille d'échantillonnage associée à un plan d'expérience. Tous les essais ont été menés sur le schéma de la séquence des phases suivante :

— **Phase 1 : Phase de pollution**, Cette première phase consiste à générer le polluant (gazeux ou particulaire) dans l'enceinte expérimentale à partir d'une ou plusieurs sources localisées.

— **Phase 2 : Phase de stabilisation**, Durant cette phase, la génération du polluant est arrêtée. Ce qui permet d'observer sa redistribution spontanée au sein de l'environnement expérimental. Cette période transitoire est déterminante pour caractériser l'état

d'équilibre temporaire des concentrations, préalablement à la phase d'évacuation.

— **Phase 3 : Phase d'épuration**, Cette dernière phase est consacrée à l'élimination progressive du polluant présent dans l'air de l'environnement expérimental grâce à un système de ventilation. En fonction de la configuration de l'essai, le système de ventilation a pu être mécanique (système de ventilation de l'enceinte expérimentale avec les débits d'entrée et sortie d'air), manuelle (à base de ventilateurs ou épurateurs classiques posés à l'intérieur de l'environnement expérimental) ou naturelle (à base de l'ouverture des fenêtres).

Pour la réalisation des différentes séries d'essais, deux plans d'expérience spécifiques ont été appliqués (Plan A et Plan B).

Plan A : on réalise l'essai **sans aucun système de ventilation ni brassage forcé** (pas de ventilateur, pas de VMC, pas de soufflage) pendant la phase de pollution ; le système de ventilation est utilisé essentiellement à la phase d'épuration.

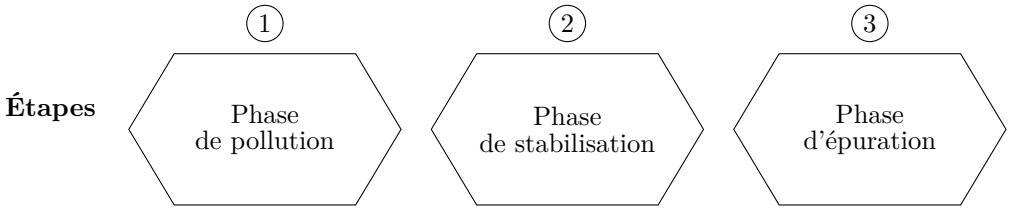


FIGURE 3.5 – Représentation du plan d'expérience A avec les différentes phases chronologiques

Plan B : on réalise l'essai **avec le fonctionnement d'un système de ventilation** en continu. Dans un environnement contrôlé tel que la chambre expérimentale de l'INRS (figure 3.2), où l'ensemble des paramètres de l'essai peuvent être configurés, il est possible de reproduire ce plan expérimental à l'aide d'une simulation en dynamique des fluides numérique (CFD - « Computational Fluid Dynamics »). Cette simulation pourrait alors constituer une référence utile pour évaluer la cohérence physico-chimique des données collectées par les instruments de mesure.

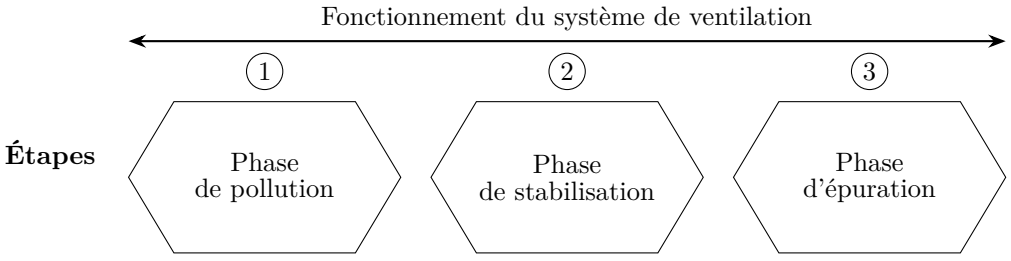


FIGURE 3.6 – Représentation des phases chronologiques du plan d'expérience B

Conformément aux protocoles définis précédemment, l'ensemble des essais menés dans le cadre de la série 1 visait à évaluer les performances du dispositif de cartographie spatio-

temporelle en fonction de divers paramètres expérimentaux. Le protocole a été décliné selon plusieurs configurations, en faisant varier la grille d'échantillonnage, le nombre de sources, leur orientation, ainsi que le débit de génération du polluant gazeux (CO_2). Le tableau 3.2 synthétise les essais réalisés, en précisant pour chacun :

— La nature de la grille spatiale (régulière ou irrégulière),
— Le nombre et l'orientation des sources de pollution,
— Les débits de génération appliqués,
— Les plans de calibration et d'expérience associés,
— L'objectif spécifique visé.

TABLEAU 3.2 – Résumé des essais de la série 1

N° Essai	**Nombre d'essais**	**Grille d'échantillonnage**	**Nombre de sources**	**Débit de génération**	**Plan de calibration et d'expérience**	**Objectif**
S1-E1	1	Régulière	1 orientée vers le bas	12,5L/min	Calibration A Expérience A	1
S1-E2	1	Régulière	1 orientée vers le haut	12,5L/min	Calibration A Expérience A	1
S1-E3	1	Régulière	1 orientée vers le haut	12,5L/min par source	Calibration A Expérience A	2
S1-E4	1	Régulière	1 orientée vers le haut	20L/min par source	Calibration A Expérience A	2
S1-E5	1	Régulière	1 orientée vers le haut	45L/min par source	Calibration A Expérience A	2
S1-E6	1	Régulière	2 orientées vers le haut	30L/min par source	Calibration A Expérience A	3
S1-E7	1	Régulière	3 orientées vers le haut	20L/min par source	Calibration A Expérience A	3
S1-E8 S1-E9	2	Régulière	1 orientée vers le haut	12,5L/min	Calibration B Expérience B	4 5
S1-E10	1	Irrégulière 1	1 orientée vers le haut	12,5L/min	Calibration B Expérience B	5
S1-E11	1	Irrégulière 2	1 orientée vers le haut	12,5L/min	Calibration B Expérience B	5
S1-E12 S1-E13	2	Irrégulière 3	1 orientée vers le haut	12,5L/min	Calibration B Expérience B	5
S1-E14	1	Régulière (à 50 cm)	1 orientée vers le haut	12,5L/min	Calibration B Expérience B	6
S1-E15	1	Régulière (à 100 cm)	1 orientée vers le haut	12,5L/min	Calibration B Expérience B	6
S1-E16	1	Régulière (à 160 cm)	1 orientée vers le haut	12,5L/min	Calibration B Expérience B	6

1 : Influence du sens de le source sur les cartographies
2 : Influence du débit de génération sur les cartographies
3 : Influence du nombre de sources sur les cartographies
4 : Répétabilité des cartographies
5 : Influence de la grille d'échantillonnage sur les cartographies
6 : Influence de la hauteur de la source par rapport aux mesures sur les cartographies

Quant à la série 2, elle est constituée d'un essai, ayant pour objectif d'évaluer la faisabilité du dispositif expérimental pour la cartographie de l'exposition aux particules fines. Cet essai a pour objectif de valider les conditions techniques et méthodologiques requises pour l'estimation spatio-temporelle des concentrations particulaires dans l'environnement expérimental.

Tableau 3.3 – Résumé de l'essai de la série 2

N° Essai	**Nombre d'essais**	**Grille d'échantillonnage**	**Nombre de sources**	**Débit de diffusion**	**Plan de calibration et d'expérience**	**Objectif**
S2-E1	1	Régulière	1 orientée vers le haut	2 bars	Calibration Particules Expérience B	Essai de faisabilité

La série 3 est dédiée à l'évaluation de la fusion de données hétérogènes(capteurs, position au cours d'un déplacement individuel) dans le cadre de la reconstruction de l'exposition individuelle à un polluant gazeux (CO_2). L'objectif est d'estimer, à partir des données spatio-temporelles mesurées dans l'environnement expérimental, la concentration cumulée de CO_2 à laquelle un individu serait exposé en fonction de sa trajectoire. Deux configurations distinctes ont été mises en œuvre, différenciées par la position et le nombre de sources d'émission. Cette variation permet d'évaluer l'influence de la distribution des sources de pollution sur la qualité de la reconstruction de l'exposition, ainsi que la performance de l'approche de fusion de données appliquée.

Tableau 3.4 – Résumé de l'ensemble des essais de la série 3

N° Essai	**Nombre d'essais**	**Grille d'échantillonnage**	**Nombre de sources**	**Débit de génération**	**Plan de calibration et d'expérience**	**Objectif**
S3-E1 à S3-E10	10	Régulière	1 orientée vers le haut	20L/min	Calibration B Expérience B	1
S3-E11 à S3-E15	5	Régulière	3 orientées vers le haut (en diffusion alternée)	20L/min	Calibration B Expérience B	2

1 : Fusion de données et reconstruction de l'exposition individuelle
2 : Influence de la distribution des sources sur la reconstruction de l'exposition individuelle

3.5 Description des sources de pollution

3.5.1 Source de pollution gaz

Le gaz utilisé comme polluant au cours des séries d'essais est le dioxyde de carbone CO_2. Le CO_2 est présent naturellement dans l'air et présente une toxicité relative (VLEP = 5000 ppm/v). Facilement disponible, son épuration ne présente pas de risque pour l'environnement et le processus de propagation peut être réversible sans difficulté. Dans l'enceinte expérimentale, sa génération a été réalisée à partir de bouteilles de CO_2 liquide de haute pureté (B20, 15 kg, Air Liquide, France – qualité : $[H_2O] < 3$ ppm, $[O_2]$, $[CnHm] < 2$ ppm, $[H_2] < 0{,}5$ ppm, $[N_2] < 8$ ppm) (Voir Figure 3.7). La bouteille de CO_2 est reliée à un tuyau en nylon de 10 mm de diamètre interne, inséré par le plafond pour diriger le gaz vers le point source. Ce dispositif assure une injection continue de dioxyde de carbone à un débit défini, grâce à un régulateur de débit massique (MFR, Brooks SLA2800 Series avec contrôleur et afficheur Brooks 0254, USA) d'une capacité de 20 ou 200 L/min (voir figure 3.7).



FIGURE 3.7 – Bouteille de CO2 (à gauche) et régulateur de débit massique [RDM] associé (à droite))

Ce processus a été privilégié pour un contrôle optimal des débits d'émission par rapport à d'autres alternatives telles que l'utilisation de carboglace. Pour la gestion du CO_2 l'installation de la bouteille de CO_2 a été effectuée en extérieur dans un cadre sécurisé car la pression en sortie de la bonbonne est d'environ 50 bar. Cette installation nous permettait de faire passer la pression de 50 à 4 bar et favorise une purge rapide du circuit afin de ne pas le laisser sous pression en continu.

3.5.2 Source de pollution particules

Dans le cadre de l'essai réalisé pour les mesures de particules fines, la source de pollution particulaire retenue est un générateur d'aérosols salins Palas AGK 2000 (5 nm à 15 μ m). Cet appareil nébulise une solution aqueuse de chlorure de sodium à 1 % (m/m) puis élimine l'excédent d'eau pour délivrer un nuage stable de particules fines solides typiquement de 70 à

100 nm de diamètre modal (voir figure 3.8).

Le dispositif offre un réglage précis du débit d'aérosol et de la concentration massique grâce à un système de buses ultrasoniques couplé à un sécheur thermique, garantissant ainsi une remarquable reproductibilité des émissions, condition sine qua non pour les mesures métrologiques. L'emploi d'un aérosol salin présente plusieurs atouts :

— Neutralité chimique et innocuité : le NaCl, non toxique aux concentrations envisagées, limite les contraintes de sécurité.

— Hygroscopicité contrôlable : les particules réagissent rapidement aux variations d'humidité relative, permettant d'étudier l'effet de la vapeur d'eau sur leur comportement.

— Traçabilité métrologique : la granulométrie et la concentration sont aisément vérifiables à l'aide d'instruments de référence (SMPS, CPC), ce qui assure la qualité des données acquises.

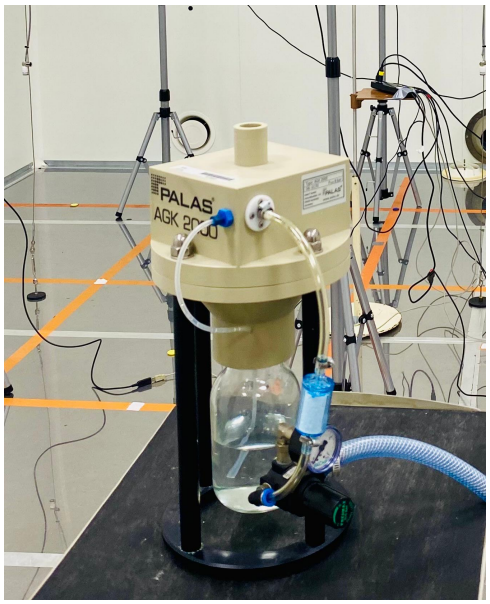


FIGURE 3.8 – Générateur de particules salines (NaCl [1%] (Palas AGK 2000))

Le recours au générateur d'aérosols salins Palas AGK 2000, alimenté par une solution aqueuse de NaCl à 1%, assure l'émission d'un aérosol salin monodisperse et stable, aisément caractérisable et adapté aux analyses spatio-temporelles exposées dans le présent manuscrit.

3.6 Méthodes et instruments de mesures

3.6.1 Réseau de capteurs sans fil gaz (RCSFg)

3.6.1.1 RCSFg conçu par l'INRS

Pour les essais de la série 1, le réseau de capteurs déployé est un système sans fil développé par l'INRS. Il comprend 20 modules autonomes, chacun relié à une station de base assurant la

transmission centralisée des données. Chaque module intègre une carte Arduino, alimentée par une batterie externe, ce qui garantit une autonomie suffisante pour les campagnes de mesure. L'Arduino est complété par deux cartes Protoshield facilitant l'interfaçage avec différents capteurs, ainsi que par une carte Xbee assurant la communication sans fil via le protocole Xbee (voir Figure 3.9). Les modules sont équipés d'un capteur infrarouge de CO_2 (C1 low power CO_2 Sensor, Eurogas LTD, Belgique), capable de mesurer des concentrations allant de 0 à 10000 ppm/v avec une précision de 50 ppm/v, dans une plage de température comprise entre 0 et 50 °C, et une humidité relative de 0 à 95 %. Un capteur combiné (Grove – Temperature and Humidity Sensor Pro DHT22/AM2302, Seeed, Chine) est également intégré pour la mesure conjointe de la température (−40 à +80 °C (±0,5 °C)) et de l'humidité relative (20 à 100 % (±2 %)).

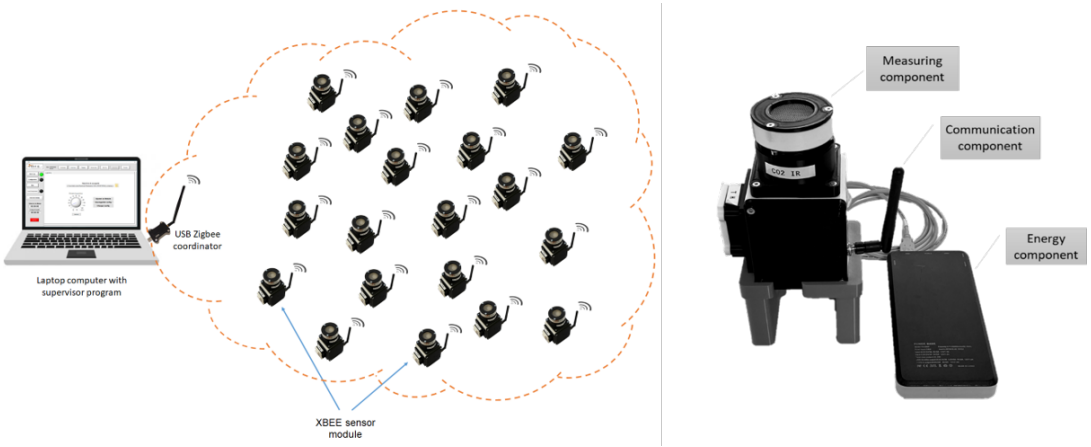


FIGURE 3.9 – Illustration du réseau de capteurs sans fil gaz utilisé pour les essais (à gauche) et d'un module constituant le réseau de capteurs avec une batterie portable pour son alimentation(à droite)

3.6.1.2 RCSFg commercial « Aranet »

Dans le cadre des essais de la série 3, le réseau de capteurs utilisé est le réseau « Aranet ». C'est un réseau de capteurs sans fil conçu pour la surveillance en temps réel des paramètres environnementaux, en particulier les concentrations en dioxyde de carbone CO_2 (0-10000ppm/v).

Il repose sur une architecture centralisée, où plusieurs capteurs communiquent avec une passerelle via une liaison radio de faible consommation (868 MHz), permettant une couverture jusqu'à 3 km en champ libre. Chaque capteur Aranet est équipé d'un capteur NDIR (Non-Dispersive Infrared) de haute précision pour la mesure du CO_2, ainsi que de capteurs internes pour la température de l'air (°C), l'humidité relative (% HR) et la pression atmosphérique (hPa). Les données sont enregistrées localement

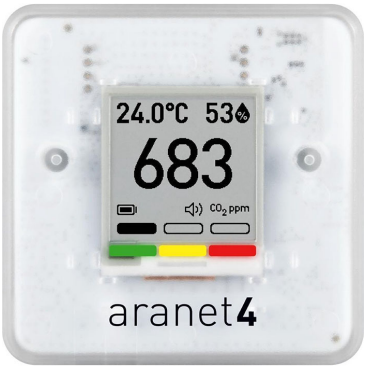


FIGURE 3.10 – Module capteur constituant le réseau Aranet 4

et transmises à la passerelle à une fréquence configurable (généralement entre 1 et 10 minutes), puis visualisées via une interface web ou une application dédiée. Les capteurs fonctionnent sur batterie, avec une autonomie pouvant atteindre plusieurs mois selon la fréquence de transmission, ce qui facilite leur déploiement dans des environnements variés sans infrastructure filaire.

3.6.2 Réseau filaire de capteurs particules (RFCp)

Le réseau filaire de capteurs utilisé est composé de 20 capteurs particules SPS30 Sensirion (0 à 1000 $\mu g/m^3$) installés en fonction de la grille d'échantillonnage sur les points de mesures, ils sont reliés via des cables USB à 2 hubs de 10 connexions. Chaque hub est relié à un PC de contrôle sur lequel est installé le soft ControlCenter V1.40.2. avec chacun une session pilotant 10 capteurs. Cette configuration a été retenue parce que la version du logiciel utilisée impose une limite de 19 capteurs par session.

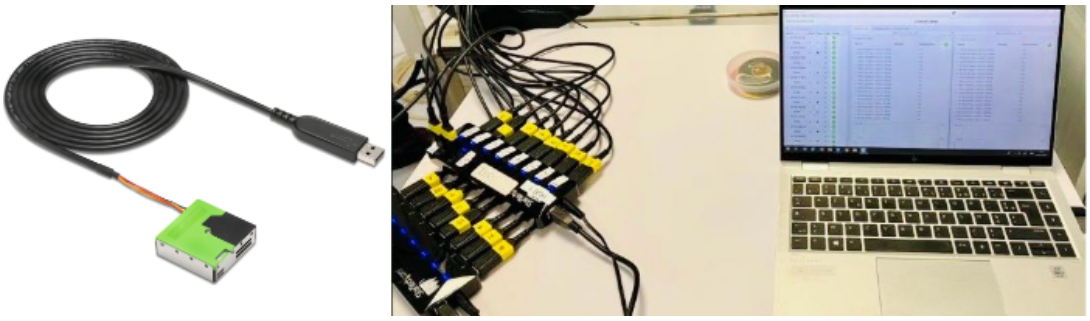


FIGURE 3.11 – Capteur SPS30 Sensirion avec le dispositif de connexion filaire

3.6.3 Détecteurs de CO_2 portables = instruments de référence

Pour les séries d'essais portant sur un polluant gazeux, un ensemble de détecteurs portables de CO_2 (ToxiRAE Pro CO_2, RAE Systems France) a été déployé en complément du réseau de capteurs sans fil. Ces dispositifs, présentés en Figure 3.12, ont servi de référence pour l'évaluation et la validation des mesures produites par le réseau. Le ToxiRAE Pro CO_2 est un détecteur personnel de dioxyde de carbone en temps réel, doté d'une fonction d'enregistrement des données. Son principe de mesure repose sur une relation proportionnelle entre l'intensité du signal fourni par le capteur infrarouge et la concentration de CO_2 présente dans l'espace de mesure. Le capteur permet de mesurer des concentrations de CO_2 comprises entre 0 et 50 000 ppmv, avec une résolution de 100 ppmv, et fonctionne dans une plage de températures allant de –20 à 55 °C.



FIGURE 3.12 – Détecteur de CO_2 portable (ToxiRAE Pro CO_2)

3.6.4 Thermo-hygromètre



FIGURE 3.13 – Thermo-hygromètre (Testo 174H)

Au sein des environnements expérimentaux, lors des différentes séries d'essais, la température de l'air (en °C) et l'humidité relative (en %) ont été enregistrées à l'aide de thermo-hygromètres de type Testo 174H – Mini Data Logger (Testo, France). Cet appareil est conçu pour mesurer des températures comprises entre –20 °C et 70 °C, ainsi qu'une humidité relative allant de 0 à 100 %. Les données enregistrées constituent une référence pour l'évaluation des mesures de température et d'humidité fournies par les modules du réseau de capteurs.

3.6.5 Scanner 3D



FIGURE 3.14 – RTC360 LT Laser scanner 3D (Leica Geosystems)

Le RTC360 LT de Leica Geosystems est un scanner 3D terrestre à technologie laser, conçu pour la capture automatisée d'environnements complexes en trois dimensions. Il permet de générer des nuages de points à haute densité, utiles pour la modélisation géométrique d'espaces intérieurs ou extérieurs. Doté d'une portée de mesure allant jusqu'à 130 mètres, le RTC360 LT offre une précision millimétrique (jusqu'à ± 1,5 mm à 10 m), ce qui le rend adapté aux applications nécessitant une restitution géométrique fine, comme la modélisation d'environnements expérimentaux, l'étude des volumes ou encore l'optimisation du positionnement d'instruments de mesure.

L'appareil est équipé de caméras panoramiques haute résolution permettant de capturer des images couleur, qui sont automatiquement associées aux données du nuage de points. Le RTC360 LT est piloté via une interface mobile intuitive (application Leica Cyclone FIELD 360) et les données sont compatibles avec les logiciels de traitement spécialisés comme Leica Cyclone REGISTER 360 ou d'autres plateformes BIM et CAO.

Dans le cadre de cette étude, ce scanner a été utilisé pour modéliser en 3D les environnements expérimentaux. Les données de nuages de points acquises permettent de reproduire fidèlement la géométrie des environnements d'essai, en vue de faciliter la fusion de données multisources (concentrations en polluants gazeux à diverses positions) et la cartographie spatio-temporelle des concentrations de polluants, en lien avec la configuration réelle de l'espace expérimental.

3.6.6 Robot mobile



FIGURE 3.15 – Robot autonome mobile SHERPA

Le robot mobile utilisé dans le cadre des essais de la série 3 est un système autonome de type AMR (Autonomous Mobile Robot) développé par Sherpa Mobile Robotics. Conçu pour évoluer dans des environnements industriels ou semi-industriels, ce robot est spécifiquement dédié aux applications de logistique interne. Il intègre un convoyeur motorisé embarqué, permettant le transport automatisé de charges légères entre plusieurs points d'un site de production : postes de travail, zones de stockage ou interfaces de ligne. Grâce à ses capacités de navigation autonome et à son module de transfert intégré, il assure une circulation fluide et sécurisée des flux matériels sans intervention humaine.

Le châssis du robot est compact et ergonomique, conçu pour évoluer dans des environnements contraints tout en garantissant une stabilité de déplacement. Le robot Sherpa repose sur une technologie de navigation autonome par SLAM (Simultaneous Localization and Mapping), couplée à un ensemble de capteurs embarqués (LIDAR, caméras, capteurs de proximité). Cette combinaison lui permet de se localiser dynamiquement dans un espace structuré ou semi-structuré ; de cartographier son environnement en temps réel ; d'éviter les obstacles fixes ou mobiles (opérateurs, équipements) et de planifier ses trajectoires de manière sûre et optimisée. Il dispose de bandes lumineuses LED bleues en façade qui indiquent visuellement l'état du robot (en attente, en déplacement, en intervention) et contribuent à la communication homme-machine dans un cadre collaboratif.

3.6.7 Station météo

La station météorologique BWS200 est un dispositif multifonction destiné à la mesure des paramètres météorologiques en environnement extérieur. Elle est équipée de capteurs intégrés permettant de mesurer, entre autres, la température de l'air, l'humidité relative, la vitesse et la direction du vent, ainsi que la pression atmosphérique et les niveaux de rayonnement solaire. Dans le cadre de cette étude, la station BWS200 a été installée à l'extérieur de la salle expérimentale durant les essais de la série 3 réalisés en période estivale. Son objectif principal était de documenter les conditions climatiques ambiantes, susceptibles d'influencer les échanges thermiques et la dynamique des polluants à l'intérieur de la salle expérimentale aménagée à l'INRS, notamment dans les situations de ventilation naturelle (ex. ouverture de fenêtres).

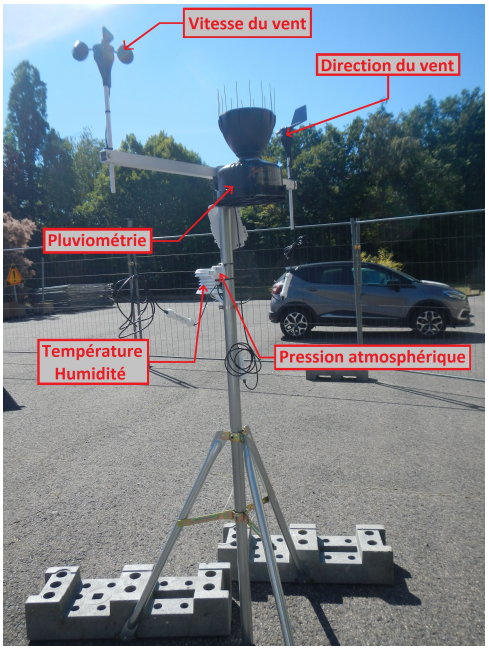


FIGURE 3.16 – Station météorologique BWS200

3.7 Vérification et étalonnage des instruments de mesures

3.7.1 Étalonnage des instruments de mesures gaz

Le protocole de calibration consiste à exposer les instruments de mesures à une série de concentrations de CO_2 comprise entre 400 et 10 000 ppmv. Cette gamme a été choisie pour correspondre tant à la plage opérationnelle prédéfinie qu'à la capacité de mesure des appareils, tout en assurant l'enregistrement des données. Le procédé s'est déroulé dans une enceinte scellée de 530 L conçue par l'INRS et équipée d'un thermo-hygromètre, d'un manomètre ainsi que d'un ventilateur permettant d'homogénéiser la distribution du CO_2 dans le volume. Cette installation permettait la calibration simultanée de 20 dispositifs (Voir figure 3.17).

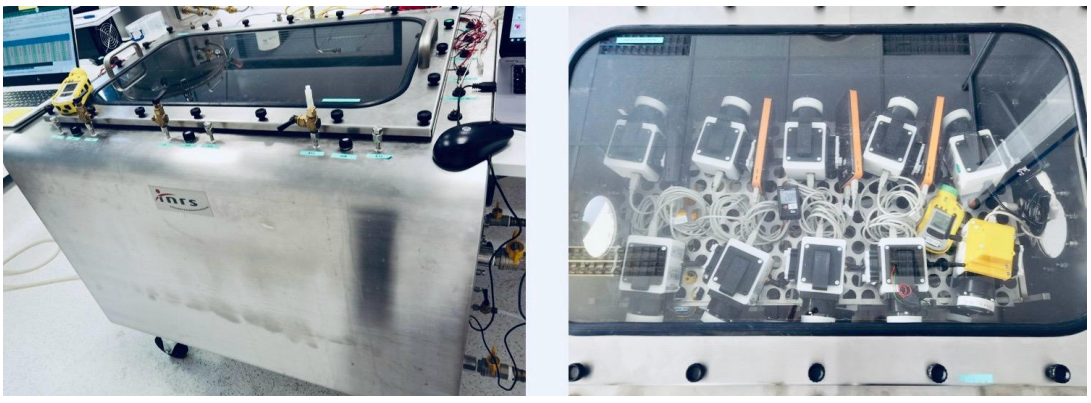


FIGURE 3.17 – Disposition des instruments de mesure dans l'enceinte de calibration

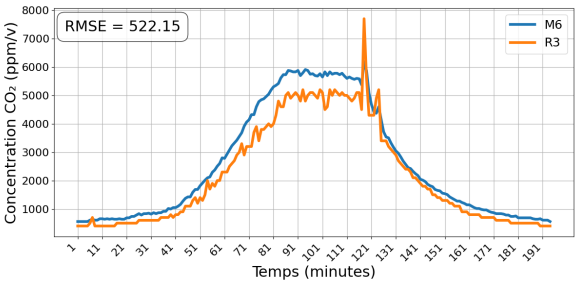
Ce protocole est compatible à la fois avec les modules du RCSFg ainsi qu'aux capteurs du réseau Aranet associés aux détecteurs de CO_2 pour la mesure du CO_2.

Le protocole expérimental s'est déroulé en plusieurs séquences successives, chacune associant un ensemble de 10 capteurs du réseau aux mêmes 10 détecteurs de CO_2. Pour chaque séquence, un gaz étalon de CO_2, caractérisé par une concentration définie, était injecté afin d'atteindre une valeur de référence spécifique.

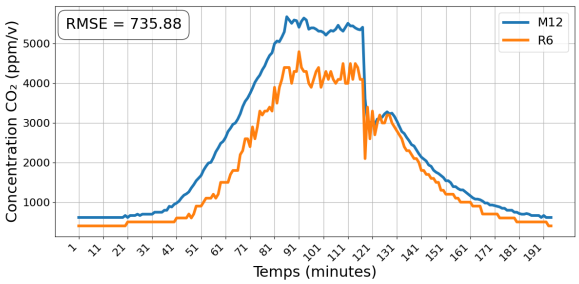
Après avoir maintenu ce niveau de concentration pendant quelques minutes, l'enceinte était ramenée à la concentration atmosphérique de CO_2. Ce cycle était répété de manière successive avec une élévation progressive des concentrations de 400 à 10 000 ppmv. Les coefficients de proportionnalité, représentant l'intensité du signal en fonction de la concentration de CO_2, ont été intégrés au logiciel de gestion pour les modules du RCSFg.

Pour les capteurs du réseau Aranet et les détecteurs de CO_2, des coefficients de correction ont été déterminés en comparant les concentrations mesurées aux concentrations de référence auxquelles les dispositifs avaient été initialement exposés, puis utilisés afin de corriger les mesures effectuées. Ce protocole d'étalonnage sera désigné dans la suite du manuscrit sous l'appellation « plan de calibration B ». Au préalable, un premier plan, intitulé « plan de calibration A », avait été mis en œuvre pour les essais initiaux de la série 1 ; il reposait essentiellement sur l'exploitation de coefficients déterminés par la comparaison des concentrations de CO_2 mesurées par les modules du RCSFg regroupés par groupes de dix avec les valeurs de référence fournies par deux détecteurs de CO_2, chacun dédié à son palier.

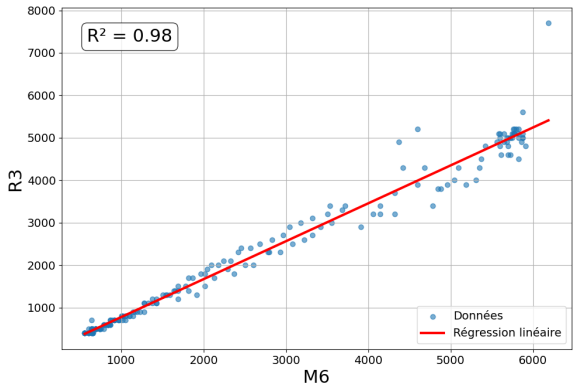
Pour chacun des plans de calibration, deux essais de faisabilité ont été réalisés, au cours desquels les modules du RCSFg ont été couplés aux détecteurs de CO_2.



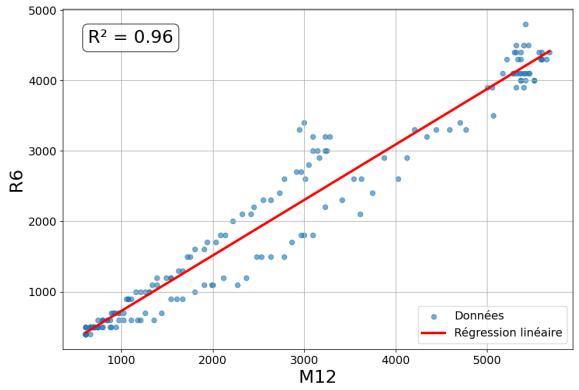
(a) Évolution temporelle de la concentration en CO_2 mesurée par M6 et R3



(b) Évolution temporelle de la concentration en CO_2 mesurée par M12 et R6

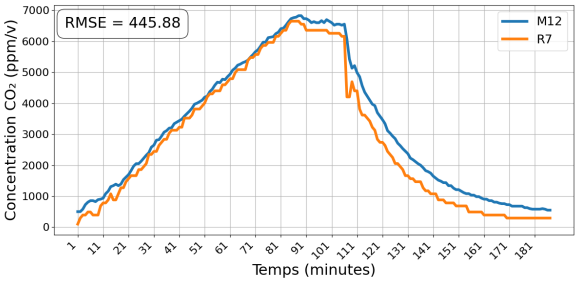


(c) Corrélation entre M6 et R3

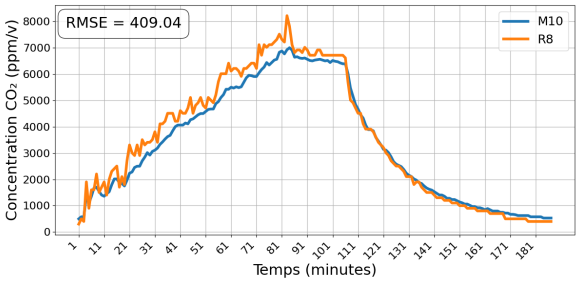


(d) Corrélation entre M12 et R6

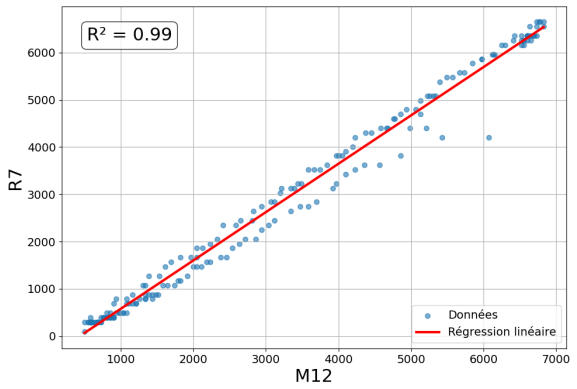
FIGURE 3.18 – Analyse comparative des signaux enregistrés par les modules de capteurs et les détecteurs de CO_2 étalonnés suivant le plan de calibration A



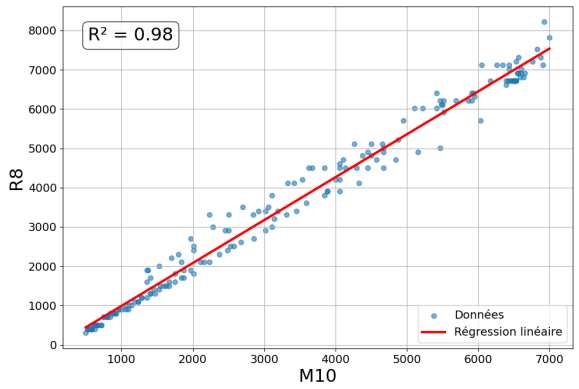
(a) Évolution temporelle de la concentration en CO_2 mesurée par M12 et R7



(b) Évolution temporelle de la concentration en CO_2 mesurée par M10 et R8

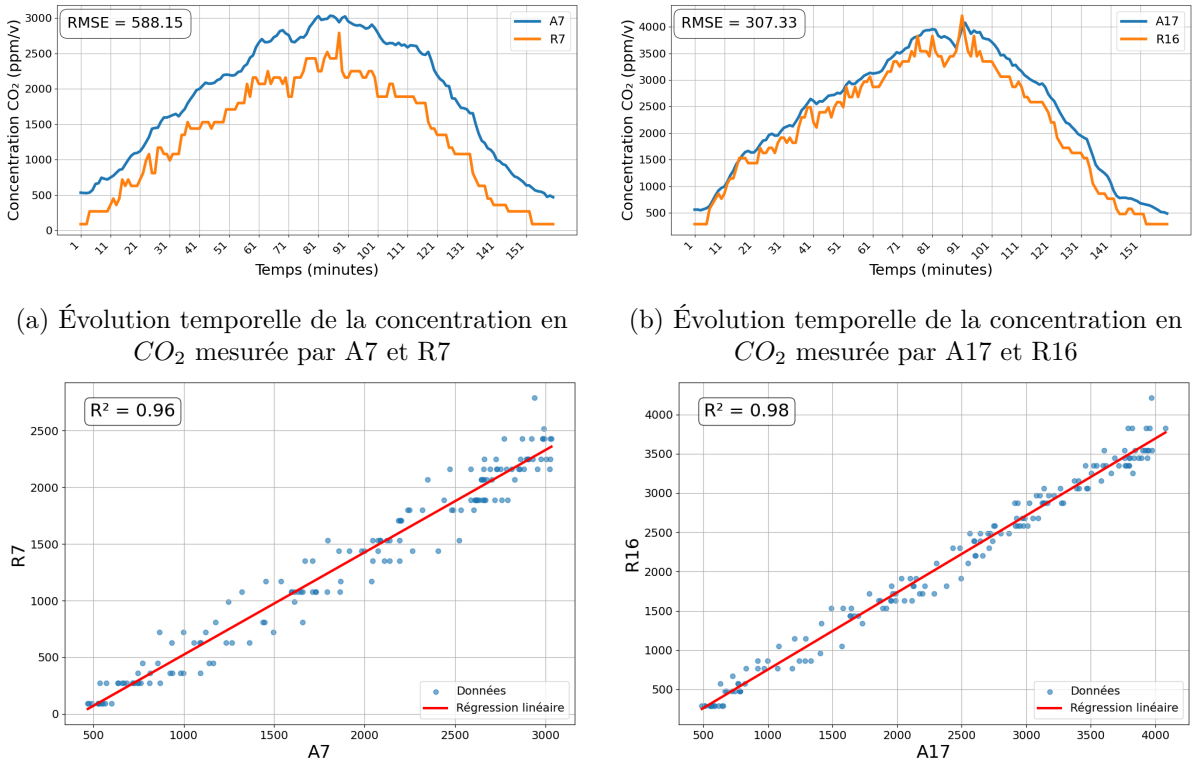


(c) Corrélation entre M12 et R7



(d) Corrélation entre M10 et R8

FIGURE 3.19 – Analyse comparative des signaux enregistrés par les modules de capteurs et les détecteurs de CO_2 étalonnés suivant le plan de calibration B



(a) Évolution temporelle de la concentration en CO_2 mesurée par A7 et R7

(b) Évolution temporelle de la concentration en CO_2 mesurée par A17 et R16

(c) Corrélation entre A7 et R7

(d) Corrélation entre A17 et R16

FIGURE 3.20 – Analyse comparative des signaux enregistrés par les capteurs du réseau Aranet et les détecteurs de CO_2 étalonnés suivant le plan de calibration B

Cette analyse comparative vise à évaluer la précision des mesures fournies par les capteurs et les appareils de référence. Elle permet également de déterminer la marge d'erreur à prendre en compte lors de l'analyse spatio-temporelle des données, en vue de leur représentation cartographique.

3.7.2 Étalonnage du réseau filaire de capteurs particules

Afin d'évaluer la dispersion et l'exactitude des données mesurées par les 20 capteurs SPS30 constituant le RFCp, les 20 capteurs SPS30 Sensirion utilisés pour l'essai de faisabilité ont été étalonnés dans une enceinte expérimentale de 1 m³. Regroupés en grappe, ils ont été positionnés approximativement au centre de l'enceinte afin de permettre une exposition homogène au flux de particules fines. L'enceinte est polluée en particules fines par un générateur de particules salines Palas AGK 2000 (solution de NaCl à 1 %), installé à la base de la face avant.
Le protocole d'étalonnage repose sur l'utilisation d'un dispositif de référence chargé de quantifier la concentration de particules fines à l'intérieur de l'enceinte ; les valeurs obtenues servent ensuite de base de comparaison pour les mesures effectuées par les capteurs SPS30. L'instrument de référence utilisé est le DustTrak DRX Aerosol Monitor 8533 (TSI), il est placé à l'extérieur de l'enceinte et aspire le flux interne de particules fines au moyen d'une tubulure raccordée à

l'angle inférieur droit de la paroi arrière (Voir figure 3.21).



FIGURE 3.21 – Enceinte de calibration équipée des capteurs SPS30, du générateur et de l'instrument de référence

Le protocole s'est déroulé en une phase de génération de solution saline [1%] à une pression de 1 bar avec une ventilation de l'enceinte pour l'homogénéisation du flux. Les mesures ont été enregistrées en continu durant une séquence de dix minutes, avec un pas d'acquisition d'une seconde.

Pour chaque capteur, les coefficients de proportionnalité ont été établis en rapportant les valeurs fournies par l'instrument de référence à celles enregistrées par le capteur ; ces coefficients ont ensuite été appliqués pour corriger l'ensemble des données recueillies durant l'essai de faisabilité.

TABLEAU 3.5 – Coefficients d'étalonnage des capteurs SPS30

Capteur	**Coefficient**	**Capteur**	**Coefficient**
C1	0.941	C11	1.145
C2	0.930	C12	1.128
C3	0.893	C13	1.017
C4	1.023	C14	0.914
C5	0.915	C15	1.100
C6	0.949	C16	1.145
C7	0.911	C17	1.040
C8	0.970	C18	1.076
C9	0.900	C19	1.031
C10	0.955	C20	1.018

3.8 Pré-traitement des données

Avant leur exploitation, les données collectées lors des séries d'essais font l'objet d'un pré-traitement. Cette étape est indispensable pour assurer la synchronisation temporelle et la bonne structuration des données, en vue de leur utilisation pour l'analyse spatio-temporelle. La Figure 3.22 en présente les principales étapes.

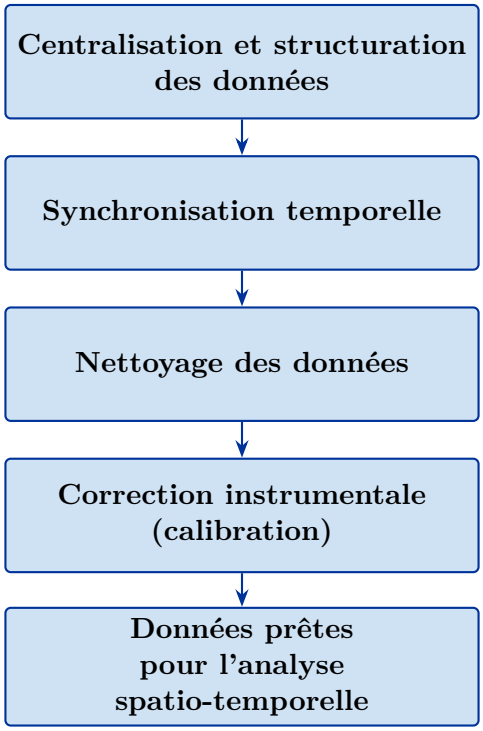


FIGURE 3.22 – Étapes du pré-traitement des données avant exploitation

— **Centralisation et structuration** : Les données issues des différents dispositifs (modules capteurs, détecteurs de référence, thermo-hygromètres, etc.) sont centralisées à partir de leurs interfaces respectives (via export en fichier .csv). Elles sont ensuite formatées selon un référentiel spatio-temporel (position,temps,valeur).

— **Synchronisation temporelle** : Afin d'assurer la compatibilité des données multicapteurs, les horodatages sont vérifiés et ajustés si nécessaire. Cette synchronisation permet d'aligner les mesures sur une échelle de temps commune, indispensable pour l'analyse.

— **Nettoyage des données** : Les données anormales, telles que les doublons, les sauts de mesure, les valeurs aberrantes(hors plage, vides ou invalides) sont identifiées puis corrigées à l'aide d'une méthode de lissage par moyenne glissante, permettant une reconstruction locale cohérente du signal.

— **Correction instrumentale (calibration)** :Des coefficients de proportionnalité ou d'étalonnage (déterminés en phase de calibration) sont appliqués aux mesures brutes pour corriger les biais instrumentaux et harmoniser les données entre capteurs de nature différente.

— **Données prêtes pour l'analyse spatio-temporelle :** Les données sont formatées pour permettre leur exploitation analytique. Les mesures issues du réseau de capteurs constituent les données d'entrée du modèle de cartographie spatio-temporelle, et celles provenant des instruments de référence sont utilisées comme données de test et de validation pour l'évaluation des performances du modèle. Les séries temporelles sont ré-échantillonnées de façon régulière pour assurer la synchronisation des données et leur intégration dans les phases de fusion et de modélisation spatio-temporelle.

Chapitre 4

Exploitation des données

Sommaire

4.1 Introduction 76

4.2 Application des méthodes de cartographie spatio-temporelle aux données réelles : choix des paramètres 76

4.3 Étude des caractéristiques de la cartographie spatio-temporelle . . 78

4.3.1 Influence de la méthode de cartographie spatio-temporelle 78

4.3.2 Influence de la grille de positionnement des points de mesure 94

4.3.3 Influence du nombre de points de mesure 101

4.3.4 Influence des caractéristiques de la source de pollution 109

4.3.5 Influence du débit d'émission 117

4.3.6 Influence du nombre de sources 123

4.3.7 Influence du type de polluant : essais avec un aérosol 127

4.4 Fusion de données : Représentation 3D de la chambre expérimentale avec les niveaux de concentrations 137

4.5 Fusion de données : Reconstruction des expositions individuelles . 140

4.5.1 Objectifs et principes des essais 140

4.5.2 Méthodologie 141

4.5.3 Résultats et discussions 144

4.6 Applications en santé au travail : retombées indirectes des travaux de thèse 162

4.6.1 Apport sur les stratégies de mesure 163

4.6.2 Périmètre de « validité » autour d'un point de mesure 165

4.6.3 Identification des zones de danger 169

4.6.4 Conclusion partielle 170

4.1 Introduction

Ce chapitre est consacré à l'exploitation des données collectées au cours des séries d'essais présentées dans le chapitre précédent. Il s'articule autour de l'analyse des performances des méthodes d'analyse spatio-temporelle appliquées aux concentrations en polluants gazeux et particulaires. L'analyse s'organise suivant deux axes complémentaires : d'une part, la réalisation de cartographies spatio-temporelles à partir des mesures collectées, et d'autre part, la reconstruction de l'exposition individuelle sur la base de ces estimations, en fusion avec des trajectoires de déplacement dans l'environnement étudié.

La première partie du chapitre est dédiée à l'étude des caractéristiques de la cartographie spatio-temporelle. Elle vise à évaluer la robustesse et la fiabilité des estimations produites, à travers plusieurs indicateurs quantitatifs, tout en analysant l'influence de paramètres clés : le modèle de cartographie utilisé, la structure de la grille de mesure, le nombre de capteurs déployés, ainsi que les propriétés physiques de la source de pollution.

La seconde partie du chapitre traite de la fusion de données pour la reconstruction de l'exposition individuelle à partir d'essais réalisés dans un environnement contrôlé avec un robot mobile embarquant des instruments de mesure. Elle évalue la capacité des cartographies spatio-temporelles à estimer la dose de polluant inhalée en fonction de trajectoires mesurées.

Enfin, dans la troisième partie, plusieurs cas d'application sont présentés, illustrant les usages concrets du dispositif de mesures pour la cartographie spatio-temporelle.

L'ensemble vise à démontrer l'intérêt des approches spatio-temporelles et de fusion de données pour la surveillance de l'air en milieu professionnel et la prévention des expositions.

4.2 Application des méthodes de cartographie spatio-temporelle aux données réelles : choix des paramètres

L'application des méthodes de cartographie spatio-temporelle développées dans la section 2.7 aux données expérimentales constitue une étape déterminante pour évaluer leur efficacité et leur adaptabilité opérationnelle.

Les méthodes de krigeage et de régularisation tensorielle ont été implémentées sous Matlab R2022a, tandis que la méthode PINNs a été développée en Python 3.11 avec PyTorch 2.2. L'ensemble des calculs a été exécuté sur le serveur SvpCalculV1 de l'INRS avec un processeur AMD EPYC 7502P à 32 cœurs 2,5 GHz, de 64 Go de RAM.

Pour déterminer les hyperparamètres optimaux λ et μ qui contrôlent les pénalités spatiales et temporelle dans la fonction 2.28, nous avons réalisé une exploration sur une grille. Concrètement, une grille logarithmique de valeurs a été définie pour chaque paramètre, couvrant plusieurs ordres de grandeur ($\lambda \in [10^{-6}, 1], \mu \in [10^{-8}, 10^{-2}]$). Pour chaque couple (λ, μ), la méthode

de régularisation spatio-temporelle a été appliquée sur les données expérimentales, et l'erreur quadratique moyenne (RMSE) a été calculée sur l'ensemble des données enregistrées par les capteurs de référence. La carte obtenue (Figure 4.1) illustre la distribution des valeurs de RMSE calculé sur un essai représentatif de la série 2 (S1-E2) en fonction des deux paramètres (λ, μ).

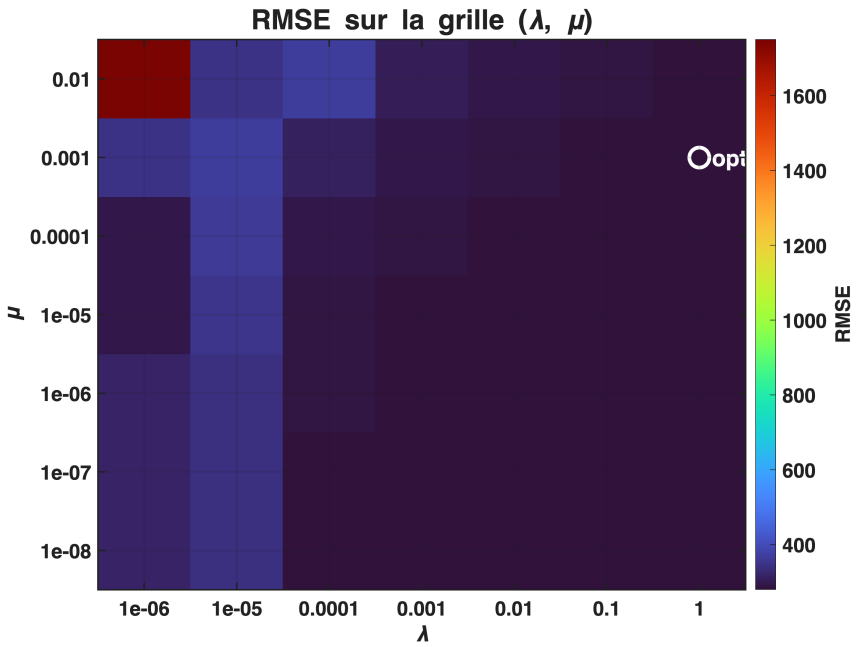


FIGURE 4.1 – RMSE en fonction des hyperparamètres de régularisation spatio-temporelle (λ, μ)

Les zones les plus foncées correspondent aux combinaisons donnant des performances élevées (erreur plus faible), tandis que les zones bleu claire et rouges indiquent une dégradation de la précision. L'optimum global, indiqué sur la carte, correspond au couple $(\lambda, \mu) = (1, 10^{-3})$ minimisant le RMSE dans ce cas.

Cette approche permet de choisir les hyperparamètres (on gardera les mêmes paramètres pour tous les jeux de données analysées), mais également de visualiser la performance de la méthode vis-à-vis de variations des hyperparamètres.

Quant aux paramètres de la méthode PINNs, le réseau a été entraîné sur 50 000 epochs. Le nombre de points de collocation a été fixé à 1,5% du domaine spatio-temporel total $N_x \times N_y \times N_t$, ces points étant échantillonnés de façon dynamique et aléatoire dans ce domaine. Pour les conditions aux frontières, un ensemble représentant 1% du domaine frontière spatio-temporel total a été utilisé.

Cette configuration, combinant un nombre élevé d'itérations avec une répartition contrôlée des points de colocation et de frontière, a été retenue afin d'assurer la stabilité de la convergence tout en maîtrisant le coût de calcul.

4.3 Étude des caractéristiques de la cartographie spatio-temporelle

Afin d'analyser les propriétés de la cartographie spatio-temporelle, nous considérons les essais expérimentaux des séries 1 et 2 (voir tableaux 3.2, 3.3), conduits dans la chambre expérimentale présentée en section 3.3.1. Cet environnement confiné et contrôlé offre la possibilité de réguler différents paramètres clés (géométrie, conditions de ventilation, modalités d'émission), ce qui contribue à améliorer la reproductibilité des mesures et à faciliter la comparaison des résultats. Ainsi, l'objectif de cette étude est d'analyser l'impact des choix de modélisation, du dispositif de mesure et des conditions de génération de la pollution sur la qualité de la reconstruction spatio-temporelle des concentrations, en examinant les principaux paramètres expérimentaux et méthodologiques.

4.3.1 Influence de la méthode de cartographie spatio-temporelle

Le choix de la méthode de cartographie constitue un facteur déterminant dans la qualité et la fiabilité des représentations spatio-temporelles des concentrations en polluants. En effet, les méthodes de reconstruction spatiale et temporelle ne se limitent pas à extrapoler les données mesurées par les capteurs : elles conditionnent également la capacité à restituer des structures fines, à lisser le bruit de mesure, ainsi qu'à préserver les dynamiques inhérentes au processus physique étudié.

Démarche expérimentale

D'un point de vue méthodologique, le choix de la méthode influe principalement sur 3 axes :

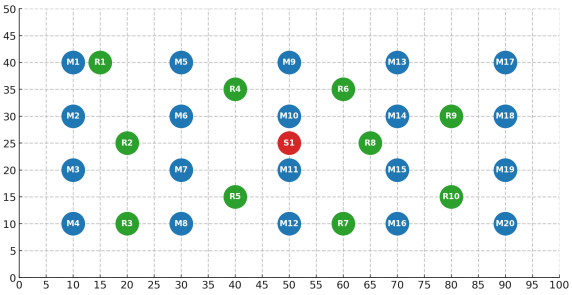
1. **Précision de l'estimation :** elle se rapporte à la capacité du modèle à générer des prédictions fiables dans les zones dépourvues de capteurs. Une méthode performante minimise l'erreur quadratique moyenne (RMSE) et maximise le coefficient de détermination (R^2), garantissant ainsi une restitution fidèle de la distribution spatio-temporelle des polluants. De plus, l'évaluation de ces métriques sur des jeux de données indépendants suivant les mêmes scénarios de mesures permet de vérifier la reproductibilité des résultats.
2. **Robustesse face aux incertitudes expérimentales :** elle désigne l'aptitude du modèle à maintenir des performances stables en présence d'une variabilité inhérente aux données de mesures, incluant le bruit instrumental et les valeurs aberrantes. L'évaluation de la dispersion des erreurs (via le RMSE) sur différents sous-ensembles de données permet d'apprécier la fiabilité du modèle face à ces perturbations.
3. **Intégration des contraintes structurelles ou physiques :** elle renvoie à l'aptitude du modèle à respecter les principes fondamentaux régissant le système étudié. Dans

notre cas, il s'agit d'évaluer la capacité de la méthode à restituer fidèlement les différentes phases de l'essai, la dynamique de la source de pollution ainsi que les conditions inhérentes à la configuration du milieu cartographié.

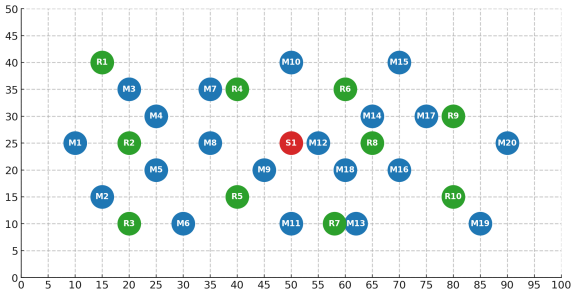
Pour étudier l'influence de la méthode de cartographie spatio-temporelle, ces 3 axes ont été considérés en utilisant les données issues des essais S1-E8, S1-E9 et S1-E12 de la série 1.

Description des essais

Les essais S1-E8, S1-E9 et S1-E12, conduits respectivement selon les grilles d'échantillonnage illustrées dans les figures 4.2a et 4.2b, ont été réalisés conformément aux plans de calibration et d'expérience B décrits dans les sections 3.7.1 et 3.4.4.

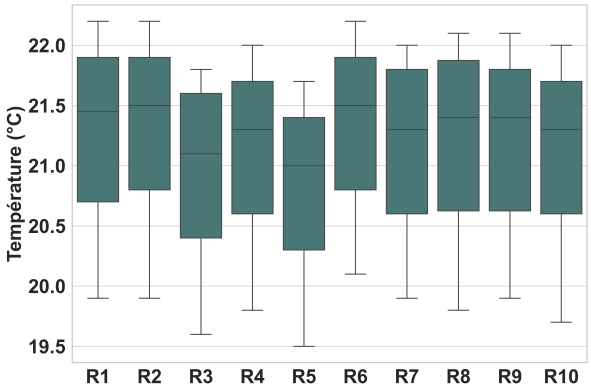


(a) Grille d'échantillonnage pour les essais S1-E8 et S1-E9

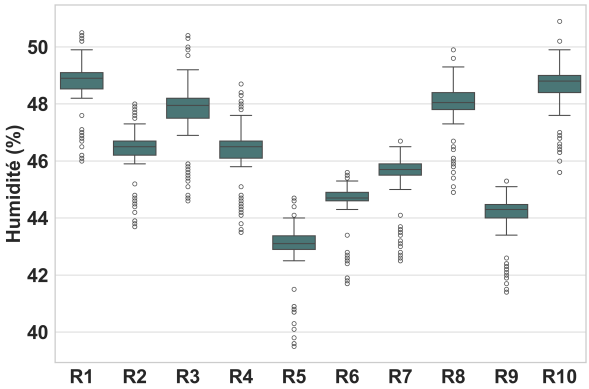


(b) Grille d'échantillonnage pour l'essai S1-E12

Ces essais ont été menés en février 2024, mais grâce au paramétrage de la chambre expérimentale, il a été possible de maintenir des conditions thermo-hygrométriques relativement stables avec une température moyenne de 21°C et une humidité relative de 46 % .



(a) Température (°C)



(b) Humidité (%))

FIGURE 4.3 – Température (°C) et Humidité relative de l'air (%) mesurées au cours de l'essai S1-E9 aux points de référence

Afin d'homogénéiser les conditions initiales de l'espace expérimental, le système de ventilation a été activé pendant 20 minutes avant chaque essai. Par la suite, les essais ont été menés avec une génération contrôlée de CO_2 au point S1 de la grille d'échantillonnage utilisée pour

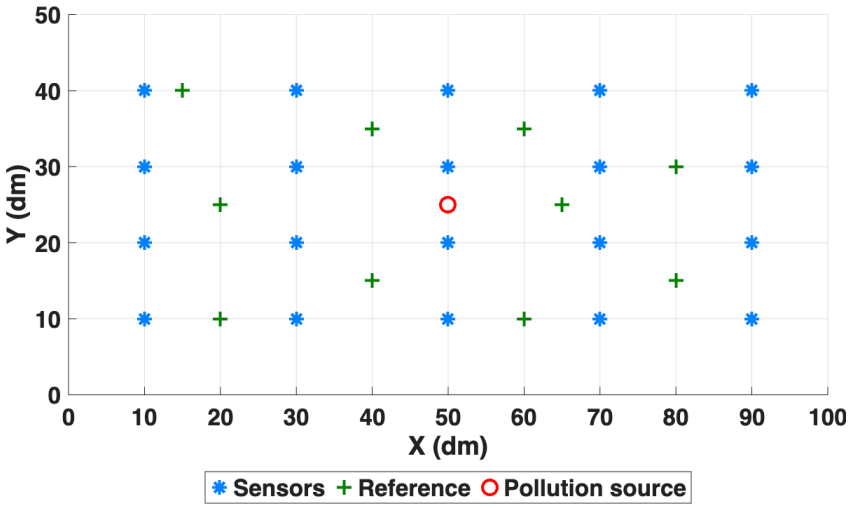
l'essai, à un débit de 12,5 L/min. Les concentrations ont été simultanément mesurées en continu aux emplacements M1 à M20 et R1 à R20, avec une fréquence d'acquisition d'une mesure par minute ; cette phase de pollution a duré $T_p = 83$ min. Cette étape a été suivie d'une phase de stabilisation d'une durée de $T_s = 20$ min, puis d'une phase d'épuration allant de $40 < T_v < 60$ min.

Des simulations numériques ont été réalisées afin de représenter l'évolution des concentrations en CO_2 dans l'espace et dans le temps selon une approche présentant la dynamique des fluides. Les caractéristiques du plan d'expérience B se prêtant à la modélisation numérique, les essais S1-E8 et S1-E9 ont été simulés à l'aide du logiciel Fluent. Ce dernier est un outil de dynamique des fluides numérique (CFD, « Computational Fluid Dynamics ») développé par ANSYS, permettant de modéliser et d'analyser les écoulements d'air, les transferts de chaleur ainsi que la dispersion de gaz dans un environnement 3D.

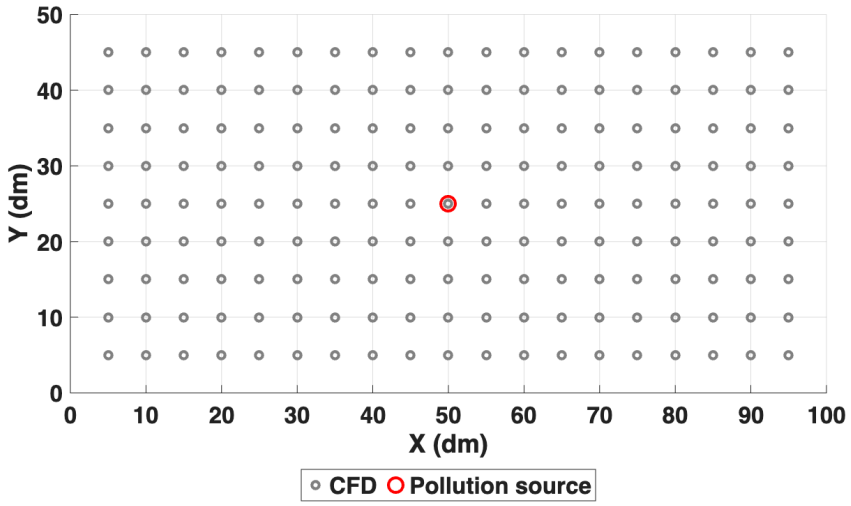
La simulation a été conduite en intégrant les paramètres techniques à savoir le débit, la position de la source et les conditions de ventilation. La chambre expérimentale, de dimensions 100×50, a été discrétisée sur une grille régulière avec un pas spatial de $\Delta_{x,y} = 5$ points. Ce maillage a permis de définir 171 positions servant de points de restitution des résultats (voir figure 4.4b). Les concentrations de CO_2 ont ainsi été extraites en chacun de ces points à chaque Δ_t fournissant une représentation discrète de la distribution spatio-temporelle du polluant dans le domaine étudié.

De façon générale avec le plan d'expérience B, l'ensemble des données obtenu est organisé en trois catégories :

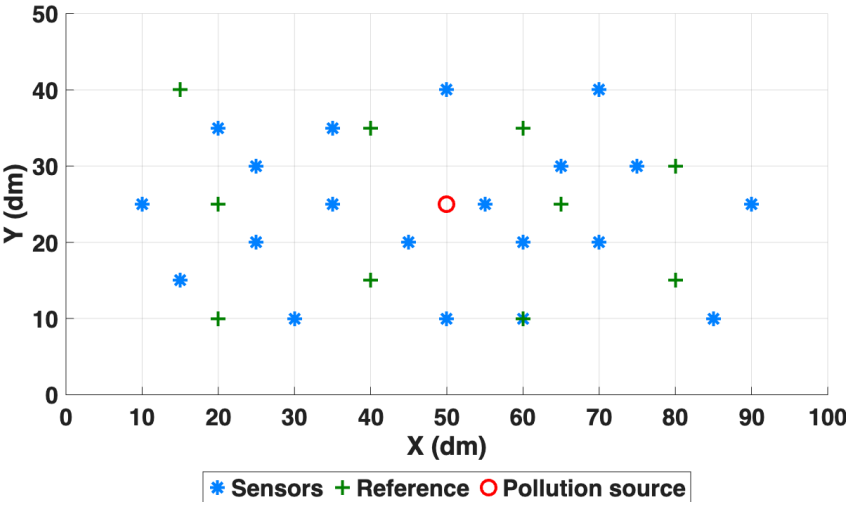
- **Données issues du réseau de capteurs (*) :** Mesures de concentrations en CO_2 enregistrées par le réseau de capteurs RCSFg, utilisées comme données d'entrées pour la génération de la cartographie spatio-temporelle.
- **Données issues des instruments de mesure utilisés comme référence (+) :** Mesures de concentration en CO_2 issues des appareils de référence positionnés en des points intermédiaires, utilisées comme données de validation de la cartographie spatio-temporelle.
- **Données issues de la simulation CFD (•) :** Concentrations en CO_2 issues des simulations de dynamique des fluides numériques (CFD) correspondant à la configuration expérimentale, utilisées comme vérité terrain.



(a) Grille synthétique des essais S1-E8 et S1-E9



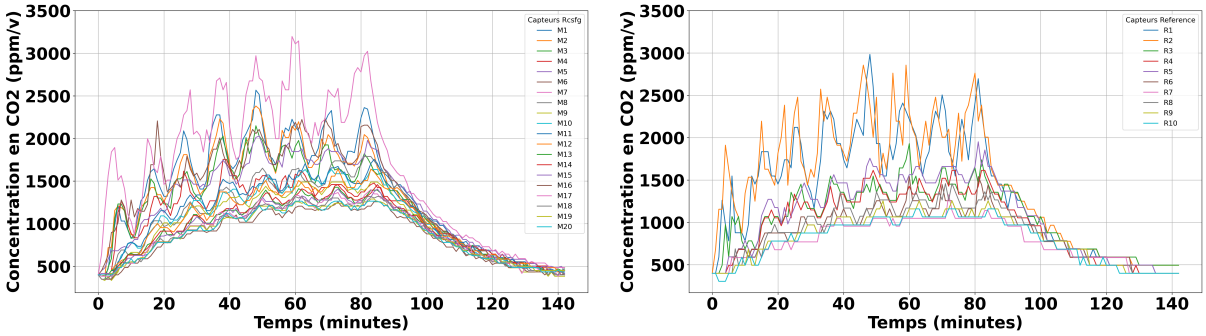
(b) Grille d'échantillonnage des points issus de la simulation CFD (Computational Fluid Dynamics)



(c) Grille synthétique de l'essai S1-E12

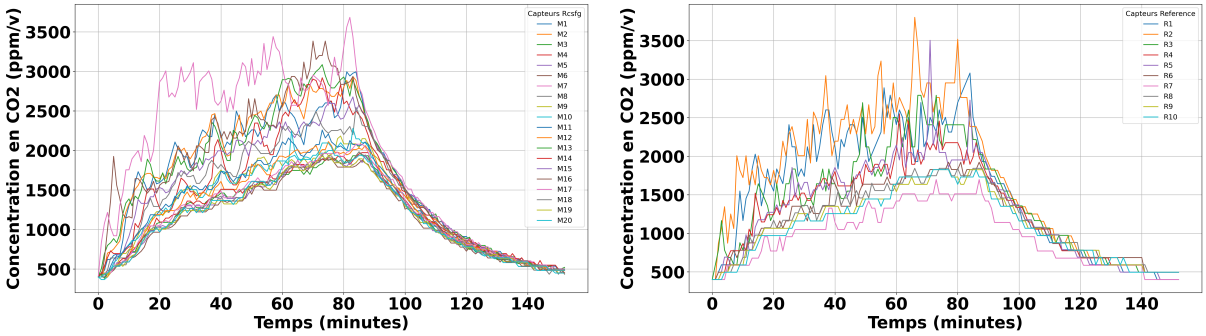
FIGURE 4.4 – Grille synthétique d'échantillonnage des points des essais S1-E8, S1-E9 (a) et S1-E12(c) et de la simulation CFD (b)

Résultats et discussions



(a) Mesures RCSFg (b) Mesures de référence

FIGURE 4.5 – Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) de l'essai S1-E8



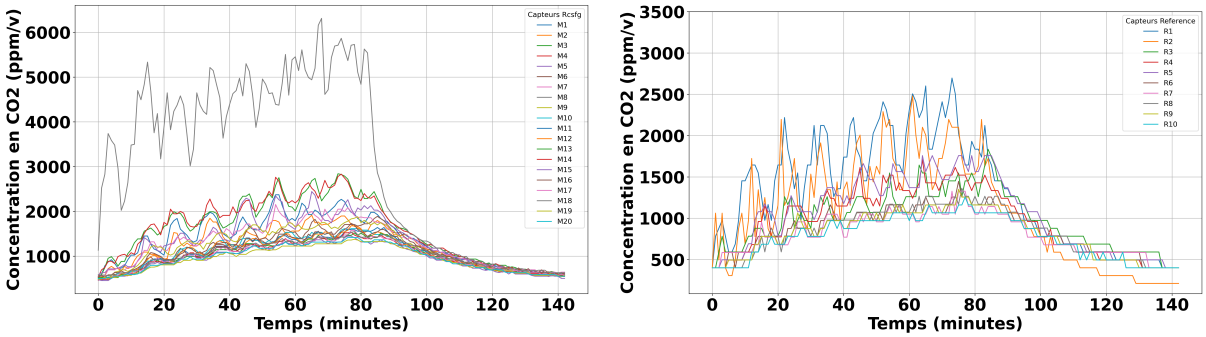
(a) Mesures RCSFg (b) Mesures de référence

FIGURE 4.6 – Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) de l'essai S1-E9

Les figures 4.5 et 4.6 présentent l'évolution temporelle des concentrations de CO_2 enregistrées respectivement par les modules capteurs M1 à M20 du réseau de capteurs et les appareils de référence R1 à R10, au cours des essais S1-E8 et S1-E9. Les résultats montrent une évolution qui met en évidence les différentes phases de l'expérience : la phase de pollution avec une montée progressive des concentrations, la phase de stabilisation avec une légère diminution due au système de ventilation qui tourne en continu, et la phase d'épuration avec une descente progressive des concentrations.

Les dynamiques observées présentent des tendances globalement similaires entre les deux essais, l'essai S1-E9 montrant une élévation plus prononcée des concentrations maximales. Les conditions expérimentales étant, en principes identiques, cette différence est certainement due à l'existence de variations liées aux propriétés des phénomènes physiques non maîtrisés dans la chambre expérimentale ou à un dysfonctionnement de la source. En effet, des fuites d'air peuvent se produire à certains endroits et il a été également constaté que la distribution spatiale du CO_2 dans la chambre pouvait être influencée par de légères hétérogénéités de circulation

d'air ou par des fluctuations locales du débit de ventilation. Cette observation met en évidence l'importance de considérer, dans l'analyse, non seulement les concentrations absolues mesurées, mais également les mécanismes de dispersion et de diffusion propres au système étudié.



(a) Mesures RCSFg

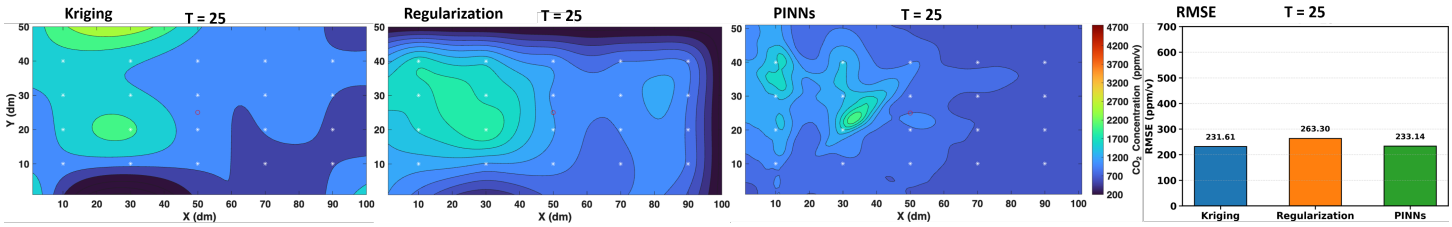
(b) Mesures de référence

FIGURE 4.7 – Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) de l'essai S1-E12

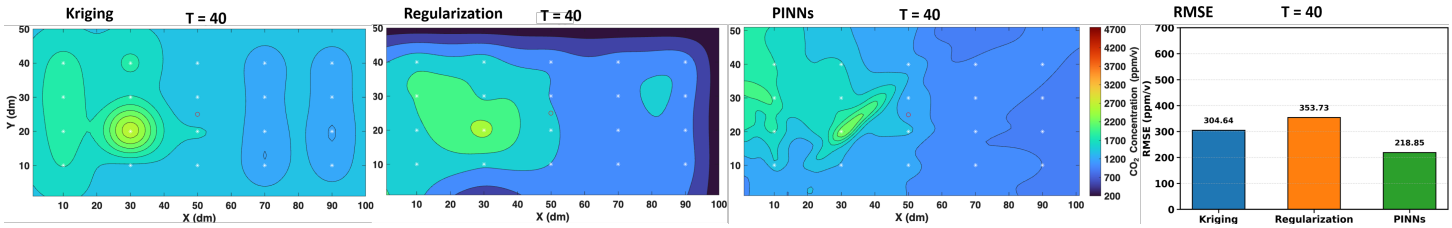
L'évolution des concentrations en CO_2 enregistrées respectivement par les capteurs M1 à M20 du réseau de capteurs et les appareils de référence R1 à R10, au cours de l'essai S1-E12, révèle aussi des différences. En effet, les mêmes dynamiques sont observées à l'exception du module capteur M8 du RCSFg qui a enregistré des valeurs plus élevées que celles relevées par les autres modules capteurs, indiquant une divergence significative par rapport à la dynamique attendue. Ce type d'écart peut résulter d'un dysfonctionnement du capteur ou de sa sensibilité aux conditions aérodynamiques locales, notamment aux turbulences de l'air couplées aux phénomènes de diffusion du CO_2.

Les cartes spatio-temporelles ont été générées avec les méthodes de cartographie spatio-temporelle décrites dans la section 2.7 appliquées aux données de concentrations de CO_2 des essais S1-E8, S1-E9 et S1-E12 (voir Figures 4.5–4.7). Ces cartographies, présentées sous forme de champs de concentration, traduisent la dynamique d'évolution du CO_2 dans l'espace de la chambre expérimentale au cours du temps (voir Figures 4.8a–4.10e). Elles représentent l'état spatio-temporel des concentrations de CO_2 pour différents instants représentatifs de l'expérience : $T = 25$ min, 40 min, 83 min pour la phase de pollution, $T = 95$ min pour la phase de stabilisation et $T = 120$ min pour la phase d'épuration. Des vidéos complémentaires illustrant les cartographies spatio-temporelles obtenues pour l'ensemble des méthodes considérées sont disponibles en ligne.

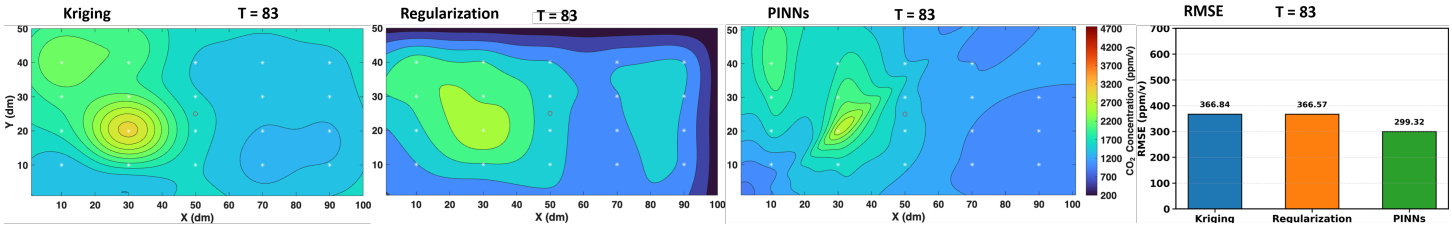
— Vidéo de la cartographie spatio-temporelle de l'essai S1-E9 avec la méthode de krigeage
— Vidéo de la cartographie spatio-temporelle de l'essai S1-E9 avec la méthode de régularisation tensorielle
— Vidéo de la cartographie spatio-temporelle de l'essai S1-E9 avec la méthode PINNs



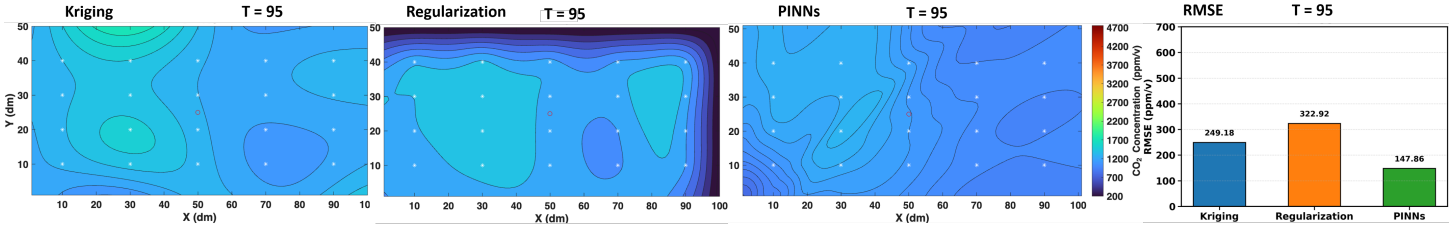
(a) $T = 25$ min | Phase de pollution



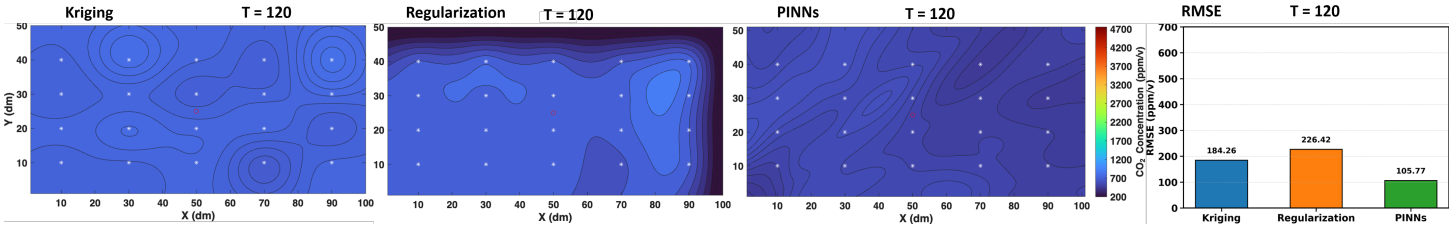
(b) $T = 40$ min | Phase de pollution



(c) $T = 83$ min | Phase de pollution

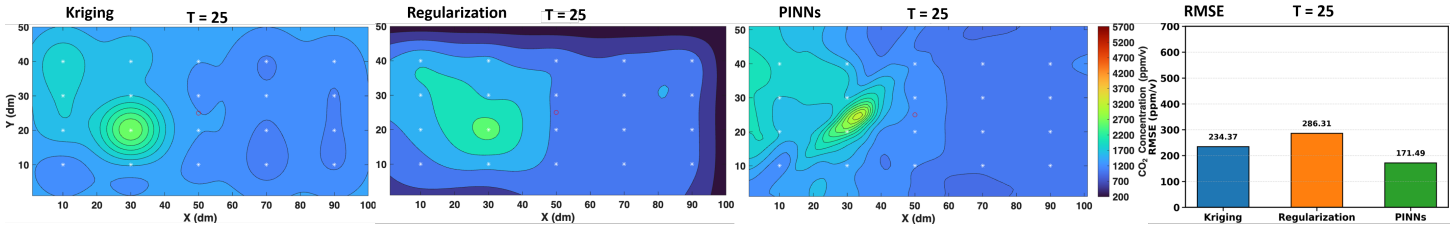


(d) $T = 95$ min | Phase de stabilisation

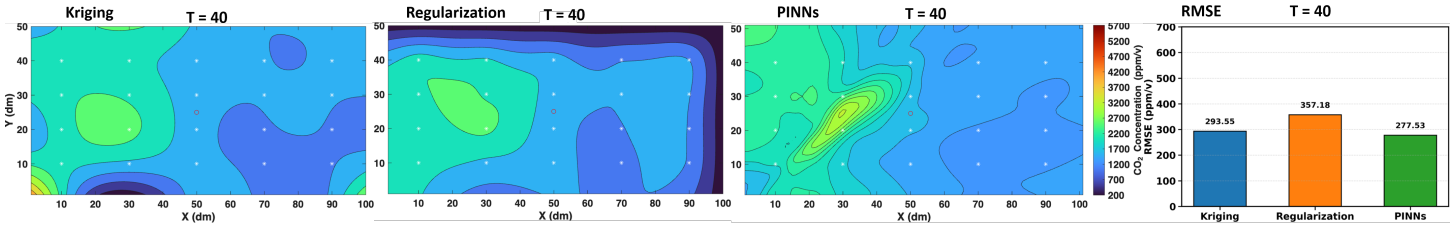


(e) $T = 120$ min | Phase d'épuration

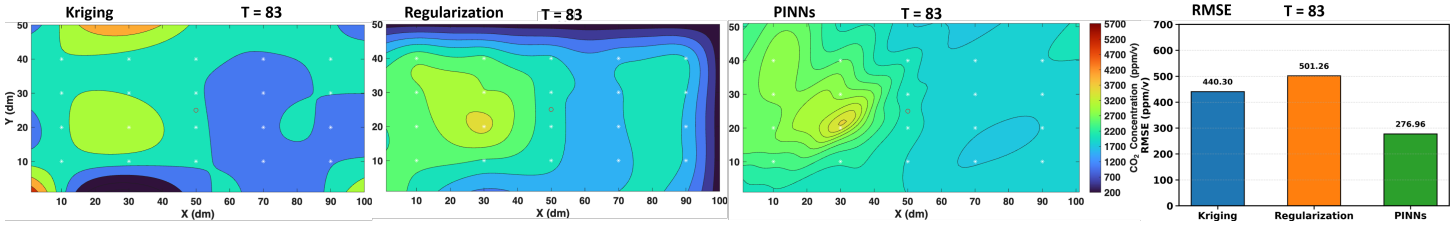
FIGURE 4.8 – Cartographies spatio-temporelles des concentrations de CO_2 et évaluation du RMSE aux instants $T = 25$ min, 40 min, 83 min, 95 min, 120 min de l'essai S1-E8



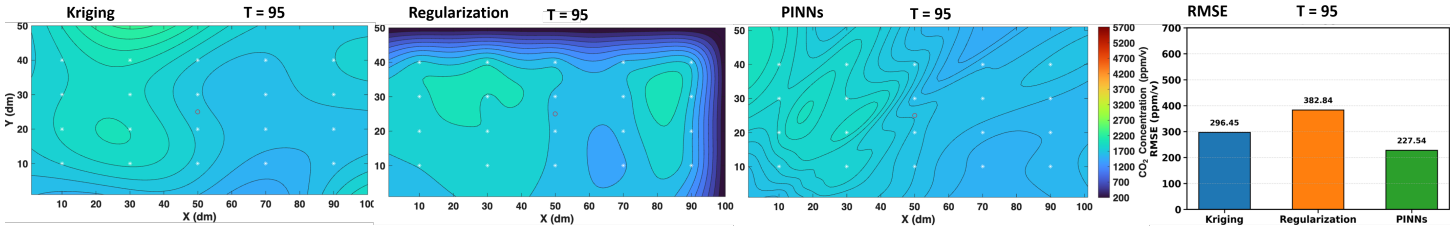
(a) $T = 25$ min | Phase de pollution



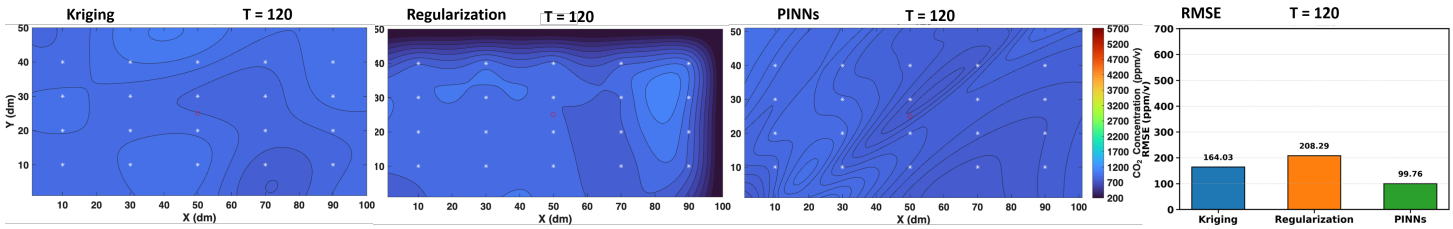
(b) $T = 40$ min | Phase de pollution



(c) $T = 83$ min | Phase de pollution

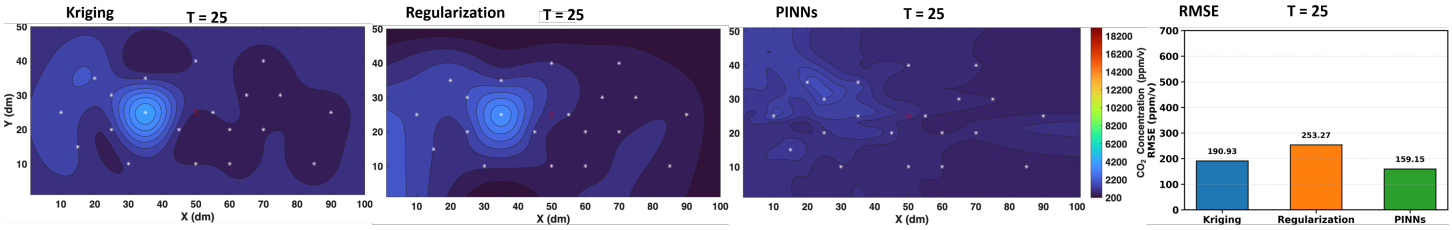


(d) $T = 95$ min | Phase de stabilisation

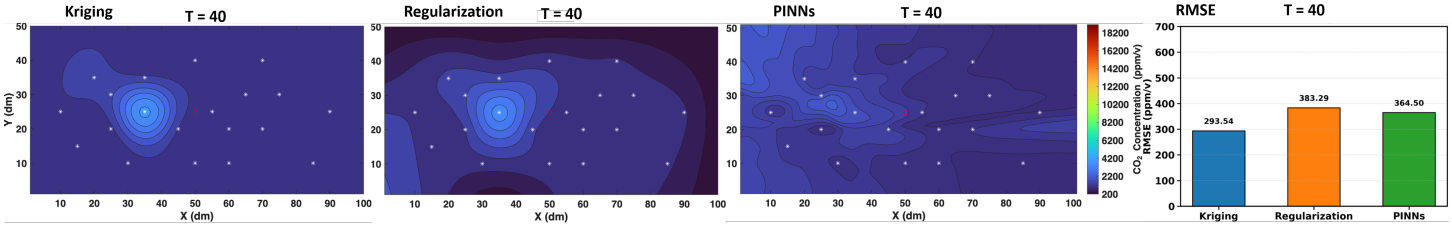


(e) $T = 120$ min | Phase d'épuration

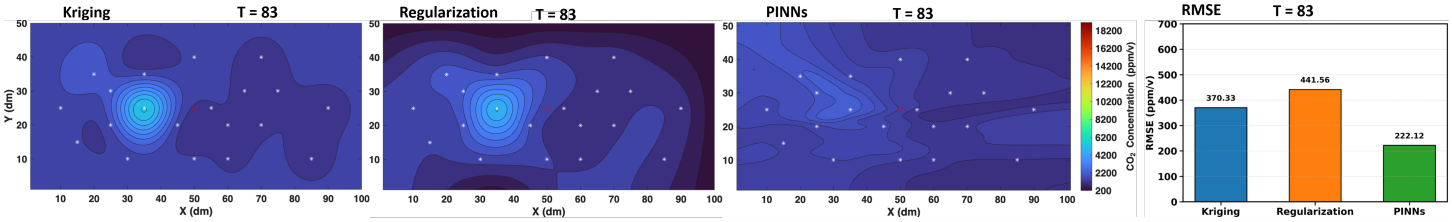
FIGURE 4.9 – Cartographies spatio-temporelles des concentrations de CO_2 et évaluation du RMSE aux instants $T = 25$ min, 40 min, 83 min, 95 min, 120 min de l'essai S1-E9



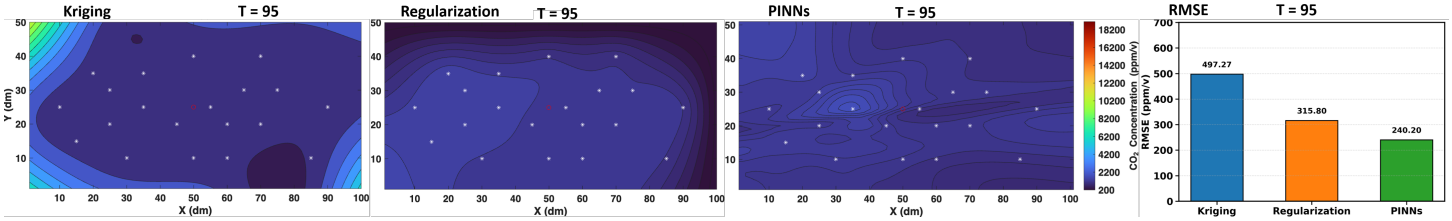
(a) $T = 25$ min | Phase de pollution



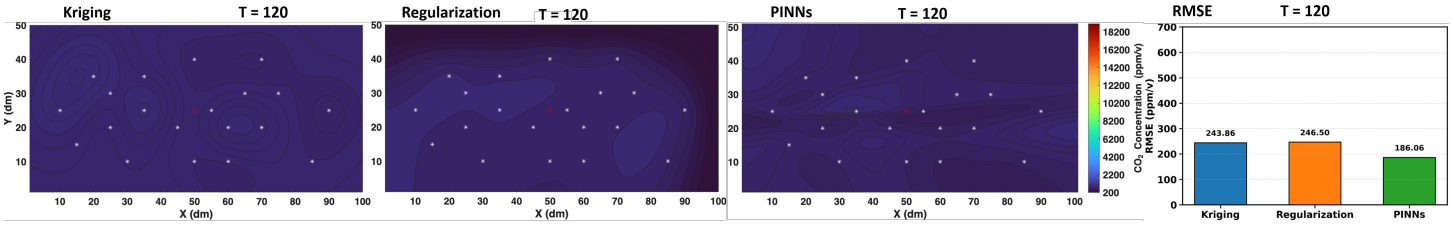
(b) $T = 40$ min | Phase de pollution



(c) $T = 83$ min | Phase de pollution



(d) $T = 95$ min | Phase de stabilisation



(e) $T = 120$ min | Phase d'épuration

FIGURE 4.10 – Cartographies spatio-temporelles des concentrations de CO_2 et évaluation du RMSE aux instants $T = 25$ min, 40 min, 83 min, 95 min, 120 min de l'essai S1-E12

Les cartes spatio-temporelles présentées dans les figures 4.8 et 4.9 , sont données pour les trois méthodes (le krigeage, la régularisation spatio-temporelle et les réseaux de neurones informés par la physique (PINNs)). On observe que, durant la phase de pollution (T = 25 min, 40 min, 83 min), les cartes générées par le krigeage présentent des zones de concentration lissées, mais avec une tendance à la diffusion excessive des gradients spatiaux. Cette limitation est couramment rapportée dans la littérature comme une conséquence du caractère globalement linéaire et isotrope de l'interpolation géostatistique, qui tend à atténuer les extrêmes locaux [Cressie15].

Les cartes produites avec la méthode de régularisation présentent une meilleure préservation des gradients locaux, mais laissent apparaître des artefacts liés à la paramétrisation du terme de lissage. Il s'agit notamment d'effets de bord nettement visibles dans la partie supérieure et droite de la carte (où les valeurs estimées sont très faibles), dus à l'application des opérateurs de différentiation selon x et y.

Les cartes obtenues avec la méthode PINNs mettent en évidence une restitution plus fidèle des panaches de pollution, notamment aux temps intermédiaires, où l'hétérogénéité spatio-temporelle atteint son maximum. L'intégration conjointe des données ponctuelles et des contraintes physiques favorise une meilleure préservation de la dynamique des fronts de concentration, conduisant à une représentation plus cohérente des processus de diffusion et de transport du CO_2 ([Raissi19], [Karniadakis21]).

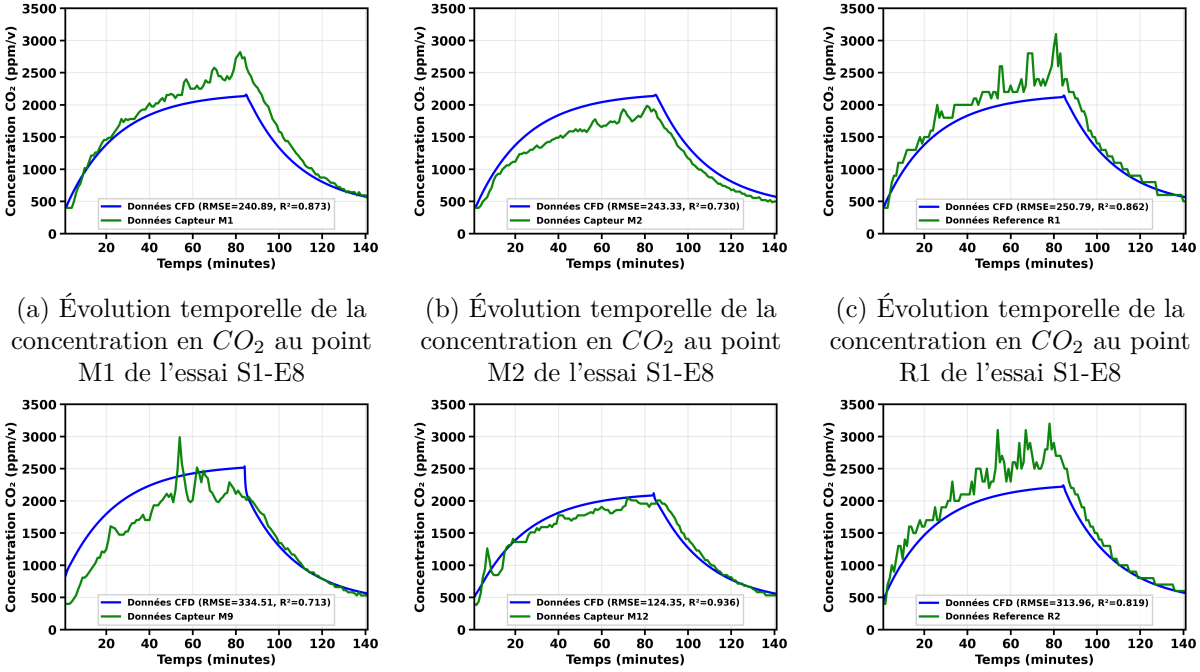
Les résultats montrent donc qu'il y a une influence de la méthode utilisée pour le cartographie sur la cartes obtenue à partir des mesures. Les cartes mettent en évidence les compromis inhérents aux familles de méthodes considérées : les approches strictement statistiques, généralement régularisées par des pénalités de lissage et des hypothèses de stationnarité, favorisent la régularité globale au prix d'une moindre fidélité aux structures locales ; à l'inverse, les schémas hybrides intégrant des contraintes physiques présentent une meilleure capacité de généralisation et restituent plus fidèlement la dynamique sous-jacente, produisant des reconstructions spatio-temporelles plus cohérentes ([Gao21], [Peng22]).

Résultats et discussions

L'évaluation de la qualité des cartographies a été réalisée par le calcul du RMSE à chaque instant pour les essais S1-E8 et S1-E9. Comme illustré dans la Figure 4.12, cette analyse a été conduite selon deux approches complémentaires : La première approche utilise des mesures expérimentales effectuées avec dix capteurs de référence, permettant d'estimer la fidélité des reconstructions (Figures 4.12a et 4.12c). La seconde approche exploite les 171 points de comparaison issus de la simulation CFD, offrant une appréciation plus globale de la cohérence spatio-temporelle (Figures 4.12b et 4.12d). Cette double approche permet ainsi de confronter les performances des méthodes de reconstruction à la fois à des données instrumentales réelles

et à un modèle numérique réaliste à haute résolution du champ de concentration.

Afin d'illustrer la comparaison entre les mesures *in situ* et les résultats de la simulation numérique, la figure 4.11 présente quelques exemples représentatifs des couples modules capteurs/CFD. Les courbes montrent, pour différents emplacements, l'évolution temporelle des concentrations en CO_2 enregistrées par les modules capteurs et celles obtenues par modélisation numérique. On observe que les signaux présentent globalement une cohérence satisfaisante, tant au niveau de la dynamique d'évolution que des ordres de grandeur atteints. Toutefois, certaines divergences locales apparaissent, notamment dans les phases transitoires de montée et de décroissance, ce qui se traduit par des valeurs de RMSE variables selon les points considérés. Ces divergences peuvent être attribuées à la limitation de la représentation des phénomènes de turbulence lors du processus de modélisation, en raison du temps de simulation extrêmement long (5 jours) avec le logiciel Fluent. Les exemples présentés (Voir figure 4.11) correspondent aux points M1 et M2 de l'essai S1-E8, aux points M9 et M12 de l'essai S1-E9, en complément des points de référence R1 et R2 de l'essai S1-E8. Pour des raisons de lisibilité, l'intégralité des comparaisons effectuées pour l'ensemble des positions instrumentées est reportée dans les tableau en annexe 19, 20, 21.



(a) Évolution temporelle de la concentration en CO_2 au point M1 de l'essai S1-E8

(b) Évolution temporelle de la concentration en CO_2 au point M2 de l'essai S1-E8

(c) Évolution temporelle de la concentration en CO_2 au point R1 de l'essai S1-E8

(d) Évolution temporelle de la concentration en CO_2 au point M9 de l'essai S1-E9

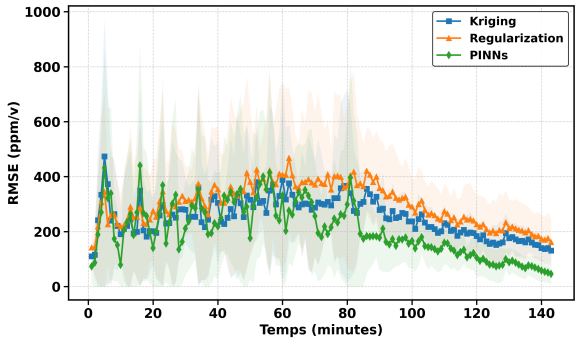
(e) Évolution temporelle de la concentration en CO_2 au point M12 de l'essai S1-E9

(f) Évolution temporelle de la concentration en CO_2 au point R1 de l'essai S1-E8

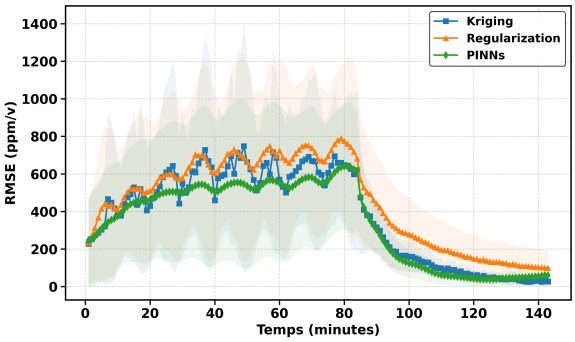
FIGURE 4.11 – Analyse comparative des signaux enregistrés par les capteurs et les données issues de la simulation CFD

L'évolution de la RMSE au cours du temps apporte des informations sur la cartographie (Voir Figure 4.12). D'abord, on observe que la qualité de la cartographie évolue au cours du temps

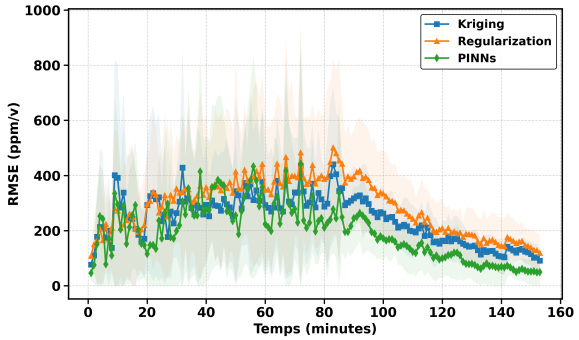
et/ou en fonction des niveaux de concentrations. Ensuite, les écarts types associés montrent que la qualité de la cartographie peut varier localement à un temps donné en fonction de l'endroit l'on se trouve. Enfin, la figure 4.12 met en évidence l'impact déterminant du choix du modèle de cartographie spatio-temporelle sur la qualité des reconstructions des concentrations de CO_2. En effet, les résultats montrent que la performance varie sensiblement selon l'approche retenue, en particulier dans les régimes transitoires où la dynamique spatio-temporelle est la plus complexe à restituer.



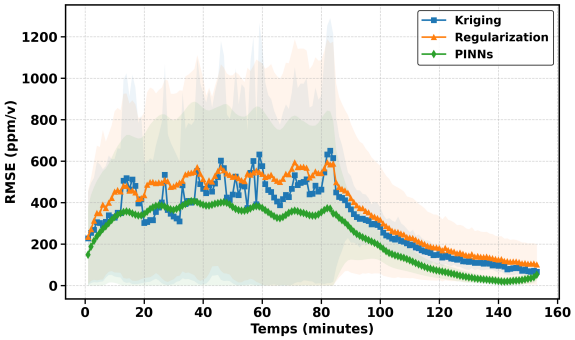
(a) Evolution temporelle du RMSE calculé avec 10 points de référence | Essai S1-E8



(b) Evolution temporelle du RMSE calculé avec 171 points issues de la simulation CFD | Essai S1-E8



(c) Evolution temporelle du RMSE calculé avec 10 points de référence | Essai S1-E9



(d) Evolution temporelle du RMSE calculé avec 171 points issues de la simulation CFD | Essai S1-E9

FIGURE 4.12 – Évolution temporelle du RMSE obtenues avec les méthodes de krigeage, de régularisation et de PINNs pour la reconstruction spatio-temporelle des concentrations de CO_2 des essais S1-E8 et S1-E9

L'évolution temporelle du RMSE fournit une lecture synthétique de la stabilité et de la précision des reconstructions spatio-temporelles (Figure 4.12). Elle montre que la qualité de la cartographie varie dans le temps et selon la localisation, soulignant l'hétérogénéité du champ de concentration et la sensibilité des modèles aux conditions locales. Les Figures 4.13 et 4.14 illustrent cette variabilité pour deux points de référence représentatifs, mettant en évidence les différences de comportement entre les méthodes de krigeage, de régularisation et de PINNs. Les résultats synthétisés dans les Tableaux 4.1 à 4.3 confirment l'influence déterminante du modèle de cartographie sur la qualité de reconstitution des concentrations, avec des performances

contrastées selon les essais et la position du point de référence.

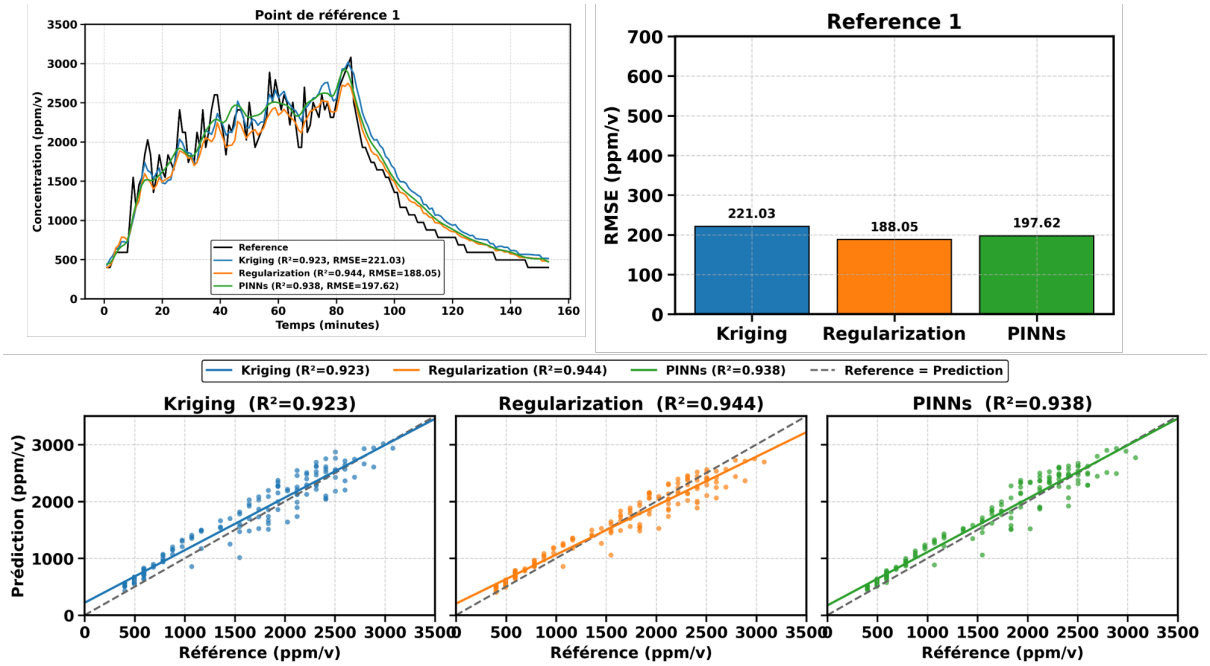


FIGURE 4.13 – Analyse des performances des méthodes (krigeage, Régularisation, PINNs) pour la reconstruction des concentrations de CO_2 au point de référence 1 | Essai S1-E9

TABLEAU 4.1 – Analyse des métriques (RMSE et R^2) pour les 10 points de référence (R1-R10) de l'essai S1-E8

Référence	Krigeage		Regularization		PINNs	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	206.04	0.907	221.86	0.893	266.75	0.845
R2	340.51	0.757	301.85	0.809	392.38	0.678
R3	162.92	0.814	200.01	0.720	304.26	0.351
R4	274.31	0.520	383.36	0.062	298.08	0.433
R5	326.39	0.477	303.69	0.547	121.31	0.928
R6	235.40	0.367	215.72	0.468	174.82	0.651
R7	296.94	-0.519	241.56	-0.006	174.16	0.477
R8	192.44	0.655	167.26	0.739	131.01	0.840
R9	278.89	-0.006	508.46	-2.345	115.86	0.826
R10	239.87	0.184	388.95	-1.147	80.26	0.909

Le krigeage présente une évolution temporelle du RMSE relativement stable, traduisant son efficacité dans les régimes stationnaires. Toutefois, sa capacité limitée à capturer les variations

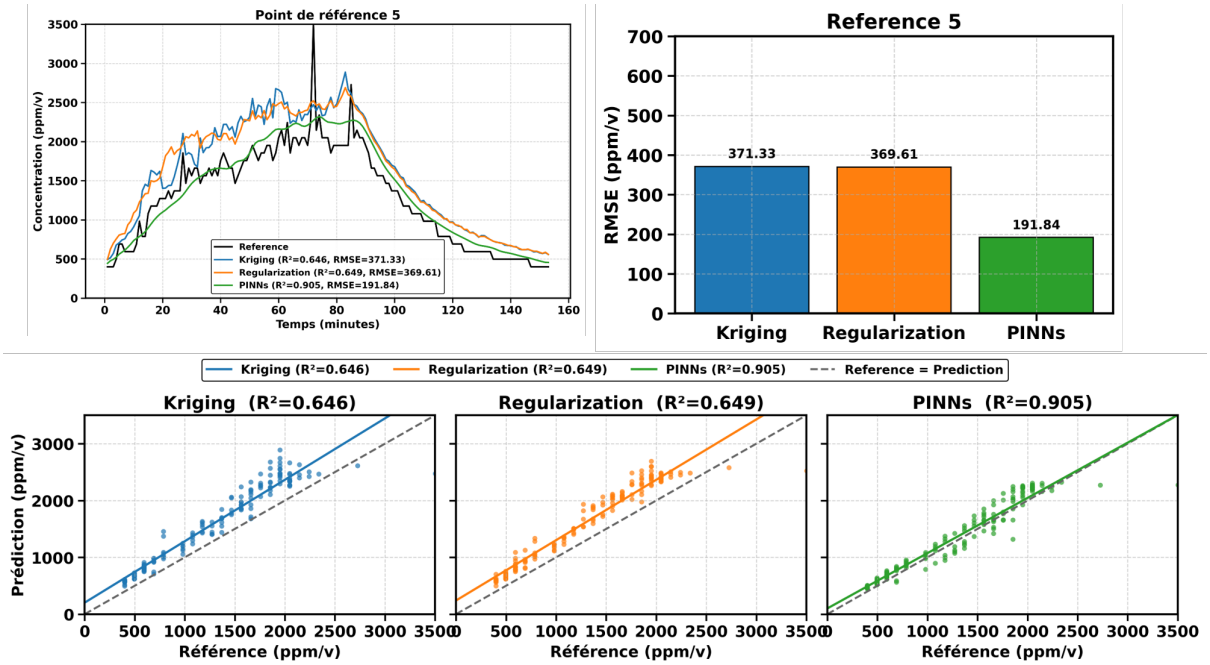


FIGURE 4.14 – Analyse des performances des méthodes (krigeage, Régularisation, PINNs) pour la reconstruction des concentrations de CO_2 au point de référence 5 | Essai S1-E9

TABLEAU 4.2 – Analyse des métriques (RMSE et R^2) pour les 10 points de référence (R1-R10) de l'essai S1-E9

Référence	Krigeage		Regularization		PINNs	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	221.03	0.923	188.05	0.944	197.62	0.938
R2	348.46	0.838	301.66	0.879	310.89	0.871
R3	198.35	0.913	222.77	0.891	338.15	0.748
R4	246.38	0.818	345.58	0.642	306.34	0.719
R5	371.33	0.646	369.61	0.649	191.84	0.905
R6	157.18	0.883	123.22	0.928	88.70	0.963
R7	369.08	0.046	293.75	0.395	260.76	0.524
R8	131.10	0.918	98.11	0.954	95.63	0.956
R9	233.84	0.710	512.69	-0.396	103.98	0.943
R10	190.84	0.816	373.98	0.292	66.38	0.978

TABLEAU 4.3 – Analyse des metriques (RMSE et R^2) pour les 10 points de référence (R1-R10) de l'essai S1-E12

Référence	Krigeage		Regularization		PINNs	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	620.52	0.085	345.03	0.717	333.68	0.735
R2	366.28	0.655	420.85	0.544	425.34	0.535
R3	232.85	0.527	170.19	0.747	199.77	0.652
R4	287.59	0.468	525.43	-0.777	235.62	0.643
R5	386.90	0.221	542.44	-0.531	192.76	0.807
R6	244.76	0.167	214.91	0.358	138.83	0.732
R7	231.17	0.102	236.20	0.063	137.40	0.683
R8	227.99	0.404	211.10	0.489	80.09	0.926
R9	242.36	0.194	228.29	0.285	123.48	0.791
R10	254.19	0.002	367.26	-1.082	218.29	0.264

rapides dans l'espace et dans le temps souligne ses faiblesses structurelles dès lors que le processus étudié s'éloigne d'un comportement quasi-statique. La régularisation spatio-temporelle améliore la restitution dans certaines phases, mais son comportement reste fortement dépendant des paramètres de pondération et de l'équilibre entre les termes spatial et temporel. Cette approche tend ainsi à présenter une variabilité plus marquée du RMSE et à surestimer l'erreur dans les phases transitoires, ce qui traduit une difficulté à maintenir une cohérence globale entre dynamique temporelle et distribution spatiale. Les PINNs se distinguent par leur capacité à intégrer simultanément les contraintes spatiales, temporelles et physiques. Elles parviennent à restituer de manière plus fidèle la dynamique transitoire, en particulier lors des phases de décroissance rapide des concentrations.

Bien que S1-E8 et S1-E9 aient été conduits sous des paramètres expérimentaux identiques, les séries de mesures brutes ne sont pas strictement répétables (Voir figures 4.5, 4.6) en raison de micro-dysfonctionnements de la chambre expérimentale. En termes métrologiques, la répétabilité instrumentale est donc limitée. En revanche, les méthodes proposées sont déterministes et parfaitement réplicables : à données constantes, elles produisent des cartes identiques. Malgré les écarts de mesure, les deux essais présentent des dynamiques spatio-temporelles comparables à toutes les phases de l'expérience. Ce qui confirme la cohérence du protocole.

Une autre question concerne la performance des méthodes lorsque l'un des capteurs enregistre des valeurs aberrantes. C'est la situation observée lors de l'essai S1-E12, qui permet d'évaluer l'impact de telles anomalies instrumentales sur la qualité et la cohérence des cartographies spatio-temporelles générées (Voir figure 4.10). En observant les cartes produites, on remarque que le krigeage, fortement dépendant des mesures ponctuelles et de leur pondération spatiale, amplifie l'effet des valeurs aberrantes : les cartes correspondantes présentent des

écarts aux zones adjacentes, avec des concentrations artificiellement élevées ou abaissées autour du capteur défectueux. Cette sensibilité excessive conduit à une distorsion locale du champ reconstruit et se traduit par une perte de cohérence spatiale globale.

La RMSE obtenu présente des valeurs particulièrement élevées à certains instants représentatifs. Pourtant, son calcul n'intègre pas de points de référence aux endroits où les concentrations sont extrêmement élevées (jusqu'à 18 200 ppm/v) qui sont aux bords de l'espace étudié. Par conséquent, ces écarts extrêmes ne se traduisent pas directement dans la dynamique temporelle du RMSE (voir figure 4.15).

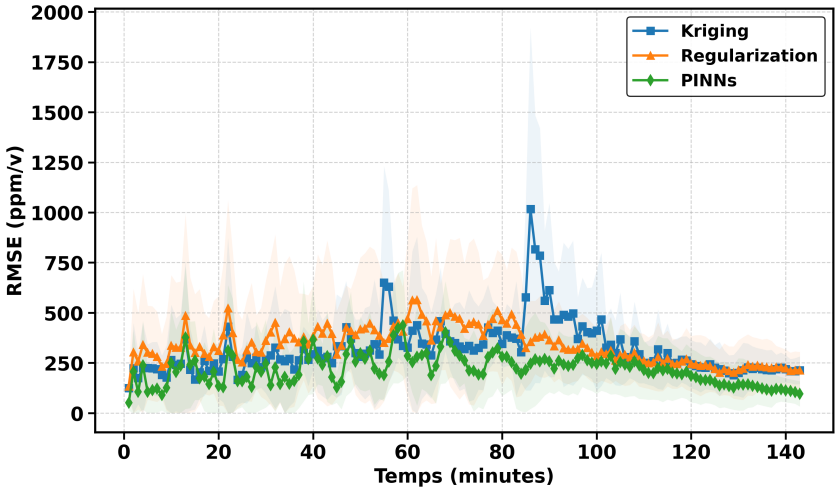


FIGURE 4.15 – Évolution temporelle du RMSE obtenues avec les méthodes de krigeage, de régularisation et de PINNs pour la reconstruction spatio-temporelle des concentrations de CO_2 de l'essai S1-E12

L'analyse comparative souligne que le choix du modèle de cartographie spatio-temporelle exerce une influence déterminante sur la précision et la robustesse des reconstructions. Comme l'ont montré [Koehler13] dans le cadre de la cartographie des expositions par interpolation géostatistique, la qualité des résultats dépend étroitement des hypothèses méthodologiques et du paramétrage retenu. Dans cette perspective, si les approches statistiques telles que le krigeage présentent une stabilité intéressante en régime stationnaire, l'introduction de méthodes de régularisation tensorielle permet d'améliorer leur capacité à représenter des dynamiques plus évolutives. Par ailleurs, les réseaux de neurones informés par la physique (PINNs) offrent un potentiel supplémentaire en intégrant directement les contraintes physiques au processus d'apprentissage, ce qui contribue à une représentation plus réaliste des phénomènes complexes de

dispersion en milieu de travail.

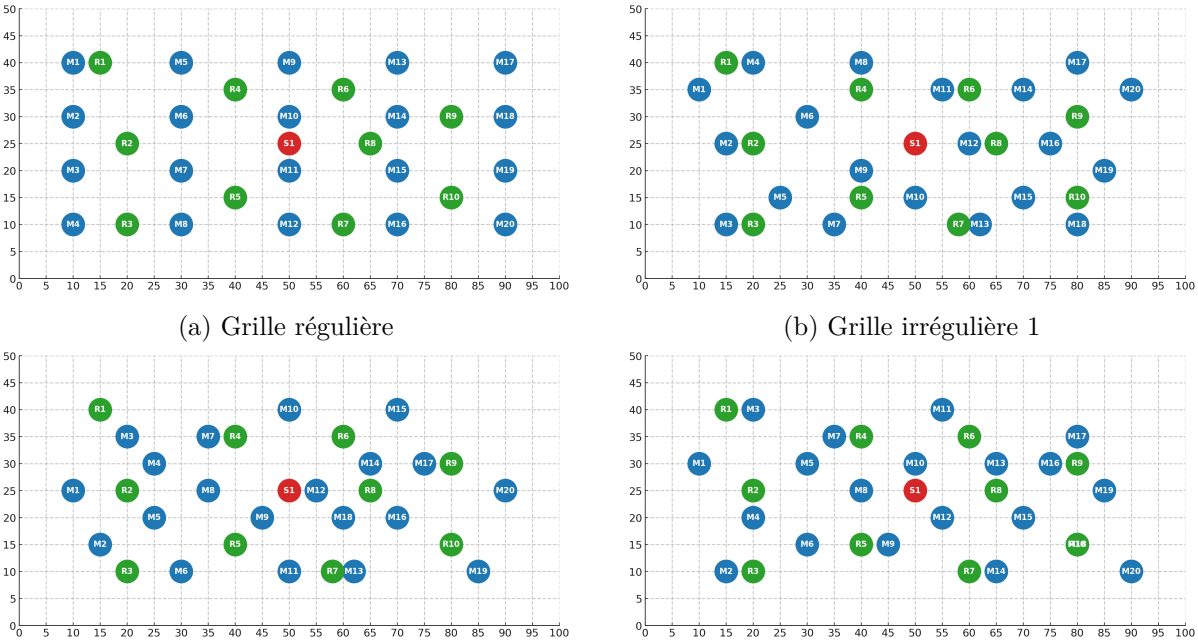
4.3.2 Influence de la grille de positionnement des points de mesure

L'échantillonnage des points de mesures constitue un paramètre essentiel dans la qualité des reconstructions spatio-temporelles des concentrations en polluants. En effet, le choix de la grille de positionnement conditionne directement la densité d'information disponible, la représentativité des phénomènes locaux et la capacité à restituer fidèlement les gradients de concentration dans l'espace étudié.

Cette section a pour objectif de quantifier l'impact de la configuration d'échantillonnage (régulière ou irrégulière) sur la précision des cartographies reconstruites et d'évaluer la robustesse des méthodes de reconstruction face aux hétérogénéités spatiales de l'échantillonnage.

Démarche expérimentale

Pour analyser l'influence de l'échantillonnage des points de mesures sur les cartographies spatio-temporelles, deux types de grilles avec une certaine homogénéité de la couverture spatiale ont été conçues ; un premier type correspond à une grille régulière qui assure une répartition équilibrée des capteurs sur l'ensemble du domaine, garantissant une couverture uniforme et limitant les zones insuffisamment échantillonnées. Une second type correspond à une grille irrégulière, issue d'une génération aléatoire et qui induit des densités de capteurs variables, pouvant entraîner une hétérogénéité dans la résolution spatiale et accroître la sensibilité des reconstructions aux biais locaux.



(a) Grille régulière

(b) Grille irrégulière 1

(c) Grille irrégulière 2

(d) Grille irrégulière 3

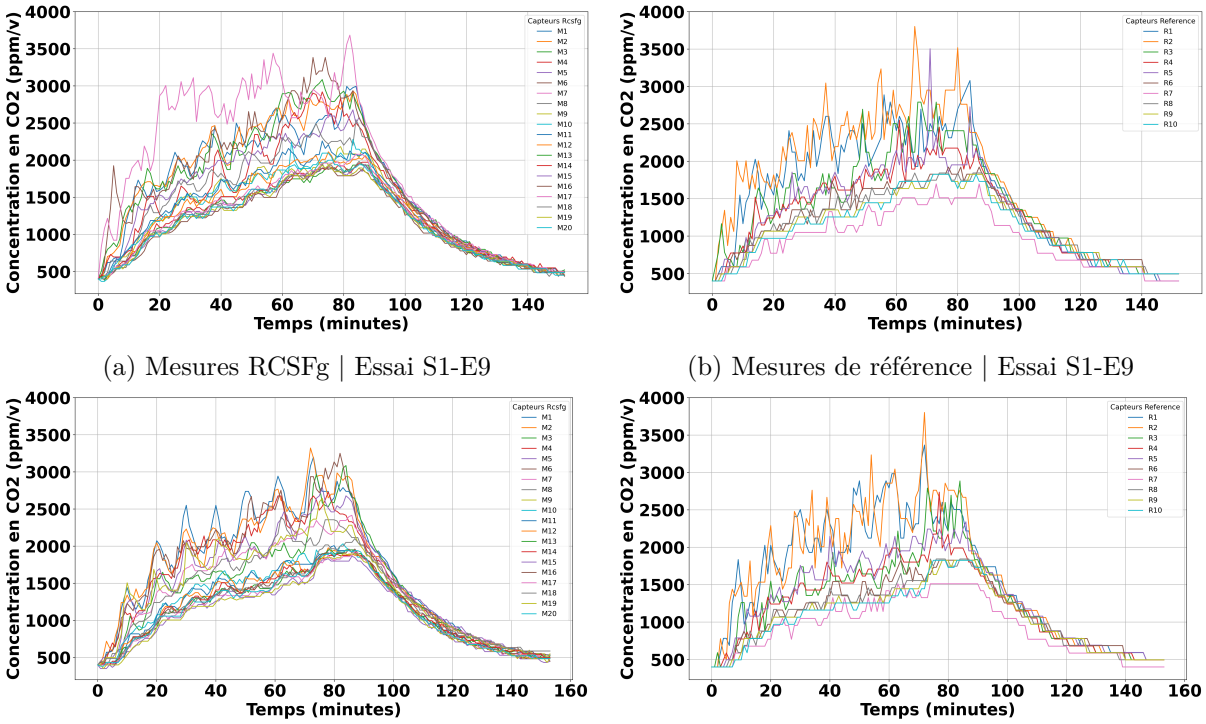
FIGURE 4.16 – Quatres grilles d'échantillonnage avec 20 points de mesures

Pour l'ensemble des scénarios, la source d'émission de CO_2 (S1 en rouge) ainsi que les instruments de référence (R1 à R10 en vert) ont été maintenus aux mêmes positions afin d'assurer une comparabilité des mesures. Seule la disposition spatiale des modules capteurs (M1 à M20 en bleu) du RCSFGg a été modifiée, selon une grille régulière et trois variantes de grilles irrégulières.

Les essais expérimentaux ont été réalisés en mars 2024 suivant les plans de calibration et d'expérience B au sein de la chambre expérimentale (Voir section 3.3.1). Les conditions thermo-hygrométriques étaient contrôlées avec une température moyenne de 21°C et une humidité relative de 46 %. Les concentrations en CO_2 mesurées étaient de deux catégories : les données issues des modules capteurs (M1–M20), utilisées pour la génération des cartographies spatio-temporelles, et les mesures issues instruments de référence (R1–R10).

Résultats et discussions

Les figures 4.17 - 4.18 présentent l'évolution temporelle de ces mesures pour chacun des essais S1-E9, S1-E10, S1-E11 et S1-E13. Les profils temporels obtenus suivent une dynamique comparable, caractérisée par une montée progressive liée à la diffusion de CO_2, suivie d'une décroissance à la phase de stabilisation et d'épuration. Malgré la variation du positionne-



(a) Mesures RCSFg | Essai S1-E9

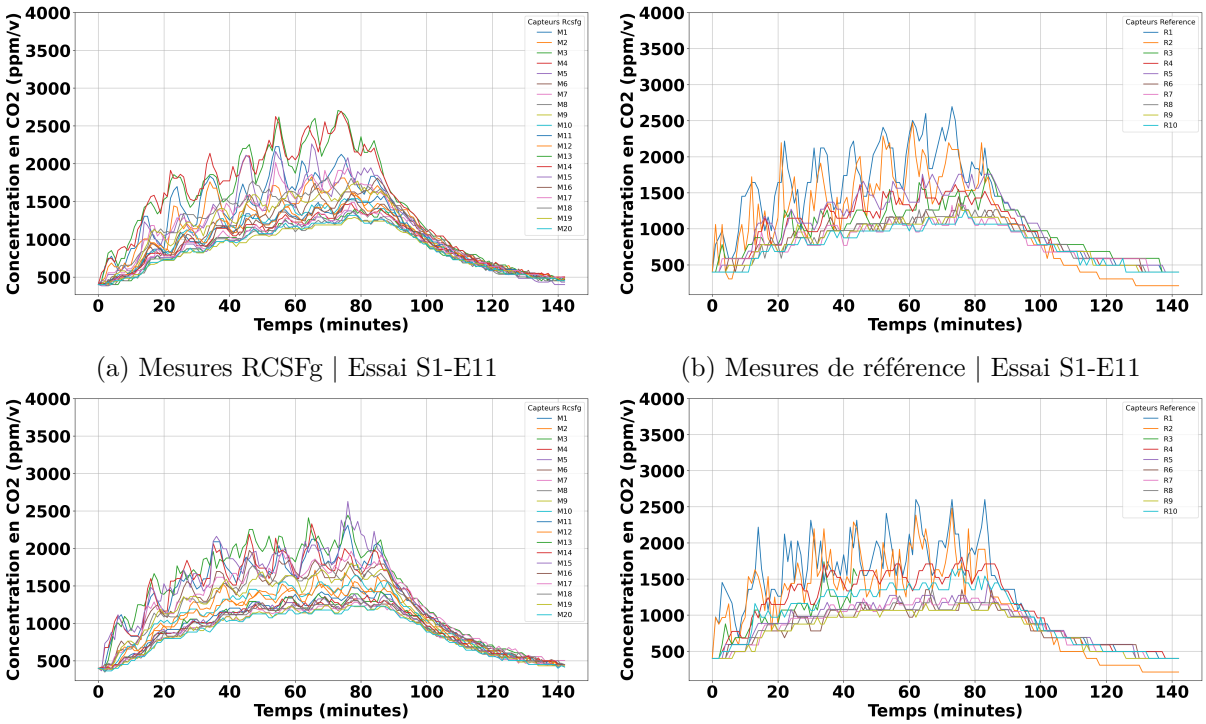
(b) Mesures de référence | Essai S1-E9

(c) Mesures RCSFg | Essai S1-E10

(d) Mesures de référence | Essai S1-E10

FIGURE 4.17 – Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) des essais S1-E9 et S1-E10

ment des capteurs, ils reproduisent globalement les tendances observées avec les instruments



(a) Mesures RCSFg | Essai S1-E11

(b) Mesures de référence | Essai S1-E11

(c) Mesures RCSFg | Essai S1-E13

(d) Mesures de référence | Essai S1-E13

FIGURE 4.18 – Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) des essais S1-E11 et S1-E13

de référence, bien que des écarts ponctuels apparaissent, notamment au niveau des amplitudes maximales et de la dispersion inter-capteurs. Ces résultats indiquent que, indépendamment de la grille d'échantillonnage (grille régulière ou irrégulière), le réseau de capteurs est en mesure de restituer de manière cohérente les dynamiques temporelles du CO_2, avec une variabilité plus marquée que celle des mesures de référence.

L'évaluation des performances des trois approches de reconstruction spatio-temporelle (krigeage, régularisation tensorielle et PINNs) a été menée sur les quatre configurations d'échantillonnage (S1-E9, S1-E10, S1-E11 et S1-E13). Les métriques utilisées sont la RMSE et le R^2, calculées pour chacun des dix points de référence (R1–R10).

Les résultats synthétisés dans les tableaux 4.4, 4.5 et 4.6 mettent en évidence une nette amélioration de la qualité de reconstruction lorsque les méthodes intègrent une contrainte physique ou de régularisation. Globalement, la méthode de krigeage présente des RMSE moyens compris entre 132,78 ppm et 1192,38 ppm, avec plusieurs valeurs négatives de R^2 indiquant une forte

TABLEAU 4.4 – Analyse des métriques (RMSE et R^2) pour les 10 points de référence (R1-R10) pour les essais S1-E9, S1-E10, S1-E11 et S1-E13 avec la méthode de krigeage

Référence	S1-E9		S1-E10		S1-E11		S1-E13	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	221.03	0.923	210.69	0.928	1192.38	-2.380	480.90	0.396
R2	348.46	0.838	240.09	0.916	606.68	0.053	615.76	0.005
R3	198.35	0.913	331.97	0.695	236.89	0.510	479.95	-1.546
R4	246.38	0.818	310.48	0.665	242.40	0.622	177.45	0.846
R5	371.33	0.646	114.98	0.960	192.66	0.807	334.42	-0.382
R6	157.18	0.883	134.84	0.906	284.15	-0.123	357.48	-0.898
R7	369.08	0.046	323.97	0.248	231.17	0.102	292.97	0.014
R8	131.10	0.918	144.45	0.891	175.71	0.646	201.24	0.481
R9	233.84	0.710	184.14	0.807	218.35	0.346	216.44	0.322
R10	190.84	0.816	132.78	0.898	606.75	-4.684	202.88	0.725

TABLEAU 4.5 – Analyse des métriques (RMSE et R^2) pour les 10 points de référence (R1-R10) pour les essais S1-E9, S1-E10, S1-E11 et S1-E13 avec la méthode de régularisation

Référence	S1-E9		S1-E10		S1-E11		S1-E13	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	188.05	0.944	238.22	0.908	361.93	0.689	317.04	0.737
R2	301.66	0.879	244.83	0.913	523.88	0.294	556.28	0.188
R3	222.77	0.891	295.62	0.758	171.06	0.745	209.17	0.516
R4	345.58	0.642	369.94	0.524	266.66	0.542	196.11	0.812
R5	369.61	0.649	124.26	0.953	148.23	0.886	400.51	-0.983
R6	123.22	0.928	80.56	0.967	221.36	0.319	342.92	-0.747
R7	293.75	0.395	319.62	0.269	236.91	0.057	282.15	0.085
R8	98.11	0.954	144.98	0.890	211.92	0.485	214.56	0.410
R9	512.69	-0.396	320.60	0.416	228.26	0.285	261.20	0.013
R10	373.98	0.292	192.52	0.786	363.18	-1.036	197.52	0.739

TABLEAU 4.6 – Analyse des metriques (RMSE et R^2) pour les 10 points de référence (R1-R10) pour les essais S1-E9, S1-E10, S1-E11 et S1-E13 avec la méthode de Pinns

Référence	S1-E9		S1-E10		S1-E11		S1-E13	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	197.62	0.938	225.71	0.917	295.86	0.792	386.74	0.609
R2	310.89	0.871	245.81	0.912	494.78	0.370	354.62	0.670
R3	338.15	0.748	247.33	0.830	338.61	-0.001	157.32	0.726
R4	306.34	0.719	221.29	0.830	175.39	0.802	135.67	0.910
R5	191.84	0.905	113.76	0.961	156.38	0.873	282.50	0.014
R6	88.70	0.963	111.38	0.936	215.51	0.354	256.85	0.020
R7	260.76	0.524	260.63	0.514	240.47	0.028	197.34	0.553
R8	95.63	0.956	87.68	0.960	95.55	0.895	97.22	0.879
R9	103.98	0.943	111.07	0.930	111.42	0.830	93.48	0.874
R10	66.38	0.978	108.13	0.932	124.65	0.760	178.21	0.788

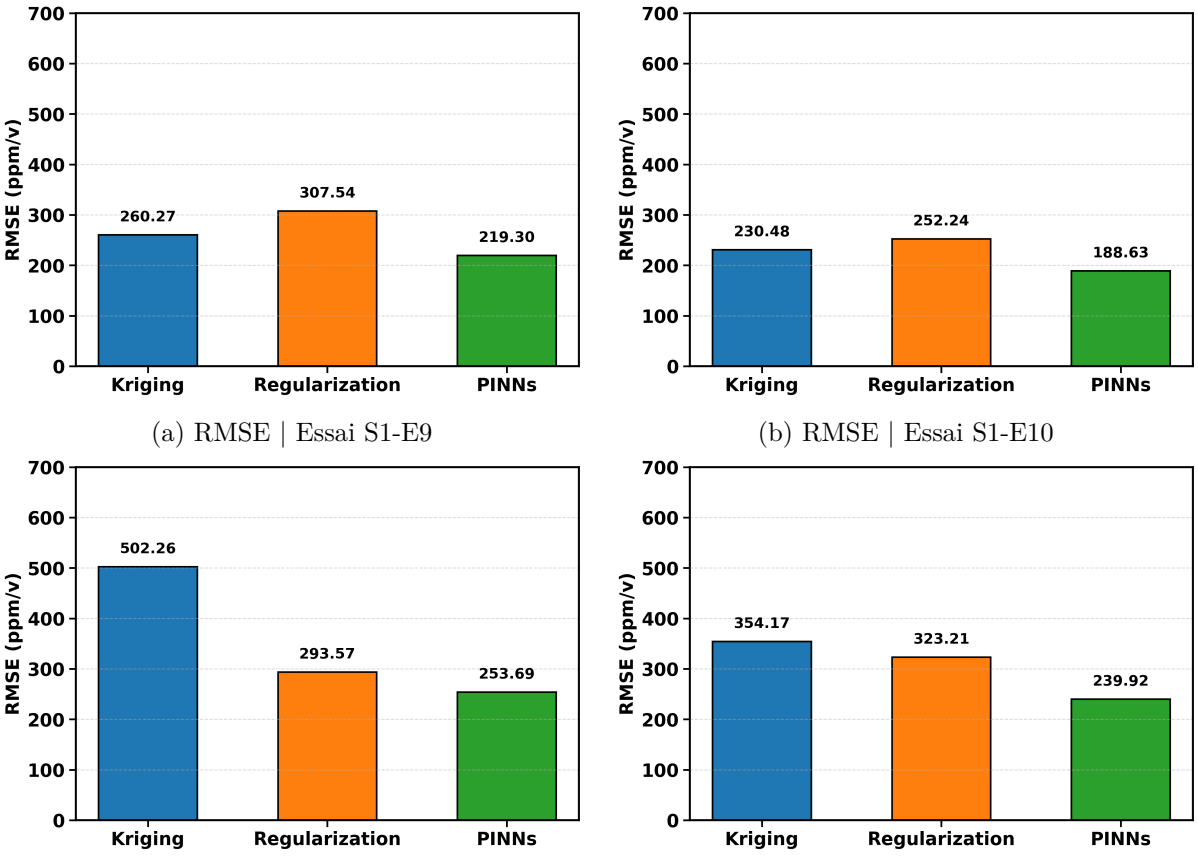
déviation des mesures observées (notamment pour R5, R6 et R10 de l'essai S1-E11).

Ces résultats traduisent les limites du krigeage lorsqu'il est appliqué à des réseaux peu dense ou irréguliers : la configuration spatiale des points conditionne directement l'erreur de prédiction et l'incertitude, un effet documenté par les travaux sur la conception optimale de réseaux de mesure ; de plus, en présence de non-stationnarité (*e.g.,* gradients locaux forts, conditions limites et dynamiques de diffusion variables), les modèles stationnaires peinent à restituer fidèlement les variations spatiales, d'où la nécessité d'approches dédiées (covariances non stationnaires, déformation spatiale) ou d'intégrer des contraintes physiques.

La régularisation tensorielle, en intégrant une pénalisation sur la variabilité spatiale, améliore significativement la stabilité des reconstructions. Les RMSE observés (80,56 ppm à 523,88 ppm) et les coefficients R^2 atteignant jusqu'à 0,967 (R6, S1-E10) confirment cette tendance. Ce comportement rejoint les conclusions selon lesquelles la régularisation permet d'atténuer la propagation du bruit et des effets d'échantillonnage irrégulier, en particulier dans les problèmes inverses mal conditionnés.

Les résultats obtenus avec les PINNs (Tableau 4.6) révèlent une amélioration supplémentaire, avec des RMSE généralement inférieurs à 200 ppm et des coefficients R^2 supérieurs à 0,9, notamment pour les points R4, R5, R8, R9 et R10. Ces performances confirment la capacité des PINNs à intégrer de manière cohérente la physique du transport diffusif dans le processus d'apprentissage, permettant une généralisation plus robuste, même en présence d'un échantillonnage irrégulier. Ces observations concordent avec les travaux récents de [Raissi19] et [Karniadakis21], qui démontrent l'efficacité des PINNs pour l'estimation de champs spatio-temporels continus lorsque les données d'entrée sont partiellement observées ou bruitées.

De façon générale, l'analyse transversale des essais (Figure 4.19) montre que la grille de



(a) RMSE | Essai S1-E9

(b) RMSE | Essai S1-E10

(c) RMSE | Essai S1-E11

(d) RMSE | Essai S1-E13

FIGURE 4.19 – Comparaison des performances (RMSE) des méthodes de krigeage, de régularisation tensorielle et de PINNs, calculées sur les 10 points de référence (R1–R10) et sur la durée totale des essais S1–E9, S1–E10, S1–E11 et S1–E13.

positionnement influence directement la précision des reconstructions. Les essais réalisés avec la grille régulière (S1-E9) présentent une variabilité spatiale plus homogène, traduite par des RMSE plus faibles pour l'ensemble des méthodes (260,27 ppm en krigeage, 247,54 ppm en régularisation et 219,30 ppm avec les PINNs). À l'inverse, dans les essais avec des grilles irrégulières (S1-E10, S1-E11 et S1-E13) l'on obtient une augmentation du RMSE, particulièrement marquée pour le krigeage (jusqu'à 752,24 ppm pour S1-E10). Cette tendance s'explique par la perte de représentativité spatiale induite par les zones sous-échantillonnées, qui limite la capacité des interpolateurs purement statistiques à restituer fidèlement les gradients locaux. Les PINNs, montrent une moindre sensibilité à l'irrégularité du maillage. Leur formulation contraint la solution à satisfaire les équations de diffusion, ce qui permet une régularisation implicite des zone à faible échantillonnage.

4.3.3 Influence du nombre de points de mesure

L'évaluation de l'influence du nombre de points de mesure sur la qualité des cartographies spatio-temporelles des concentrations de CO_2 constitue un enjeu majeur pour l'optimisation des réseaux de capteurs. En effet, la densité d'échantillonnage conditionne directement la précision des reconstructions tout en impactant le coût d'instrumentation et la complexité opérationnelle. L'objectif de cette section est d'analyser et de quantifier la dégradation de la précision des cartographies spatio-temporelles lorsque la densité diminue et d'identifier dans notre contexte le seuil minimal permettant de garantir une représentation fidèle des distributions spatio-temporelles du polluant, tout en identifiant la sensibilité des méthodes de reconstruction (krigeage, régularisation tensorielle et PINNs) à une réduction progressive du nombre de points.

Démarche expérimentale

L'analyse de l'influence du nombre de points de mesure sur les cartographies spatio-temporelles générées a été réalisée dans un premier temps à l'aide de données simulées, en considérant différentes configurations avec 10, 15, 20 et 25 points de mesures (Voir section 2.6).

Ces analyses préliminaires ont montré que l'utilisation de 20 points de mesures avec une bonne couverture spatiale dans notre espace expérimental représente le compromis le plus pertinent entre la précision de la reconstruction spatio-temporelle et le coût d'instrumentation. Sur cette base, l'évaluation a été transposée aux données expérimentales collectées en partant d'un essai sur la grille régulière (Voir figure 4.16a), qui est une configuration homogène avec des points de mesures équidistants. Cette grille d'échantillonnage a été choisie comme grille de base pour notre analyse suite aux résultats issus de la section précédente sur l'influence de la grille de positionnement sur les cartographies spatio-temporelles (Voir section 4.3.2) qui montre que qu'elle est la plus performante comparativement aux autres grilles d'échantillonnage irrégulières testées, en termes de qualité de la reconstruction et de stabilité des estimations spatiales.

Afin de quantifier l'effet de la densité du réseau, nous partons de l'essai S1-E9 réalisé sous la grille d'échantillonnage régulière (Voir figure 4.2a) avec 20 modules capteurs sur une surface de 50 m^2 (Chambre expérimentale de dimension 10 m $\times$ 5 m), soit une densité de déploiement de 0,4 module capteur /m^2. Cette configuration constitue la référence du dispositif expérimental. Par la suite, sous les mêmes conditions expérimentales, trois grilles d'échantillonnages ont été définies par suppression aléatoire de certains modules capteurs tout en maintenant une répartition spatiale homogène, conduisant à des densités respectives de 0,3 , 0,2 et 0,1 module capteur/m^2 pour des réseaux de 15, 10 et 5 points (figures 4.20a–4.20c).



(a) Grille régulière avec 15 points de mesures



(b) Grille régulière avec 10 points de mesures



(c) Grille régulière avec 5 points de mesures

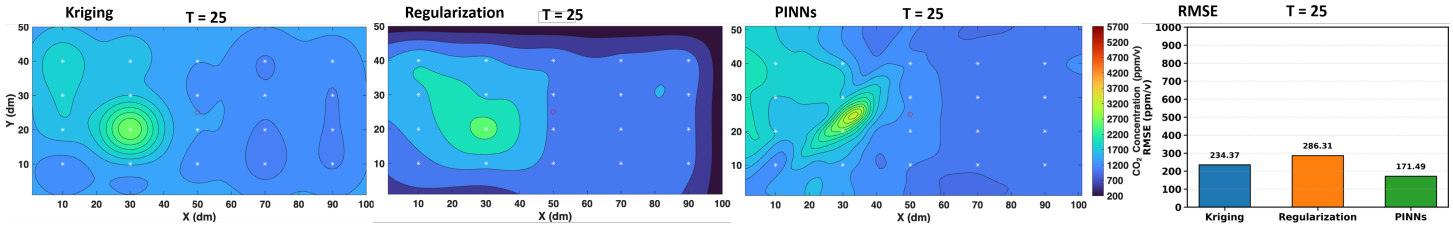
Pour évaluer l'influence du nombre de points de mesure sur la qualité des reconstructions spatio-temporelles, les données expérimentales issues des grilles d'échantillonnage (15, 10 et 5 points de mesures) ont été utilisées pour générer les cartographies spatio-temporelles des concentrations de CO_2 avec les trois méthodes (krigeage, régularisation tensorielle et PINNs). Les performances pour chacune des configurations (RMSE et R^2) ont ensuite été quantifiées à partir des dix points de référence puis comparées à celles obtenues avec la grille initiale à 20 points. Cette analyse permet d'examiner conjointement la structure spatiale des cartographies reconstruites et l'évolution des performances selon la densité d'échantillonnage.

Résultats et discussions

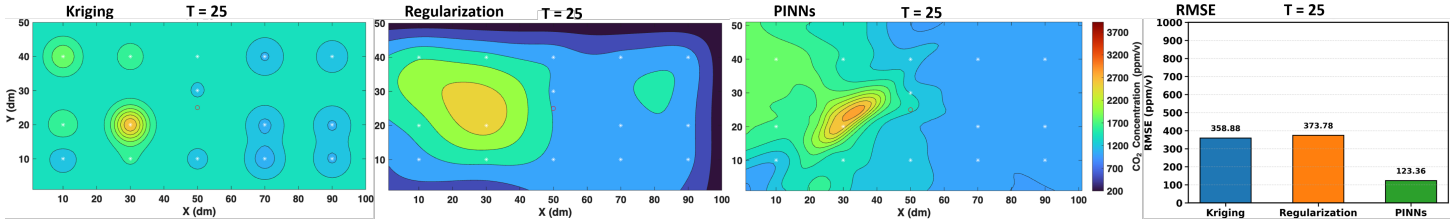
Les cartographies spatio-temporelles de la Figure 4.21 illustrent l'influence du nombre de points de mesure sur la qualité des reconstructions à $T = 25$ min (Phase de pollution) pour les trois méthodes considérées. Pour la grille initiale à 20 points (Figure 4.21a), les trois méthodes restituent globalement la distribution du CO_2, avec une cohérence spatiale satisfaisante.

Lorsque la densité d'échantillonnage diminue (15 et 10 points), la dégradation progressive de la précision est clairement perceptible. Les reconstructions obtenues par krigeage deviennent plus diffuses, traduisant une perte de résolution spatiale, tandis que la régularisation tensorielle montre une sur-lissage accentué des zones de forte concentration. Les PINNs conservent, dans ces configurations intermédiaires, une meilleure continuité spatiale et une représentation plus fidèle des zones à forte concentration, confirmée par des valeurs de RMSE plus faibles (Voir figure 4.21b, 4.21c).

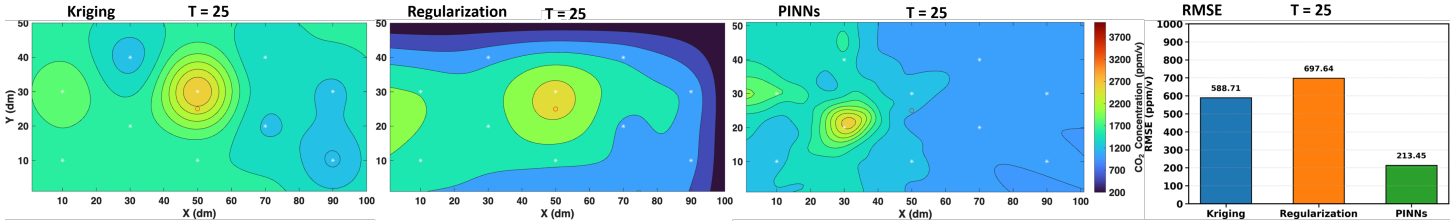
Avec seulement 5 points de mesure, les trois méthodes perdent en capacité de représentation des concentrations. Les cartes présentent des artefacts marqués et des concentrations atténuées, indiquant une forte sensibilité à la sous-détermination de l'échantillonnage. Dans cette configuration, les performances chutent significativement pour toutes les méthodes de cartographie (Voir figure 4.21d).



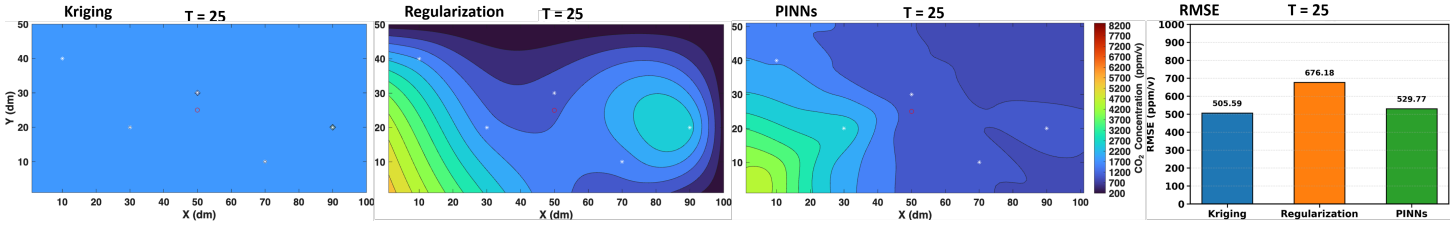
(a) $T = 25$ min | Essai S1-E9 avec 20 points de mesures



(b) $T = 25$ min | Essai S1-E9 avec 15 points de mesures



(c) $T = 25$ min | Essai S1-E9 avec 10 points de mesures



(d) $T = 25$ min | Essai S1-E9 avec 5 points de mesures

FIGURE 4.21 – Cartographies spatio-temporelles des concentrations de CO_2 et évaluation du RMSE à $T = 25$ min de l'essai S1-E9 avec 20, 15, 10 et 5 points de mesures

Les résultats obtenus (RMSE et R^2) pour chacun des 10 points de référence sont synthétisés dans les tableaux 4.7-4.9.

TABLEAU 4.7 – Comparaison des performances (RMSE et R^2) pour 10 points de référence – Krigeage

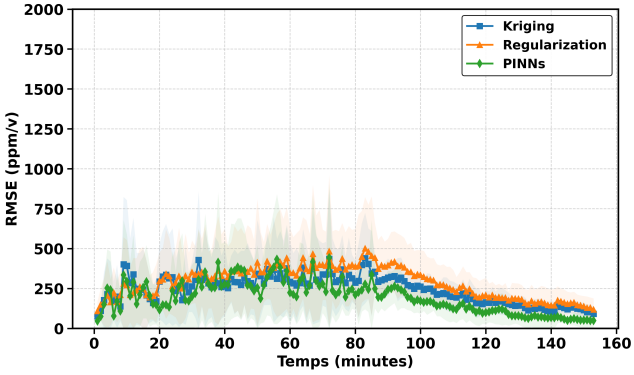
Réf.	20 points		15 points		10 points		5 points	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	221.03	0.923	269.72	0.885	232.54	0.914	254.18	0.898
R2	348.46	0.838	448.32	0.732	354.49	0.832	318.34	0.865
R3	198.35	0.913	196.46	0.915	301.14	0.800	303.18	0.797
R4	246.38	0.818	216.40	0.860	385.94	0.554	402.72	0.514
R5	371.33	0.646	267.92	0.815	427.58	0.530	409.31	0.569
R6	157.18	0.883	230.17	0.749	770.03	-1.811	558.86	-0.481
R7	369.08	0.046	463.87	-0.507	833.23	-3.864	791.09	-3.384
R8	131.10	0.918	224.73	0.760	565.18	-0.519	580.79	-0.604
R9	233.84	0.710	308.70	0.494	532.65	-0.506	630.01	-1.108
R10	190.84	0.816	310.08	0.514	331.66	0.443	644.23	-1.100

TABLEAU 4.8 – Comparaison des performances (RMSE et R^2) pour 10 points de référence – Regularization

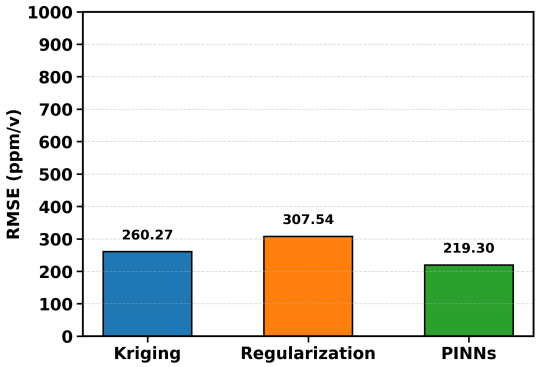
Réf.	20 points		15 points		10 points		5 points	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	188.05	0.944	178.45	0.949	480.78	0.633	291.90	0.865
R2	301.66	0.879	451.51	0.728	321.98	0.862	422.87	0.762
R3	222.77	0.891	211.19	0.902	223.87	0.890	1211.46	-2.235
R4	345.58	0.642	429.96	0.446	326.91	0.680	184.89	0.898
R5	369.61	0.649	386.27	0.616	449.48	0.481	624.44	-0.002
R6	123.22	0.928	175.12	0.855	662.61	-1.082	375.92	0.330
R7	293.75	0.395	288.00	0.419	843.54	-3.985	851.09	-4.075
R8	98.11	0.954	205.40	0.799	707.83	-1.383	791.81	-1.982
R9	512.69	-0.396	609.44	-0.972	903.69	-3.336	608.05	-0.963
R10	373.98	0.292	384.38	0.253	651.46	-1.147	776.59	-2.051

TABLEAU 4.9 – Comparaison des performances (RMSE et R^2) pour 10 points de référence – PINNs

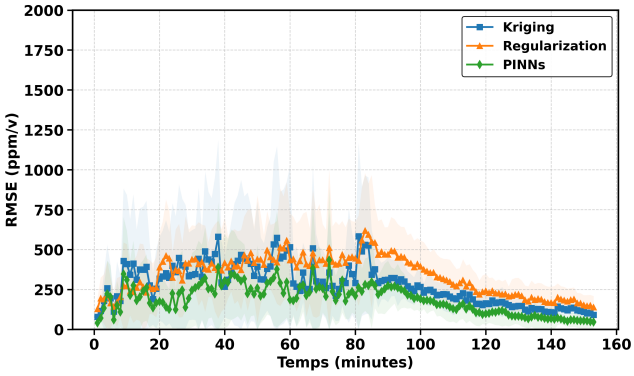
Réf.	20 points		15 points		10 points		5 points	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	197.62	0.938	201.27	0.936	362.69	0.791	189.00	0.943
R2	310.89	0.871	287.11	0.890	326.96	0.857	393.62	0.793
R3	338.15	0.748	318.75	0.776	370.86	0.697	997.82	-1.194
R4	306.34	0.719	287.07	0.753	322.81	0.688	143.10	0.939
R5	191.84	0.905	185.49	0.912	265.47	0.819	389.37	0.610
R6	88.70	0.963	86.08	0.965	96.38	0.956	100.88	0.952
R7	260.76	0.524	258.56	0.532	308.17	0.335	262.33	0.518
R8	95.63	0.956	97.14	0.955	77.18	0.972	89.77	0.962
R9	103.98	0.943	103.81	0.943	86.00	0.961	101.51	0.945
R10	66.38	0.978	78.08	0.969	110.80	0.938	83.14	0.965



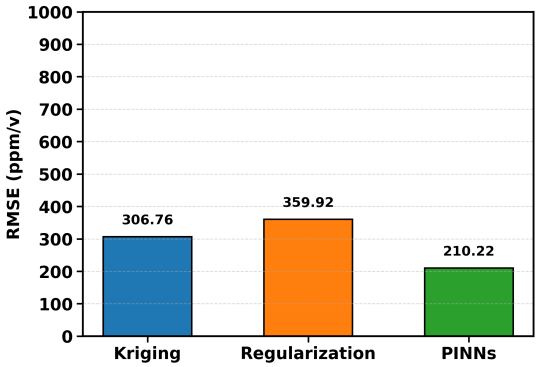
(a) Évolution temporelle RMSE | Essai S1-E9 avec 20 points de mesures



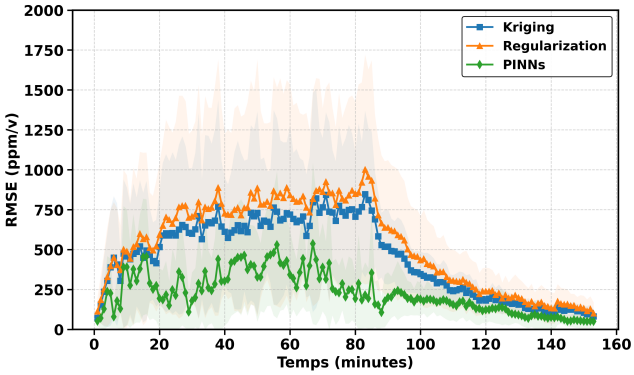
(b) RMSE | Essai S1-E9 avec 20 points de mesures



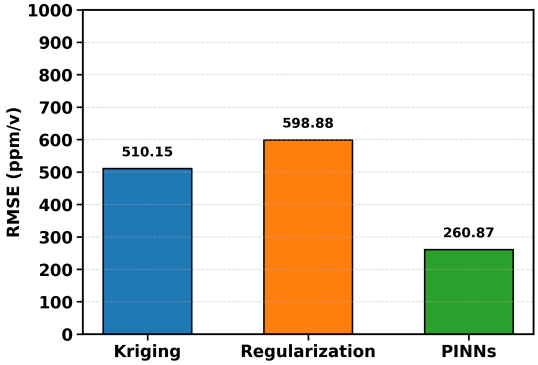
(c) Évolution temporelle RMSE | Essai S1-E9 avec 15 points de mesures



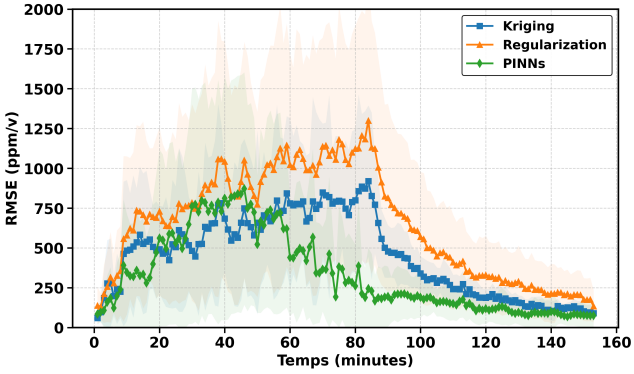
(d) RMSE | Essai S1-E9 avec 15 points de mesures



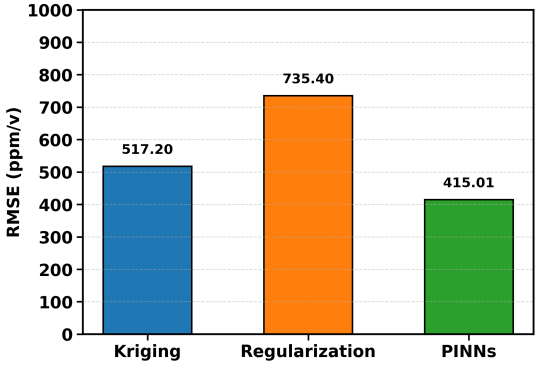
(e) Évolution temporelle RMSE | Essai S1-E9 avec 10 points de mesures



(f) RMSE | Essai S1-E9 avec 10 points de mesures



(g) Évolution temporelle RMSE | Essai S1-E9 avec 5 points de mesures



(h) RMSE | Essai S1-E9 avec 5 points de mesures

FIGURE 4.22 – Évolution temporelle du RMSE obtenues avec les méthodes de krigeage, de régularisation et de PINNs pour la reconstruction spatio-temporelle des concentrations de CO_2 de l'essai S1-E9 avec 20, 15, 10 et 5 points de mesures

Les résultats des quatre scénarios (20, 15, 10 et 5 points) confirment que la densification du réseau de mesure améliore la fidélité des reconstructions spatio-temporelles, mais avec une sensibilité très marquée au choix de la méthode (Voir Figures 4.21–4.22 ; Tableaux 4.7–4.9).

Krigage Avec 20 points (grille régulière), le krigeage fournit des estimations acceptables sur une majorité de références (par ex. $R^2 > 0.9$ pour R1, R3, R6, R8, R10). Lorsque la densité décroît, la dégradation est rapide : dès 15 points, plusieurs R^2 chutent (R7 devient négatif), et avec 10 puis 5 points, on observe des $R^2 < 0$ et des RMSE très élevés (par exemple. R6 : 770 ppm avec 10 points ; R7 : $R^2 = -3.864$ avec 10 points). Les cartes à $T = 25$ min montrent des gradients mal restitués et des artefacts spatiaux (Figures 4.21c–4.21d). Cette sensibilité au plan d'échantillonnage est cohérente avec la littérature sur la conception de réseaux et la prédiction géostatistique, qui montre que la performance du krigeage dépend fortement de la configuration spatiale et de la densité des points, et qu'une hétérogénéité de couverture peut induire des biais et une variabilité accrue de l'erreur.

Régularisation tensorielle La régularisation tensorielle améliore la stabilité des reconstructions par rapport au krigeage pour des densités moyennes (20–15 points) : on observe, par exemple, des R^2 élevés (jusqu'à 0,949 sur R1 avec 15 points) et des RMSE réduits sur plusieurs références (Voir tableau 4.8). Néanmoins, lorsque le réseau devient très parcimonieux (10 puis 5 points), la méthode perd en robustesse, avec de nombreux R^2 négatifs (R7–R10) et des RMSE élevés (R3 : 1211 ppm avec 5 points). Lorsque la densité de l'échantillonnage devient faible (10 puis 5 points), la méthode conserve une certaine continuité spatio-temporelle, mais sa capacité à restituer fidèlement les gradients locaux se dégrade, comme l'indiquent les R^2 négatifs observés pour R7 à R10 et les RMSE élevés (jusqu'à 1211 ppm pour R3 avec 5 points). Cette perte de précision traduit le fait que, malgré la régularisation, l'information que l'on peut tirer dans les données devient insuffisante pour contraindre localement la solution.

Un effet de bord observé dans les extrémités supérieures et droite de la carte spatiale (concentrations mal estimées). Cela provient de la structure unilatérale des opérateurs de dérivation discrets utilisés dans la régularisation. En effet, les différences finies sont orientées vers le bas et la gauche ce qui implique que les valeurs situées sur la première ligne et la dernière colonne ne sont pas régularisées par des différences « en amont ». Ces zones restent donc faiblement contraintes, engendrant un effet de bord asymétrique caractéristique lorsque la pénalisation est appliquée séparément sur les lignes et sur les colonnes.

PINNs La méthode PINNs présentent la meilleure robustesse lorsque le nombre de points diminue. À 20 et 15 points, les performances sont élevées et homogènes ($R^2 \geq 0.9$ pour R1, R5, R6, R8–R10 ; RMSE < 200 ppm). Avec 10 points, la dégradation reste modérée sur la plupart

des références (par ex. R6 : $R^2 = 0.956$; R8 : $R^2 = 0.972$; R9 : $R^2 = 0.961$) alors que le krigeage et la régularisation se détériorent fortement. Même avec 5 points, plusieurs valeurs de référence demeurent bien restituées (R1, R4, R6, R8–R10 avec $R^2 \approx 0.94$–0.97), à l'exception notable de R3 (RMSE $\approx$ 998 ppm, $R^2 < 0$), illustrant une limite locale lorsque la physique seule ne compense pas une lacune d'information majeure. L'évolution temporelle du RMSE (Voir figure 4.22) montre que les erreurs des PINNs restent stables durant les phases transitoires (montée de concentration et début d'évacuation), alors que celles des approches purement statistiques explosent avec un faible échantillonnage des points. Ce comportement est cohérent avec le principe des PINNs, qui contraignent l'approximation à satisfaire les équations de transport/diffusion (termes physiques dans la fonction de coût), agissant comme une régularisation informée par la physique et améliorant la généralisation sous données parcimonieuses.

De manière générale, l'analyse comparative menée sur les quatre configurations d'échantillonnage met en évidence l'influence déterminante de la densité du réseau de mesure sur la qualité des reconstructions spatio-temporelles. La réduction progressive du nombre de points de mesure de vingt à quinze, dix puis cinq capteurs induit une perte graduelle d'information spatiale, se traduisant par une dégradation des performances globales, tant en termes de fidélité des champs locaux que de stabilité temporelle. Lorsque la densité passe de 0,4 à 0,3 module capteur/m², la précision des estimations demeure globalement satisfaisante, indiquant qu'une légère diminution de la densité peut être tolérée sans altérer significativement la représentation des distributions.

En revanche, au-dessous de ce seuil, c'est-à-dire pour des densités de 0,2 et 0,1 module capteur/m², la diminution de l'échantillonnage spatial provoque une dégradation marquée des reconstructions : perte de cohérence spatiale, apparition de zones sous-échantillonnées et amplification des erreurs. Ces effets sont particulièrement visibles lors des phases transitoires du champ de concentration (voir figure 4.22), où la dynamique temporelle rapide du CO_2 nécessite une couverture dense pour assurer une représentation fidèle des variations spatio-temporelles. Les écarts de performance observés entre configurations traduisent donc principalement la capacité de chaque densité d'échantillonnage à capturer la variabilité du phénomène, plutôt que des différences intrinsèques entre méthodes de reconstruction.

Ces observations s'inscrivent dans les constats déjà rapportés dans la littérature. À titre d'exemple, une densité de $1,6 \times 10^2$ a été calculée dans l'air intérieur d'un atelier de menuiserie [Li18] contre 10^{-6} dans l'air extérieur [Shin24]. Toutefois, aucune valeur de référence universelle ne peut être définie, la densité optimale dépendant du contexte expérimental et de l'environnement étudié [Akyildiz02], [Mansouri10]. Il est donc difficile de fixer un seuil de densité applicable à l'ensemble des configurations.

Ainsi, les résultats obtenus indiquent qu'un échantillonnage de vingt points de mesure, correspondant à une densité de 0,4 module capteur/m² avec une couverture spatiale homogène,

constitue un compromis optimal entre la précision de la reconstruction et le coût d'instrumentation. En dessous de cette densité, la stabilité et la cohérence des cartographies ne peuvent être maintenues qu'au prix de l'introduction de contraintes physiques explicites dans les modèles de reconstruction. Ces observations soulignent enfin l'importance d'une conception rigoureuse du réseau de capteurs : minimiser les zones sous-échantillonnées et garantir une couverture homogène de l'espace, condition essentielle à la fiabilité des cartographies spatio-temporelles.

4.3.4 Influence des caractéristiques de la source de pollution

Les caractéristiques de la source du polluant dans l'atmosphère (*i.e.* nature du polluant émis, orientation et position de la source, nombre de sources, etc.) est un facteur important à prendre en considération dans l'étude des cartographies spatio-temporelles. Pour autant, ces caractéristiques ont été assez peu abordées dans les travaux publiés [Ash'aari20]. L'objectif de la présente section est d'évaluer l'influence des caractéristiques de la source sur la performance des modèles de cartographie spatio-temporelle pour la reconstitution de la dynamique de concentration en polluant. Dans ce but, plusieurs essais expérimentaux ont été menés en faisant varier la grille d'échantillonnage ainsi que les paramètres de la source d'émission. Ces paramètres ont été choisis de manière à explorer différentes configurations représentatives des conditions d'émission rencontrées en milieu de travail.

Les essais expérimentaux considérés dans cette section pour étudier l'influence de l'orientation de la source, du débit d'émission et du nombre de sources ont été menées avec le CO_2 comme polluant modèle. Un essai a été mené avec des aérosols de particules fines afin de comparer leur dynamique de diffusion à celle du CO_2 et d'étudier l'influence du type de polluant sur les cartographies spatio-temporelles obtenues.

4.3.4.1 Influence de l'orientation de la source

L'orientation de la source d'émission constitue un paramètre déterminant dans la dynamique de dispersion des polluants en environnement intérieur. En effet, la direction initiale du flux de gaz influence directement les trajectoires de transport, les zones d'accumulation. Dans le cadre de cette étude, l'objectif est d'évaluer l'effet de l'orientation de la source sur les cartographies spatio-temporelles reconstruites des concentrations en CO_2, obtenues à partir du réseau de capteurs RCSFg. Pour ce faire, deux essais représentatifs de la série 1 ont été sélectionnés : S1-E1, correspondant à une émission de CO_2 dirigée vers le bas, et S1-E2, correspondant à une émission orientée vers le haut. Ces deux configurations ont été choisies afin d'analyser, à environnement et conditions expérimentales identiques, les différences induites sur la cartographie spatio-temporelle des concentration en polluant et sur la performance des modèles de reconstruction utilisés (Krigeage, Régularisation tensorielle et PINNs).

Démarche expérimentale

Les essais S1-E1 et S1-E2 ont été réalisés en août 2023 dans la chambre expérimentale (Cf section 3.3.1) suivant la grille d'échantillonnage à la figure 4.23 et conformément aux plans de calibration et d'expérience A décrits dans la section 3.7.1 et 3.4.4.



FIGURE 4.23 – Grille d'échantillonnage régulière avec les modules capteurs du RCSFg en bleu (M1-M20), les détecteurs de référence en vert (R1-R10) et la source de pollution en rouge S1

Ainsi, seule l'orientation de la source S1 a varié, avec une émission avec un débit de 12,5 L/min dirigée vers le bas (S1-E1) et vers le haut (S1-E2) (Voir figure 4.24) afin de quantifier les différences induites dans la dynamique de dispersion et d'évaluer la sensibilité du modèle de reconstruction à ce paramètre. Les conditions de température et d'humidité relative de l'air au cours de ces essais était respectivement en moyenne 22 ^{o}C et 50 %.

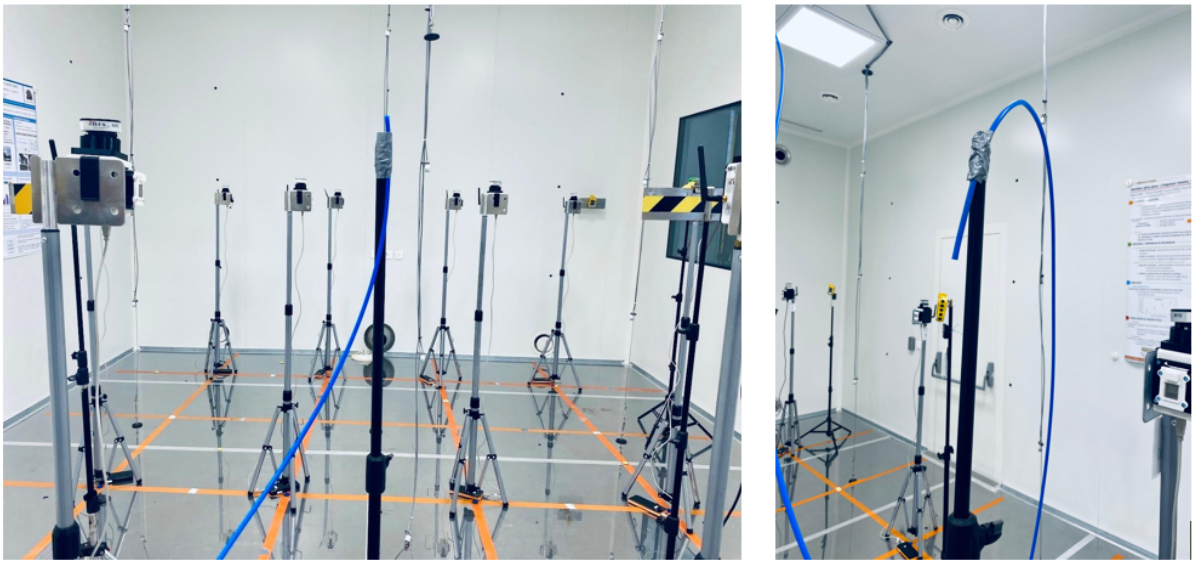


FIGURE 4.24 – Dispositif expérimental et source installés pour les essais S1-E1 à gauche (Source orientée vers le bas) et S1-E2 à droite (Source orientée vers le haut) dans la chambre expérimentale

L'influence de l'orientation sur la cartographie spatio-temporelle a été étudiée par le biais

de plusieurs approches :

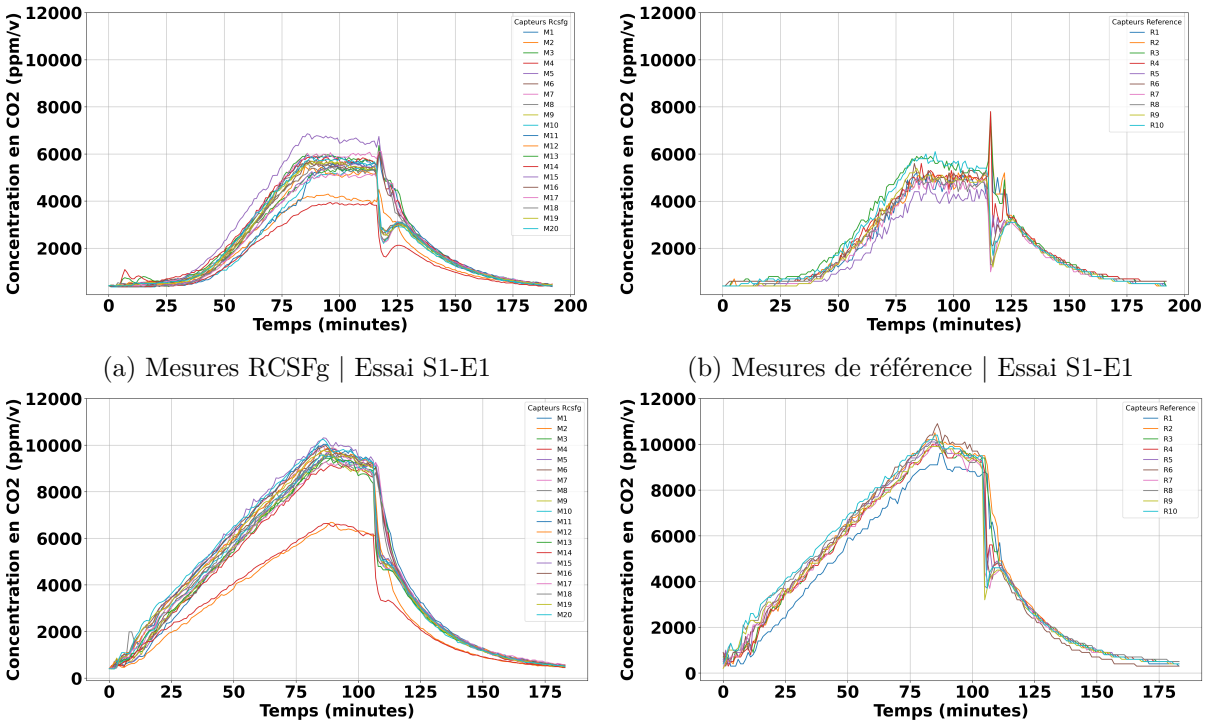
— La comparaison des profils temporels des concentrations mesurées (figure 4.25)
— La cartographie spatio-temporelle reconstruite à partir des trois méthodes considérées (figure 4.27)
— L'évaluation quantitative des performances de la cartographie par les indicateurs RMSE et R^2 aux points de référence (tableaux 4.10 et 4.11)

Résultats et discussions

Les courbes de la figure 4.25 présentent l'évolution temporelle des concentrations de CO_2 mesurées par les capteurs du RCSFg (M1–M20) et les instruments de référence (R1–R10) lors des essais S1-E1 et S1-E2. Les résultats indiquent que l'orientation de la source émettrice influence à la fois la cinétique d'évolution des concentrations en CO_2 et les niveaux atteints. En effet, lorsque la source d'émission est dirigée vers le bas (essai S1-E1), la montée en concentration enregistrée par les capteurs est plus progressive et les valeurs maximales sont modérées (Environ 7000 ppm/v). Par opposition, lorsque la source d'émission est dirigée vers le haut (essai S1-E2), la montée en concentration est plus rapide et les pics observés sont plus élevés (Environ 10000 ppm/v).

Ces résultats peuvent être attribués aux phénomènes physiques de diffusion et de convection du CO_2 dans la chambre expérimentale et au fait que la densité relative du CO_2 gazeux est de 1,53 (air = 1). Lorsque la source d'émission est dirigée vers le bas, la direction descendante du flux de CO_2 favoriserait d'abord une accumulation des gaz vers le sol, en dessous du plan de mesure des capteurs. Ensuite, le gaz remonterait par diffusion/convection à la hauteur des capteurs, ce qui produirait un léger décalage temporel dans l'évolution de la concentration. Ces phénomènes expliqueraient une élévation plus lente des concentrations enregistrées et une dispersion initiale restreinte dans l'environnement étudié. Au contraire, la configuration « source dirigée sers le haut » favoriserait un brassage plus important du CO_2, lié à une dynamique ascendante du gaz. Cette dynamique conduirait à une accumulation transitoire plus marquée à hauteur des capteurs avant dilution dans l'espace expérimental.

— Remarque 1 : Au-delà des différences induites par l'orientation de la source, les profils temporels des concentrations mettent en évidence la dynamique caractéristique du plan d'expérience A, qui se décompose en trois phases successives : Pollution ($0 < T_p < 85$ min), Stabilisation ($85 < T_s < 105$ min), Épuration ($105 < T_v < 185$ min).
— Remarque 2 : L'ensemble des modules capteurs étant étalonné selon le protocole de calibration A (Voir 3.7.1), une variabilité inter-capteurs peut subsister lors de la comparaison aux détecteurs de référence en raison de la limitation de ce plan d'étalonnage reposant sur seulement deux détecteurs de référence pour l'ensemble du réseau.



(a) Mesures RCSFg | Essai S1-E1

(b) Mesures de référence | Essai S1-E1

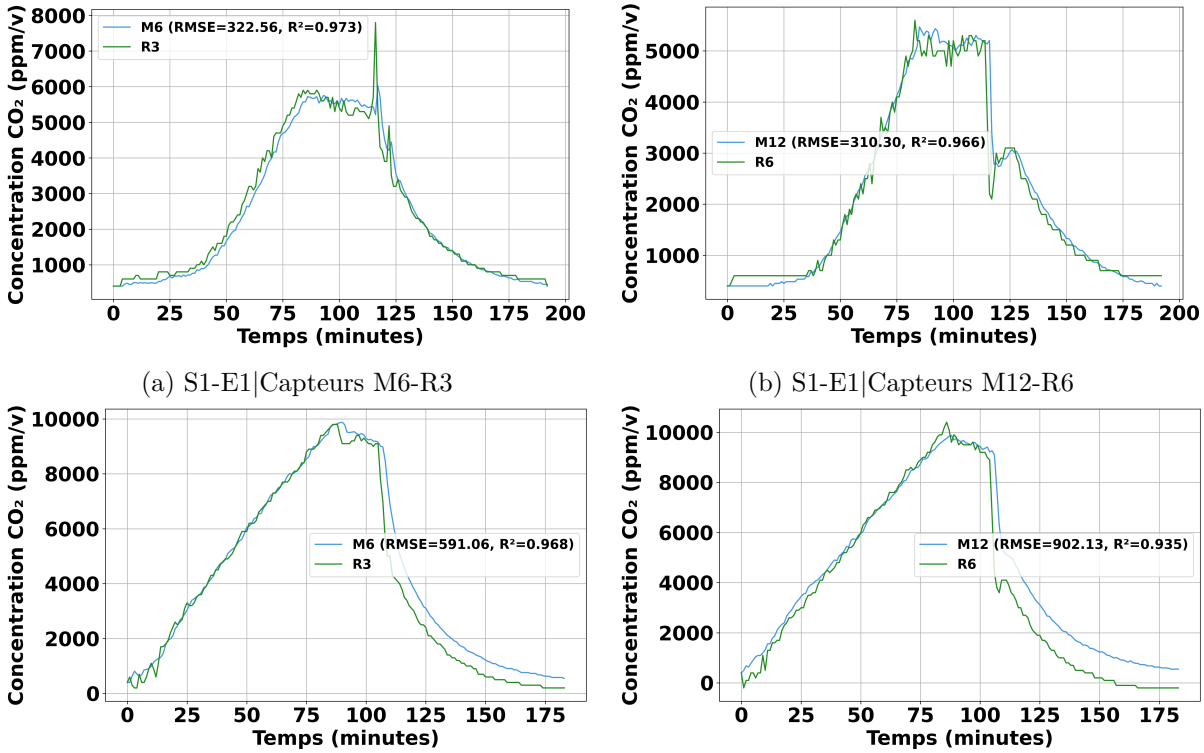
(c) Mesures RCSFg | Essai S1-E2

(d) Mesures de référence | Essai S1-E2

FIGURE 4.25 – Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) au cours des essais S1-E1 et S1-E2

La figure 4.26 présente les écarts obtenus aux points de co-localisation entre les modules capteurs et les détecteurs de référence, exprimés en RMSE et R^2. Ces résultats montrent la cohérence inter-capteurs et permettent de quantifier les biais résiduels entre les mesures du RCSFg et les mesures de référence avant l'analyse des reconstructions spatio-temporelles.

Les cartographies spatio-temporelles à $T = 35$ min (Figure 4.27) confirment les profils temporels observés avec les modules capteurs du RCSFg. Dans le cas de l'essai S1-E1 (émission dirigée vers le bas), les valeurs reconstruites de concentration demeurent globalement faibles et spatialement homogènes. Aucun « point chaud » ne permet d'identifier la position réelle de la source. Ce comportement traduit le fait que, durant cette phase expérimentale, la majeure partie du CO_2 s'est accumulée en bas de la chambre expérimentale, en dessous du plan d'échantillonnage des capteurs. Malgré une émission pendant 35 min, le transport du CO_2 à hauteur des modules capteurs reste limitée, d'où une sous-représentation des gradients de concentration sur les cartographies.



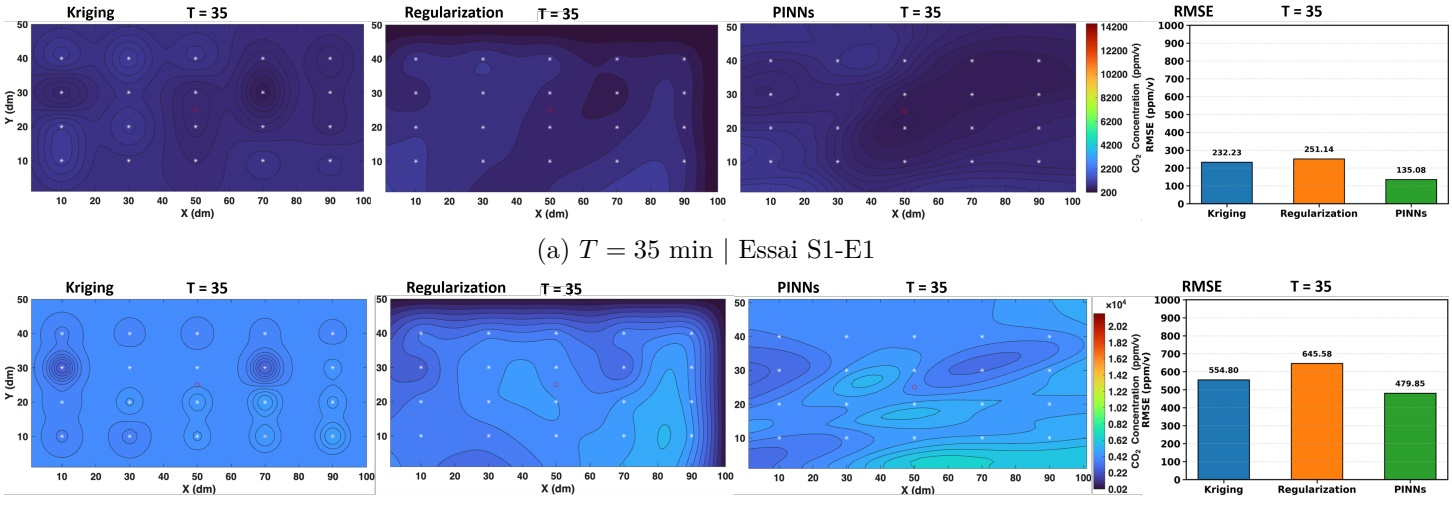
(a) S1-E1|Capteurs M6-R3

(b) S1-E1|Capteurs M12-R6

(c) S1-E2|Capteurs M6-R3

(d) S1-E2|Capteurs M12-R6

FIGURE 4.26 – Comparaison des mesures de concentrations de CO_2 enregistrées au cours des essais S1-E1 et S1-E2 aux points de co-localisation des capteurs des réseaux RCSFg et les détecteurs de référence ; M6-R3 (a), M12-R6 (b)



(a) $T = 35$ min | Essai S1-E1

(b) $T = 35$ min | Essai S1-E2

FIGURE 4.27 – Cartographies spatio-temporelles des concentrations de CO_2 et évaluation du RMSE à $T = 35$ min de l'essai a)S1-E1 et b) S1-E2

Les cartographies spatio-temporelles reconstruites à l'aide des trois méthodes (krigeage, régularisation tensorielle et PINNs) présentent des cartes globalement comparables pour les essais S1-E1 et S1-E2. Cette similarité s'explique principalement par les conditions expérimentales imposées par le plan d'expérience A, qui consiste en l'absence de ventilation active durant la phase de pollution. Dans ce contexte, l'air ambiant demeure quasi statique, limitant ainsi les effets de convection et favorisant un régime par la diffusion. Cette stabilité conduit mécaniquement à une répartition homogène du polluant et, par conséquent, à des reconstructions proches quelle que soit la méthode utilisée.

Les résultats présentés dans les tableaux 4.10 et 4.11 mettent en évidence l'impact de l'orientation de la source d'émission sur la distribution spatiale des concentrations et la précision des reconstructions spatio-temporelles à travers le RMSE et le R^2. Dans le cas de l'essai S1-E1, où la source est orientée vers le bas, les valeurs de RMSE restent globalement modérées (entre 320 et 900 ppm) et les R^2 atteignent des niveaux élevés, souvent supérieurs à 0,9. Les concentrations observées sont relativement homogènes dans l'espace. Comme l'a montré les profils temporels des essais (Voir figure 4.25), Cette configuration traduit un régime de diffusion lent et stable. Quant à l'essai S1-E2, correspondant à une émission orientée vers le haut, les valeurs de RMSE augmentent, dépassant localement 1000 ppm, tandis que les R^2 restent élevés (entre 0,90 et 0,97 selon les points). Cette hausse du RMSE reflète une plus forte variabilité spatiale des concentrations liée à la dynamique ascendante du flux de CO_2. Le gaz, émis vers le haut est à la même hauteur que le plan d'échantillonnage, ce qui donne des pics de concentration marqués dans les zones situées à côté de la source avant de se diffuser dans l'espace.

Tableau 4.10 – Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E1

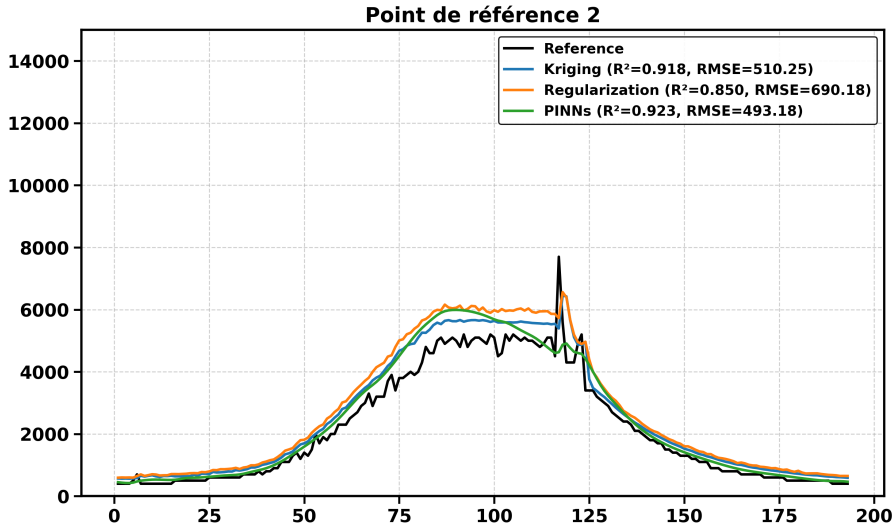
Référence	Krigeage		Regularization		PINNs	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	918.44	0.716	720.48	0.825	814.52	0.777
R2	510.25	0.918	690.18	0.850	493.18	0.923
R3	326.00	0.972	367.08	0.965	312.14	0.974
R4	389.66	0.951	692.33	0.844	563.15	0.897
R5	724.55	0.744	873.72	0.628	538.64	0.859
R6	397.41	0.945	394.34	0.946	263.41	0.976
R7	458.76	0.918	388.61	0.941	613.56	0.853
R8	580.64	0.871	1103.77	0.533	485.40	0.910
R9	490.52	0.920	934.25	0.708	318.15	0.966
R10	365.76	0.964	319.70	0.972	619.34	0.896

TABLEAU 4.11 – Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E2

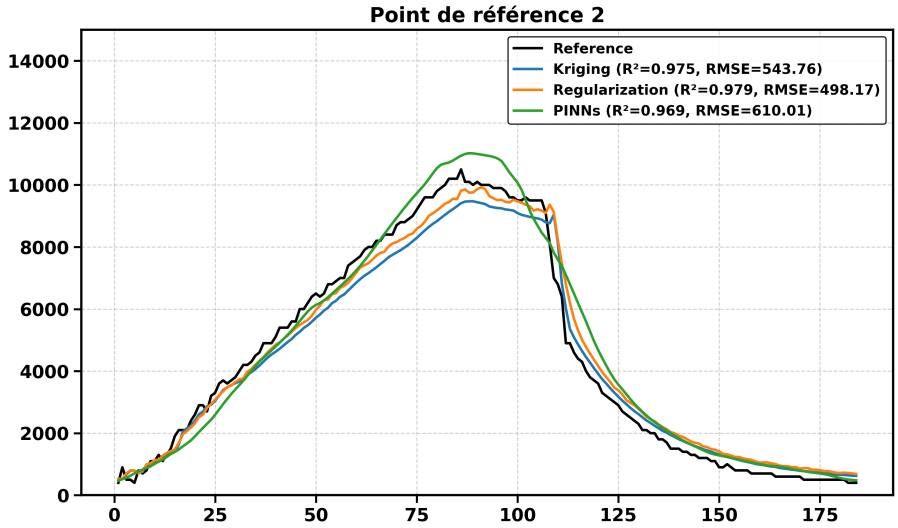
Référence	Krigeage		Regularization		PINNs	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	959.03	0.906	1637.35	0.726	971.04	0.904
R2	543.76	0.975	498.17	0.979	610.01	0.969
R3	658.96	0.961	724.51	0.952	583.68	0.969
R4	739.00	0.949	1000.52	0.907	879.20	0.928
R5	617.22	0.966	686.15	0.957	934.79	0.921
R6	978.47	0.924	974.53	0.924	787.83	0.950
R7	880.95	0.926	1304.15	0.839	2017.00	0.614
R8	804.42	0.940	1244.55	0.856	647.43	0.961
R9	772.39	0.943	1076.98	0.890	787.20	0.941
R10	769.35	0.948	756.40	0.950	656.66	0.962

Cette dynamique favorise l'apparition de zones de concentration plus prononcés, ce qui rend les distributions locales plus hétérogènes. Les différences entre les points de mesure s'en trouvent accentuées, et les reconstructions présentent des écarts plus importants aux références, surtout dans les zones de transition rapide entre forte et faible concentration (Phase de ventilation).

Ainsi, la comparaison directe entre les deux essais montre que l'émission vers le bas (S1-E1) engendre des concentrations plus faibles, des variations spatiales limitées et une évolution plus progressive du champ de concentration, alors que l'émission vers le haut (S1-E2) conduit à une dispersion plus rapide et plus complexe, marquée par une accumulation transitoire du CO_2 à hauteur des capteurs.



(a) Concentrations de CO_2 estimées au point de référence 2 de l'essai S1-E1 (Source orientée vers le bas)



(b) Concentrations de CO_2 estimées au point de référence 2 de l'essai S1-E2 (Source orientée vers le haut)

FIGURE 4.28 – Evolution des concentrations de CO_2 de référence et estimées par les méthodes de cartographie spatio-temporelle (Krigeage, Regularisation tensorielle, Pinns) au point de référence 2 pour l'essai S1-E1 (a) et l'essai S1-E2 (b)

Les différences en fonction de l'orientation de la source se manifestent principalement par une augmentation des concentrations moyennes et une modification de la structure des champs de concentration. Néanmoins, malgré ces variations, les performances des modèles de reconstruction spatio-temporelle demeurent globalement stables. Cette stabilité s'explique par le caractère stationnaire du plan d'expérience A réalisé en l'absence de ventilation active pendant la phase de pollution, ce qui limite les phénomènes de convection et favorise un régime de diffusion homogène, propice à la robustesse des trois approches considérées (Voir figure 4.28).

En résumé, l'orientation de la source modifie la dynamique de dispersion du CO_2 et la structure spatiale des cartographies reconstruites. Une orientation vers le bas favorise un régime de diffusion lent et homogène, tandis qu'une orientation vers le haut amplifie les phénomènes de mélange et les contrastes de concentration à hauteur des capteurs. Cette influence doit être prise en compte dans l'interprétation des cartes spatio-temporelles, car elle conditionne à la fois les niveaux de concentration observés, la rapidité des variations et la précision globale des reconstructions.

4.3.5 Influence du débit d'émission

Une autre caractéristique à étudier dans la modélisation spatio-temporelle est le débit d'émission de la source. Ce paramètre détermine la dynamique de dispersion et la formation des champs de concentration en milieu intérieur. Il conditionne à la fois l'intensité, la portée et la structuration du polluant, influençant ainsi la distribution spatio-temporelle observée et la qualité des cartographies reconstruites à partir des mesures du réseau de capteurs. Un débit élevé tend à générer des zones de forte concentration localisée et à accentuer les contrastes spatiaux, tandis qu'un débit plus faible favorise une dispersion plus homogène mais à des niveaux de concentration moindres. Ces effets se traduisent par des différences notables dans la stabilité des champs reconstruits, ainsi que dans la sensibilité des modèles de cartographie aux conditions d'émission.

L'objectif de cette section est d'analyser l'influence du débit d'émission sur la précision et la représentativité des cartes spatio-temporelles de concentration obtenues. Dans le contexte de la modélisation des expositions professionnelles, où la dynamique des sources d'émission conditionne la validité et la capacité de généralisation des modèles développés, des études antérieures ont en effet montré que la fidélité des cartographies dépend fortement de la représentativité des régimes d'émission et de leur adéquation avec les conditions réelles d'aération et de ventilation [Zuidema19b], [Ludwig17].

Démarche expérimentale

Afin d'évaluer l'influence du débit d'émission sur la qualité de la cartographie spatio-temporelle des concentrations, trois essais expérimentaux ont été menés avec des débits de 12,5 L/min, 20 L/min et 45 L/min. Il s'agit des essais S1-E3, S1-E4 et S1-E5 de la série 1 (Voir tableau 3.2) réalisés suivant une même grille d'échantillonnage (Voir figure 4.29) et les plans d'expérience et de calibration A (Voir sections 3.7.1 et 3.4.4).

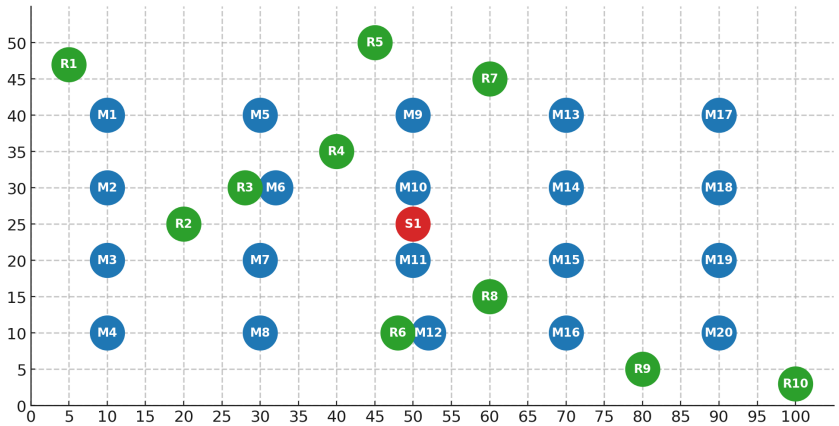


FIGURE 4.29 – Grille d'échantillonnage régulière avec les modules capteurs du RCSFg en bleu (M1-M20), les détecteurs de référence en vert (R1-R10) et la source de pollution en rouge S1

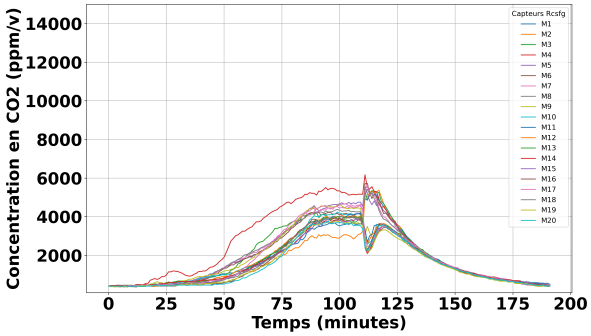
Les durées d'émission pour la phase de pollution ont été ajustées de manière à garantir une

même quantité totale de polluant émise pour l'ensemble des essais, soit respectivement 84 min, 58 min et 30 min pour les débits de 12,5 L/min, 20 L/min et 45 L/min. Il s'en ai suivi la phase de stabilisation d'une durée de 20 minutes et la phase de ventilation qui a duré environ 70 minutes.

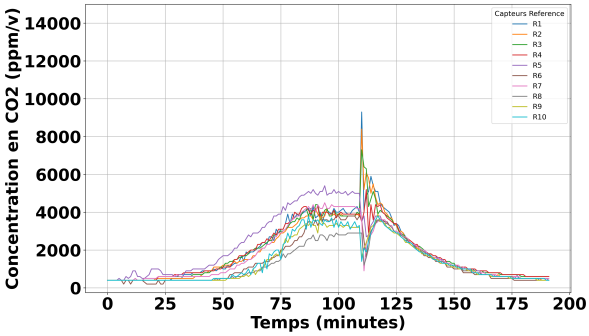
Les essais se sont déroulés en mai 2023 dans des conditions thermo-hygrométriques contrôlées, avec une température moyenne et une humidité relative de l'air respectivement de 22°C et 55% .

L'analyse de l'influence du débit sur la cartographie spatio-temporelle repose sur une approche combinée, associant l'étude des profils temporels de concentration, des structures spatiales reconstruites et l'évaluation quantitative des performances avec les différentes méthodes de cartographie. Pour chaque méthode, les champs reconstruits ont été comparés aux mesures issues des dix points de référence, permettant d'évaluer la fidélité de la reconstitution à l'aide du RMSE et du R^2. Cette analyse conjointe des dynamiques temporelles, des structures spatiales et des indicateurs de performance offre une compréhension approfondie de l'impact du débit d'émission sur la précision et la robustesse des différentes approches de cartographie spatio-temporelle.

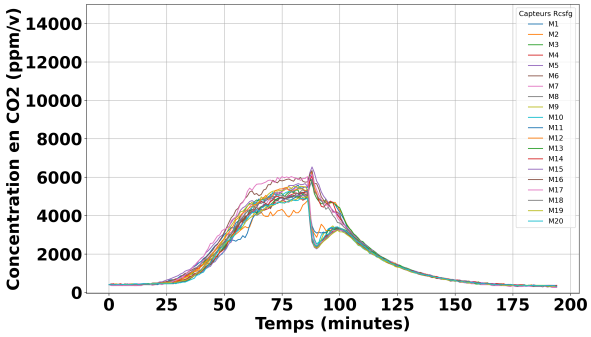
Résultats et discussions



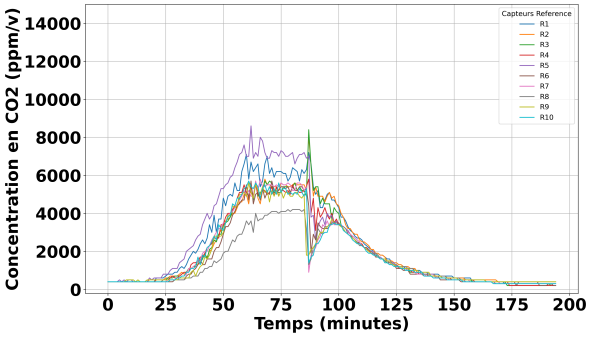
(a) Mesures RCSFg | Essai S1-E3



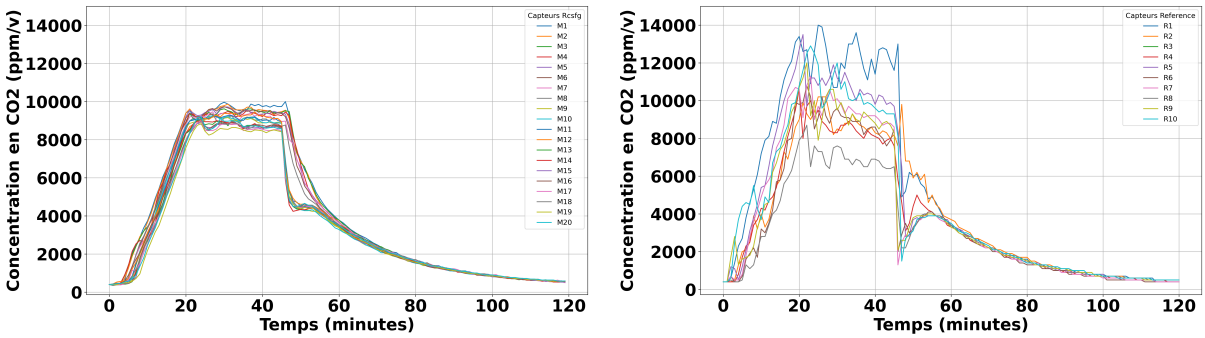
(b) Mesures de référence | Essai S1-E3



(c) Mesures RCSFg | Essai S1-E4



(d) Mesures de référence | Essai S1-E4



(e) Mesures RCSFg | Essai S1-E5 (f) Mesures de référence | Essai S1-E5

FIGURE 4.30 – Evolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) au cours des essais S1-E3, S1-E4 et S1-E5

La figure 4.30 présente l'évolution temporelle des concentrations de CO_2 mesurées par les modules capteurs du RCSFg (M1–M20) et les détecteurs de référence (R1–R10) pour les débits d'émission de 12,5, 20 et 45 L/min. Ces profils illustrent la dynamique de diffusion du polluant dans la chambre expérimentale de la phase de pollution à la phase de ventilation. À 12,5 L/min, la montée en concentration est progressive (atteint un maximum d'environ 6000 ppm/v) et spatialement homogène. À 20 L/min, la structuration du champ s'intensifie, atteignant 10 000 ppm/v avec des gradients marqués dûs à la saturation des modules capteurs. À 45 L/min, les pics dépassent 14 000 ppm/v et révèlent une forte hétérogénéité locale. Comme attendu, l'augmentation du débit d'émission accroît à la fois le niveau maximal de concentration et l'hétérogénéité spatio-temporelle du champ, conformément aux hypothèses : un débit élevé tend à engendrer des zones localisées de concentration élevée et des gradients marqués, tandis qu'un débit plus faible favorise une dispersion plus homogène.

TABLEAU 4.12 – Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E3

Référence	Krigeage		Regularization		PINNs	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	491.20	0.902	1674.25	-0.144	541.67	0.880
R2	428.33	0.920	476.71	0.901	368.48	0.941
R3	355.04	0.942	386.03	0.931	298.64	0.959
R4	272.22	0.960	381.41	0.922	331.78	0.941
R5	498.91	0.905	2417.31	-1.226	388.82	0.942
R6	439.77	0.869	434.25	0.872	316.35	0.932
R7	340.11	0.939	1009.92	0.463	291.07	0.955
R8	634.03	0.590	524.82	0.719	516.30	0.728
R9	438.92	0.854	631.43	0.697	342.72	0.911
R10	497.42	0.820	1682.12	-1.064	389.38	0.889

TABLEAU 4.13 – Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E4

Référence	Krigeage		Regularization		PINNs	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	730.49	0.892	2181.70	0.041	605.62	0.926
R2	304.50	0.976	296.68	0.977	375.67	0.963
R3	363.90	0.965	393.34	0.960	278.28	0.980
R4	340.81	0.965	385.63	0.956	336.82	0.966
R5	1181.37	0.777	3235.81	-0.669	1140.33	0.793
R6	313.61	0.968	314.41	0.968	262.27	0.978
R7	409.43	0.949	1317.89	0.473	743.28	0.832
R8	572.47	0.808	452.87	0.880	536.36	0.831
R9	446.08	0.930	522.66	0.904	516.77	0.906
R10	499.58	0.921	2327.32	-0.725	622.91	0.876

TABLEAU 4.14 – Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E5

Référence	Krigeage		Regularization		PINNs	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	1974.18	0.827	4857.75	-0.050	1829.62	0.851
R2	617.40	0.965	675.57	0.958	576.93	0.969
R3	803.60	0.935	869.70	0.924	745.08	0.944
R4	1492.97	0.869	5494.52	-0.775	1452.51	0.876
R5	768.31	0.946	766.49	0.947	794.40	0.943
R6	1168.55	0.896	2777.41	0.412	1051.20	0.916
R7	1097.77	0.806	916.14	0.865	1195.49	0.770
R8	833.22	0.942	973.81	0.921	1314.28	0.857
R9	1376.16	0.874	5327.08	-0.893	1863.66	0.768

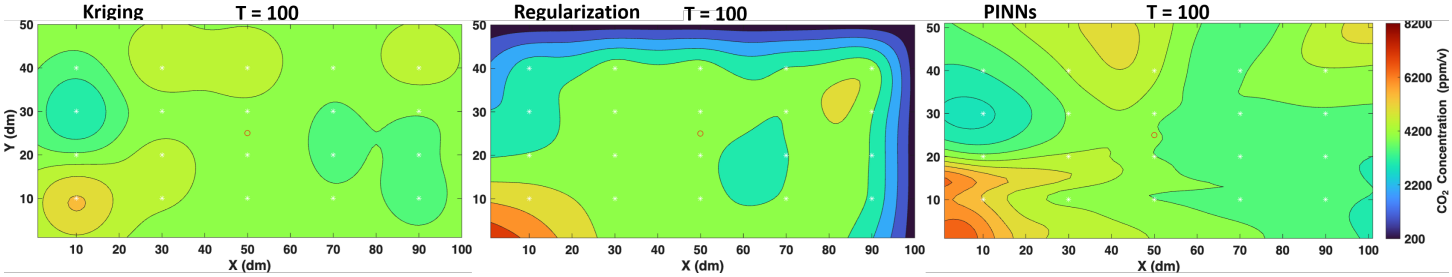
Les tableaux 4.12- 4.13 présentent les valeurs de RMSE et de coefficient de détermination (R^2) pour les dix points de référence, selon les trois méthodes de reconstruction (krigeage, régularisation tensorielle, PINNs).

— Pour le débit de 12,5 L/min, les performances sont globalement élevées : le krigeage affiche un R^2 moyen d'environ 0,90 avec un RMSE d'environ 430 ppm, tandis que les PINNs atteignent des R^2 jusqu'à 0,955 (R7). Ce résultat traduit la bonne représentativité des champs dans ce régime de diffusion relativement homogène.

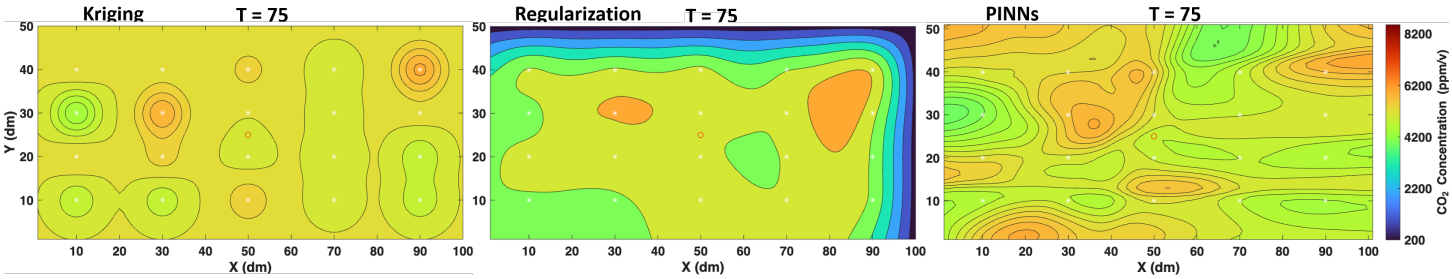
— À 20 L/min, la complexité spatio-temporelle du champ de concentration s'accroît : le krigeage conserve un R^2 moyen proche de 0,90, bien que le RMSE présente des fluctuations plus marquées. En revanche, la régularisation tensorielle montre une dégradation localisée de ses performances (RMSE $>$ 2000 ppm et R^2 parfois négatif), principalement

aux points situés en bordure du domaine, où les effets de bord liés à la formulation matricielle du problème atténuent la contrainte de régularisation. Les PINNs, conservent des performances stables avec des R^2 supérieurs à 0,92.

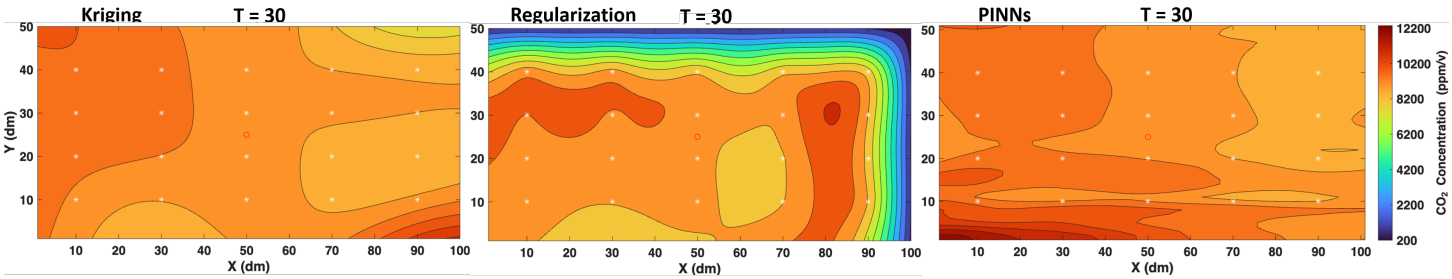
— Au débit de 45 L/min, les reconstructions se dégradent : le krigeage présente des R^2 parfois $< 0{,}85$ et un RMSE moyen > 1000 ppm, les PINNs restant légèrement supérieures avec un RMSE de 900 ppm et un R^2 moyen de 0,88. Cette détérioration traduit l'effet de gradients forts, de variations rapides et d'hétérogénéité accrue du champ.



(a) $T = 100$ min | Essai S1-E3 | Phase de stabilisation pour le débit de 12,5 L/min



(b) $T = 75$ min | Essai S1-E4 | Phase de stabilisation pour le débit de 20 L/min



(c) $T = 30$ min | Essai S1-E4 | Phase de stabilisation pour le débit de 45 L/min

FIGURE 4.31 – Cartographies spatio-temporelles des concentrations de CO_2 à la phase de stabilisation pour les essais S1-E3 (a), S1-E4 (b), S1-E5 (c)

La figure 4.31 présente les cartographies spatiales des concentrations au cours de la phase de stabilisation pour chaque débit :

— À 12,5 L/min : les cartes montrent un champ homogène, zone de concentration élevée proche de la source et gradients doux vers la périphérie. Les reconstructions par PINNs sont particulièrement lissées.

— À 20 L/min : la structuration spatiale devient plus prononcée. Le krigeage et la régularisation tensorielle restitue les zones principales mais atténue les contrastes. Les PINNs

capturent mieux les gradients marqués, reflétant leur capacité à modéliser la dynamique temporelle et l'advection.

— À 45 L/min, la complexité spatio-temporelle du panache s'accroît fortement : les gradients deviennent plus abrupts, les zones de forte concentration plus localisées et les transitions plus rapides. Ces conditions accentuent les écarts de reconstruction et mettent en évidence les limites des modèles dans la restitution fidèle des variations intenses et des structures fines du champ de concentration.

L'ensemble des résultats montre que le débit d'émission constitue un levier prépondérant de la morphologie et de la stabilité des champs de concentration :

— À faible débit, la diffusion domine, le champ est homogène et les méthodes de cartographie obtiennent de bonnes performances.

— À débit intermédiaire, l'advection et la convection deviennent significatives, générant des gradients spatiaux plus forts et exigeant une modélisation plus sophistiquée.

— À débit élevé, l'intensification des gradients spatio-temporels et la variabilité rapide des concentrations complexifient la reconstruction du champ, diminuant la précision des estimations dans les zones à forte dynamique.

Sur le plan méthodologique, les résultats suggèrent que les modèles hybrides intégrant des contraintes physiques (comme dans les PINNs) ont un avantage certain dans les régimes de dispersion complexes, tandis que les méthodes purement statistiques sont adéquates dans les régimes diffusifs quasi stationnaires. Cela corrobore l'importance d'utiliser dans le domaine professionnel des stratégies de surveillance et d'évaluation des expositions capables de tenir compte de la variabilité spatio-temporelle, d'après la littérature sur la qualité de l'air intérieur dans les environnements de travail [Felgueiras23].

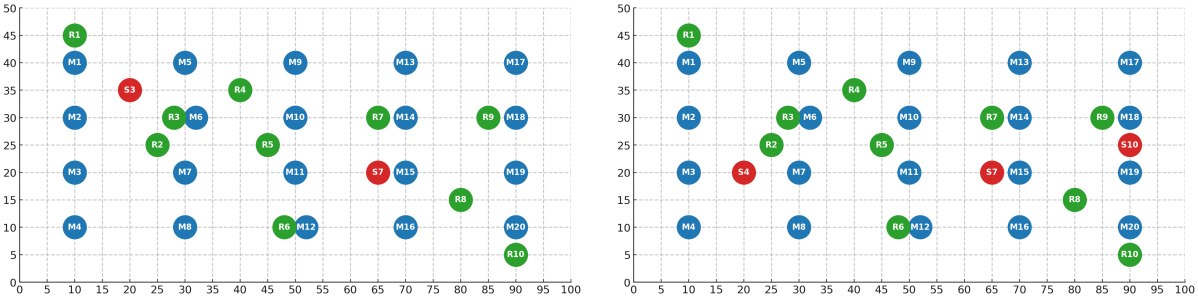
4.3.6 Influence du nombre de sources

L'évaluation de l'influence du nombre de sources d'émission constitue une étape essentielle pour apprécier la robustesse et la capacité de généralisation du modèle de cartographie spatio-temporelle des concentrations de polluants. En effet, la complexité des champs de concentration augmente avec le nombre de sources, notamment en raison des phénomènes d'interaction et de superposition. L'objectif de cette section est d'analyser la capacité des différentes approches de modélisation spatio-temporelle à reconstruire fidèlement les champs de concentration dans des configurations comportant un nombre variable de sources. Deux scénarios expérimentaux ont été considérés : le premier avec deux sources d'émission et le second avec trois sources, tout en maintenant identique la grille d'échantillonnage. Dans un souci de comparabilité, le débit d'émission de chaque source et la durée d'injection ont été ajustés de manière à garantir une quantité totale de polluant équivalente dans la chambre expérimentale, indépendamment du nombre de sources.

Démarche expérimentale

Les essais ont été réalisés dans la chambre expérimentale suivant les grilles d'échantillonnage présentées aux figures 4.32a et 4.32b, avec vingt modules capteurs (M1–M20) positionnés sur une grille régulière, dix détecteurs de référence (R1–R10) et les sources de pollution représentées en rouge.

Dans la configuration à deux sources, chaque source émet un débit de 30 L/min de CO_2, tandis que dans la configuration à trois sources, le débit est à 20 L/min afin de conserver une même quantité totale émise.



(a) Grille régulière avec 2 sources de pollution (b) Grille régulière avec 3 sources de pollution

FIGURE 4.32 – Grilles d'échantillonnages régulières avec 2 sources de pollution (a) et 3 sources de pollution (b)

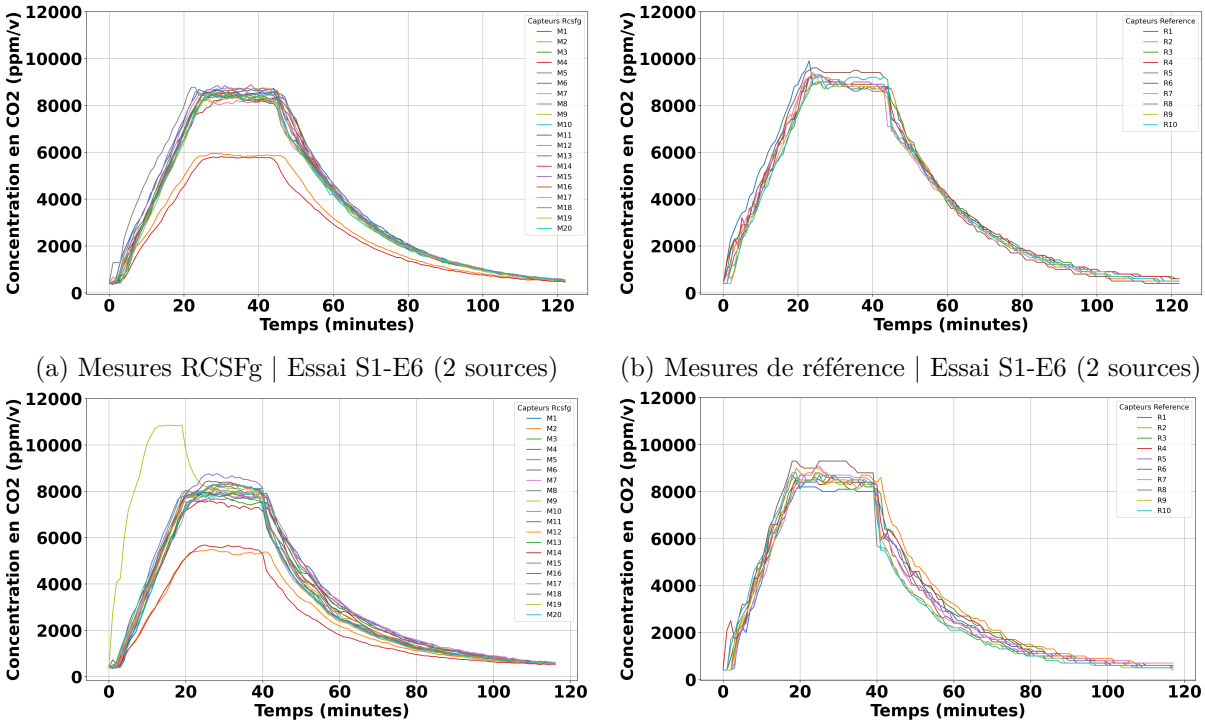
Il s'agit des essais S1-E6 et S1-E7 de la série 1 (Voir tableau 3.2 réalisés suivant les plans d'expérience et de calibration A (Voir sections 3.7.1 et 3.4.4). Les essais se sont déroulés au mois d'août 2023, dans des conditions thermo-hygrométriques contrôlées, la température moyenne de l'air étant de 22 °C et l'humidité relative à environ 50 %. Les durées d'émission pour la phase

de pollution ont été ajustées de manière à garantir une même quantité totale de polluant émise pour les deux essais, soit 25 minutes pour l'essai S1-E6 (2 sources) et 17 minutes pour l'essai S1-E7 (3 sources). Il s'en est suivi la phase de stabilisation d'une durée de 20 minutes et la phase d'épuration qui a duré environ 60 minutes.

Afin d'évaluer l'influence du nombre de sources d'émission sur la qualité de la cartographie spatio-temporelle, l'analyse a porté conjointement sur les profils temporels de concentration enregistrés par les modules capteurs et les détecteurs de référence, ainsi que sur la structure spatiale des champs de concentration reconstruits à l'aide des différentes approches de modélisation. Pour chaque méthode, les concentrations estimées ont été comparées aux mesures de référence, et la performance de la reconstitution a été quantifiée avec le RMSE et le R^2.

Résultats et discussions

La figure 4.33 présente l'évolution temporelle des concentrations en CO_2 mesurées par les modules du RCSF (M1–M20) et les détecteurs de référence (R1–R10) lors des essais S1–E6 (deux sources) et S1–E7 (trois sources).



(a) Mesures RCSFg | Essai S1-E6 (2 sources)

(b) Mesures de référence | Essai S1-E6 (2 sources)

(c) Mesures RCSFg | Essai S1-E7 (3 sources)

(d) Mesures de référence | Essai S1-E7 (3 sources)

FIGURE 4.33 – Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) et les détecteurs de référence (R1-R10) au cours des essais S1-E6 (2 sources), S1-E7 (3 sources)

Dans les deux configurations, les concentrations mesurées mettent en évidence les différentes phases attendues du plan d'expérience A. Une montée (Phase de pollution), Une stabilisation et une décroissance progressive (Phase d'épuration). Toutefois, l'essai à trois sources se caractérise

par des pics de concentration plus élevés (jusqu'à saturation du module à 12 000 ppm/v,) et une dispersion plus importante des mesures, ce qui reflète une distribution spatiale du CO_2 plus hétérogène. L'essai à deux sources présente des profils temporels plus homogènes et une meilleure cohérence inter-capteurs (courbes similaires), traduisant une répartition plus uniforme du mélange gazeux dans l'espace expérimental.

Les performances quantitatives des méthodes de cartographies spatio-temporelles (Voir tableaux 4.15 et 4.16) confirment cette tendance. Pour les essais à deux sources (S1-E6), les trois approches présentent des R^2 élevés (généralement supérieurs à 0,94) et des valeurs de RMSE comprises entre 500 et 800 ppm pour la majorité des points de référence. La méthode de régularisation conduit localement à de meilleures performances (R^2 = 0,97), notamment dans les zones où les gradients sont faibles, traduisant sa robustesse en conditions quasi homogènes (Voir tableau 4.15) .

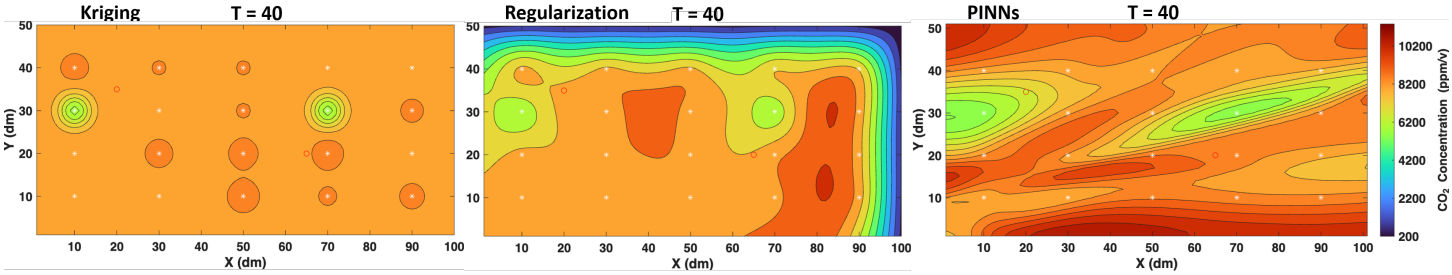
Tableau 4.15 – Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E6 (2 sources)

Référence	Krigeage		Regularization		PINNs	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	732.55	0.947	2128.68	0.550	747.34	0.945
R2	621.82	0.963	614.55	0.964	777.51	0.942
R3	569.73	0.966	528.81	0.971	664.15	0.954
R4	566.62	0.966	496.73	0.974	543.47	0.969
R5	607.52	0.963	569.75	0.967	966.17	0.905
R6	776.62	0.947	769.92	0.948	728.28	0.954
R7	823.38	0.931	1253.96	0.839	1736.19	0.691
R8	608.02	0.960	912.62	0.910	476.75	0.975
R9	647.72	0.955	929.01	0.908	502.03	0.973
R10	696.25	0.951	768.45	0.940	514.07	0.973

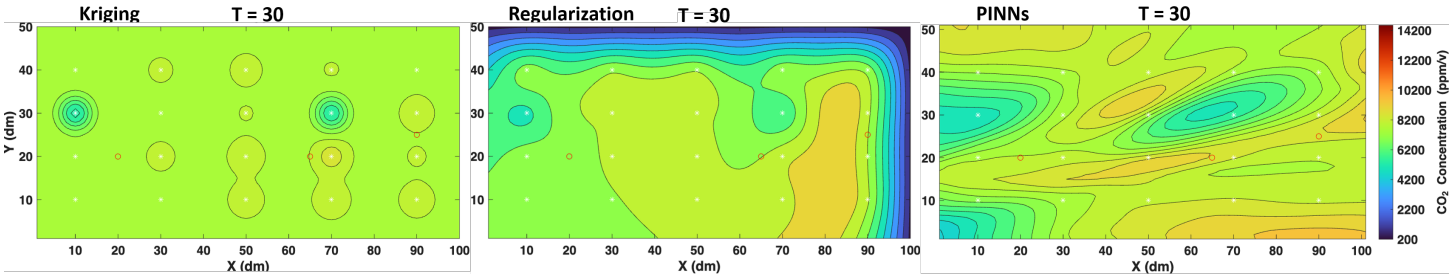
Pour l'essai S1-E7 (3 sources), une légère baisse des performances est observée, en particulier pour les méthodes fondées sur des hypothèses de stationnarité spatiale comme le krigeage. Ceci peut être expliqué par le fait que la présence de plusieurs sources d'émission entraîne des gradients plus abrupts et des anisotropies locales du champ de concentration ; dans ces circonstances, les modèles purement géostatistiques peinent à estimer fidèlement les concentrations (Voir tableau 4.15). La méthode PINNs conserve une cohérence globale supérieure (R^2 moyen d'environ 0,96) grâce à sa capacité à intégrer les phénomènes physiques dffusifs, offrant ainsi une meilleure généralisation en conditions complexes.

TABLEAU 4.16 – Comparaison des performances (RMSE et R^2) pour 10 points de référence | Essai S1-E7 (3 sources)

Référence	Krigeage		Regularization		PINNs	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
R1	249.41	0.992	1777.52	0.613	648.08	0.949
R2	672.77	0.952	577.54	0.964	634.93	0.957
R3	428.73	0.979	387.08	0.983	468.31	0.975
R4	0.977	499.67	0.971	449.00	419.59	0.979
R5	595.50	0.960	431.19	0.979	832.77	0.923
R6	730.86	0.949	724.08	0.950	718.27	0.951
R7	949.23	0.898	1059.44	0.874	1775.11	0.645
R8	648.60	0.953	833.60	0.922	591.89	0.961
R9	587.55	0.961	847.36	0.918	560.22	0.964
R10	658.65	0.954	716.83	0.945	594.48	0.962



(a) $T = 40$ min | Essai S1-E6 | Phase de stabilisation pour 2 sources de 30L/min



(b) $T = 30$ min | Essai S1-E7 | Phase de stabilisation pour 3 sources de 20L/min

FIGURE 4.34 – Cartographies spatio-temporelles des concentrations de CO_2 à la phase de stabilisation pour les esssais S1-E6 (a), S1-E7 (b)

Sur le plan spatial (Voir figure 4.34), les cartographies obtenues durant la phase de stabilisation mettent clairement en évidence l'impact du nombre de sources d'émission sur la structure des champs de concentration. En effet, pour la configuration à deux sources (essai S1-E6), les concentrations présentent une organisation relativement régulière autour des zones d'émission. Les gradients de concentration s'établissent de manière continue depuis les points sources vers la périphérie du domaine. Lorsque le nombre de sources augmente (essai S1-E7), la morphologie du champ de concentration se complexifie sensiblement. Les panaches issus des différentes zones d'émission se superposent et interagissent, générant des zones de recouvrement où les

concentrations locales atteignent des niveaux supérieurs à la moyenne. Ainsi, l'augmentation du nombre de sources se traduit par une redistribution plus complexe du CO_2 dans la chambre expérimentale, avec des zones de concentration persistantes et des gradients moins prévisibles. Ce comportement met en évidence la sensibilité du champ de concentration aux interactions entre panaches, qui deviennent un facteur déterminant dans la dynamique de mélange et la structure spatiale des polluants en milieu intérieur.

D'un point de vue physique, il est aussi probable que l'augmentation du nombre de sources modifie la structure de turbulence induite dans la chambre expérimentale, ce qui affecte la dispersion et le mélange des polluants. Des études en santé au travail ont montré que la distribution spatiale des contaminants dans des environnements clos dépend fortement du nombre et de la localisation des sources, mais également de la dynamique de l'air ambiant et des obstacles présents [Borgini20], [Cherrie18]. Par exemple, Borgini *et al.*, [Borgini20] ont observé qu'une augmentation du nombre de sources ponctuelles de CO_2 dans des espaces confinés entraînait une variabilité spatiale accrue des concentrations, pouvant atteindre un facteur 2 à 3 entre des zones proches, même en conditions de renouvellement d'air limité.

De même, Cherrie *et al.*, [Cherrie18] ont montré que, dans des environnements industriels, la superposition de plusieurs émissions locales pouvait compromettre la représentativité des mesures ponctuelles et nécessiter des approches de modélisation plus intégrées pour évaluer l'exposition réelle des travailleurs.

En conclusion, le nombre de sources d'émission constitue un paramètre déterminant dans la structure et la dynamique des champs de concentration en milieu confiné. Les méthodes classiques de reconstruction (telles que le krigeage) conservent de bonnes performances pour des configurations simples, mais montrent leurs limites lorsque la complexité spatiale augmente. Les autres méthodes apparaissent plus adaptées pour traiter des scénarios multi-sources, en raison de leur capacité à intégrer de manière cohérente les interactions non linéaires entre diffusion, superposition et continuité spatiale et temporelle. Ce sont des éléments dont il faudra tenir compte pour des applications en pratiques de la cartographie spatio-temporelle en milieux professionnels, par exemple pour la recherche de sources d'émission.

4.3.7 Influence du type de polluant : essais avec un aérosol

Plusieurs types de polluants atmosphériques, gazeux ([Jumpponen13],[Monsé14]) et particulaires ([Schlünssen23]), sont susceptibles d'être rencontrés dans les environnements de travail. Il n'est d'ailleurs pas rare que les deux types de polluants soient présents dans en entreprise ([Davidson18]). Aussi, il peut être nécessaire d'avoir recours à la cartographie spatio-temporelle pour chacun de ces deux types de pollution, voire pour les deux. Les travaux publiés antérieurement rapportent d'autres applications de la cartographies spatio-temporelle aux problématiques de la santé au travail pour les gaz et les particules [Peters12],[Zuidema19b]. Cependant, ces tra-

vaux n'ont pas clairement abordé la problématique de l'adéquation des modèles employés pour la cartographie spatio-temporelle et les caractéristiques des polluants dont les concentrations sont à cartographier. Pourtant, la nature du polluant et les propriétés physico-chimiques qui lui sont propres influencent fortement la dynamique de sa dispersion. Elle peut donc affecter la représentativité des cartographies spatio-temporelles reconstruites. Les propriétés physico-chimiques propres à chaque polluant telles que la densité, la diffusivité, la réactivité chimique, la propension à s'adsorber sur les surfaces ou encore la taille des particules, influencent directement les mécanismes de transport, de dépôt, de mélange et d'accumulation en milieu intérieur. Ces différences peuvent se traduire par des structures spatio-temporelles distinctes. L'objectif de cette section est d'évaluer l'influence du type de polluant sur la qualité et la structure des cartes reconstruites à partir des données de concentration acquises par un réseau de capteurs. Il s'agit notamment de déterminer dans quelle mesure la similarité de leurs comportements à petite échelle permet d'exploiter des approches de modélisation communes pour la cartographie spatio-temporelle de polluants de nature hétérogène. Pour aborder ce sujet, le programme a été complété avec des essais de cartographie en prenant un aérosol comme modèle expérimental. Les résultats et conclusions relatifs à ces essais seront comparés à ceux obtenus avec CO_2, largement étudié dans les sections précédentes.

Démarche expérimentale

La démarche expérimentale est fondée sur l'essai S1-E9 de la série 1 (voir tableau 3.2), réalisé sur une grille d'échantillonnage régulière conformément au plan d'expérience B. Cet essai a été réalisé avec le CO_2 et a servi de base pour la mise en place d'un protocole expérimental équivalent dédié aux aérosols. Afin d'assurer une comparaison optimale entre les deux types de polluants, un essai analogue S2-E1 a été réalisé (voir tableau 3.3), cette fois-ci avec des particules fines. La génération des particules a été réalisée à l'aide d'un générateur d'aérosols salins Palas AGK 2000, selon la procédure décrite dans la section 3.5.2. Le réseau de capteurs particulaires RFCp (voir section 3.6.2) a été déployé selon la même grille d'échantillonnage que celle utilisée pour le CO_2, comprenant 20 capteurs positionnés de manière régulière (Voir figure 4.16a). Contrairement à l'essai mené avec le CO_2, aucun détecteur de référence n'a pu être intégré au plan de mesure pour l'aérosol.

Les essais reposent sur l'utilisation d'une grille simplifiée avec les 20 modules capteurs du RFCp en bleu (*) et le point de pollution en rouge (°) (Voir figure 4.36).

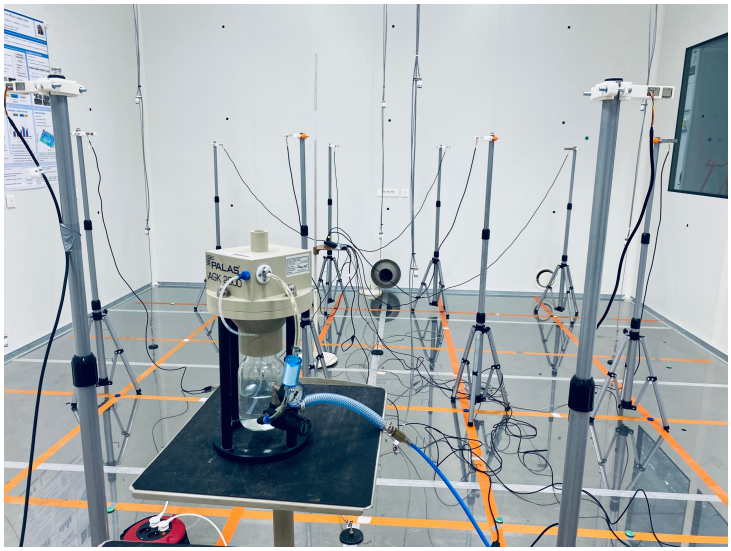


FIGURE 4.35 – Installation du dispositif expérimental pour la réalisation de l'essai avec les microaérosols S2-E1

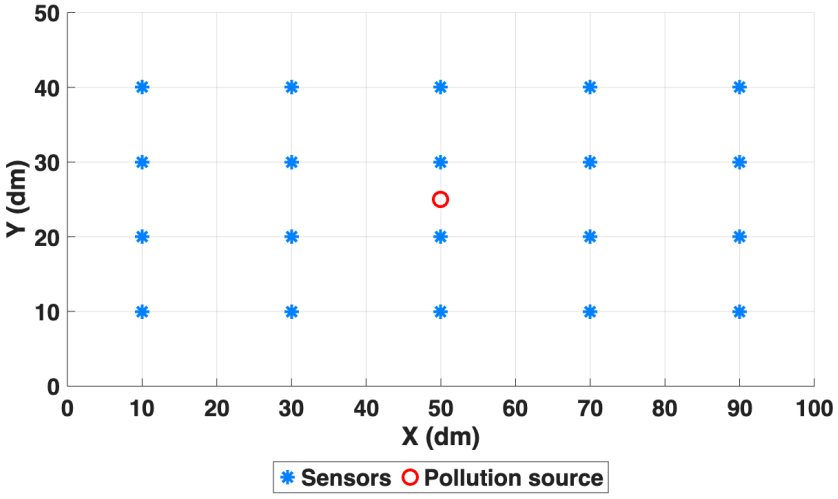


FIGURE 4.36 – Grille de mesures de l'essai S2-E1 avec les microaérosols

L'essai S2-E1 a été conduit suivant le plan d'expérience B, en utilisant une ventilation ajoutée via 2 petits ventilateurs et un épurateur disposés à l'intérieur de la chambre expérimentale. Cette configuration vise à reproduire des conditions de brassage et d'épuration représentatives d'un environnement réel, tout en tenant compte des phénomènes de dépôts des aérosols. Les petits ventilateurs visent à assurer le mélange homogène durant la phase de pollution, tandis que l'épurateur facilite la réduction progressive de la concentration lors de la phase d'épuration. Comme l'essai avec le CO_2 S1-E9, l'essai avec les aérosols a été réalisé en février 2024 sous les mêmes conditions thermo-hygrométriques, relativement stables, avec une température moyenne de l'air de 21 °C et une humidité relative de 46 % .

L'essai s'est déroulé sur une durée totale de 60 minutes, comprenant une phase de pollution (30 min) et une phase d'épuration (30 min). Les mesures ont été enregistrées à une fréquence d'acquisition initiale de 1 seconde. Cependant, pour assurer la cohérence des comparaisons avec l'essai CO_2, les données ont été exploitées à une fréquence d'acquisition de 1min. Les mesures enregistrées par le réseau de capteurs RFCp ont été synchronisées et normalisées suivant les coefficients d'étalonnage (Voir tableau 3.5). Les données acquises ont été traitées et classifiées selon des indicateurs représentatifs des particules émises, basés sur leurs concentrations massique et numérique. Les variables MassConc_{10p} et $\text{MassConc}_{2.5p}$ désignent respectivement les concentrations massiques des particules de diamètre aérodynamique inférieur à 10 µm (PM_{10}) et 2,5 µm ($\text{PM}_{2.5}$), exprimées en microgrammes par mètre cube ($\mu g/m^3$). Elles traduisent la charge massique totale en particules contenue dans un volume d'air donné et constituent des indicateurs de référence de la qualité de l'air. Les paramètres NumbConc_{10p} et $\text{NumbConc}_{2.5p}$ correspondent aux concentrations numériques, exprimant le nombre de particules par unité de volume (particules/m³) pour les mêmes classes granulométriques. Ces mesures renseignent sur la distribution des particules émises.

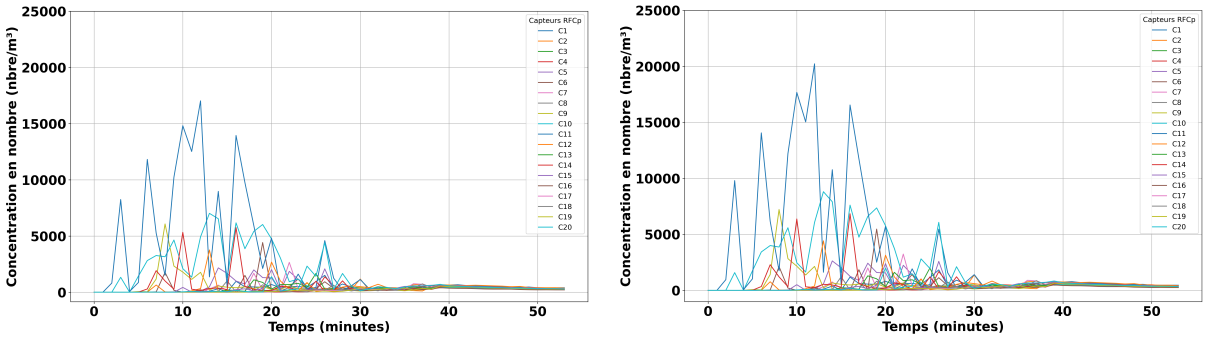
Les cartographies spatio-temporelles des concentrations ont ensuite été établies à partir de ces données avec les trois modèles considérés. L'étude de l'influence du type de polluant sur la cartographie spatio-temporelle a été effectuée en observant la structure des cartes à des instants donnés pendant l'essai. Le recours à des indicateurs de fiabilité comme le RMSE n'a pas été possible pour l'aérosol en l'absence de détecteur de référence pour ce polluant.

Résultats et discussions

L'analyse conjointe des concentrations massiques et numériques, associée aux seuils de 2,5 et 10 µm, a permis une caractérisation complète de la charge particulaire et de la dynamique de dispersion des aérosols dans l'environnement expérimental. En premier lieu, les résultats obtenus lors des essais indiquent que l'aérosol généré dans l'espace expérimental est infra-micrométrique. Pour les besoins des essais, le choix s'est porté sur un aérosol modèle à dominante micrométrique généré par un nébuliseur assez bien caractérisé (AGK 2000). La distribution granulométrique

des aérosols produits par le générateur AGK 2000 peut être ajusté de manière reproductible dans une plage comprise entre environ 5 nm et 15 µm, en fonction de la concentration de la solution saline utilisée. Les résultats corroborent ceux déjà publiés avec le même générateur, le même soluté (NaCl), la même concentration (1%) et avec des conditions de génération similaires ([Bémer24]).

Par ailleurs, il apparaît que l'évolution temporelle des concentrations en PM_{10} et en $PM_{2.5}$ dans les aérosols est complètement différente de celle observée avec le CO_2. En effet, elle révèle plusieurs pics de concentration transitoires et successifs au cours des 30 premières minutes correspondant à la phase de pollution (Voir Figure 4.38). Ensuite, une décroissance rapide vers un régime quasi stationnaire est observée, après le démarrage de la phase d'épuration (les 30 minutes suivantes). Il n'y a donc pas d'accumulation de particules dans l'air pendant la phase de pollution comme cela était observé pour le CO_2 (Voir Figure 4.6). Par ailleurs, les concentrations maximales atteignent environ 3000 µg/m^3 et 20×10^3 nbre/m^3 pour les PM_{10}, contre 2000 µg/m^3 et 15×10^3 nbre/m^3 pour les $PM_{2.5}$. Ces valeurs révèlent une charge particulaire initiale élevée, dominée par les plus grosses particules, tandis que les particules les plus fines présentent une décroissance plus lente, ce qui indique une meilleure stabilité en suspension. Il faut noter que les courbes de concentration en PM2,5 et en PM10 présentent des allures très similaires (*i.e.* les pics de concentration ont la même forme).



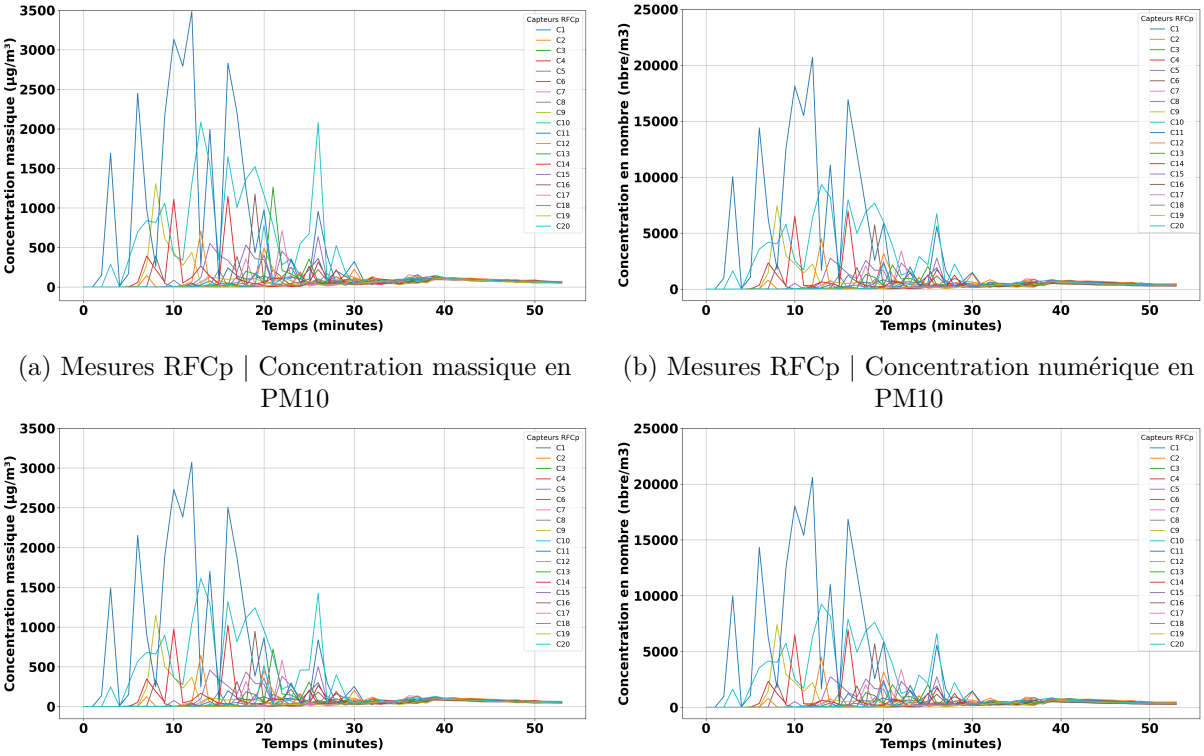
(a) Mesures RFCp | Concentration numérique en PM0.5

(b) Mesures RFCp | Concentration numérique en PM1

FIGURE 4.37 – Évolution des concentrations numérique en particules de taille 0.5 µm (PM 0.5) et 1 µm (PM 1) enregistrées par les modules capteurs du RFCp (C1 à C20) au cours de l'essai S2-E1

Les figures 4.40 et 4.41 présentent les cartographies spatio-temporelles reconstruites des concentrations massiques et numériques de nanoaérosols aux instants représentatifs des phases de pollution ($T = 21$ min) pour les PM_{10} et $PM_{2.5}$.

L'examen des cartes spatio-temporelles construites à partir des données de mesures du RFCp (concentration massique et en nombre) révèle une forte similarité dans la structure globale des champs de concentration lorsque le krigeage et la régularisation sont utilisés pour la reconstruc-



(a) Mesures RFCp | Concentration massique en PM10

(b) Mesures RFCp | Concentration numérique en PM10

(c) Mesures RFCp | Concentration massique en PM2,5

(d) Mesures RFCp | Concentration numérique en PM2,5

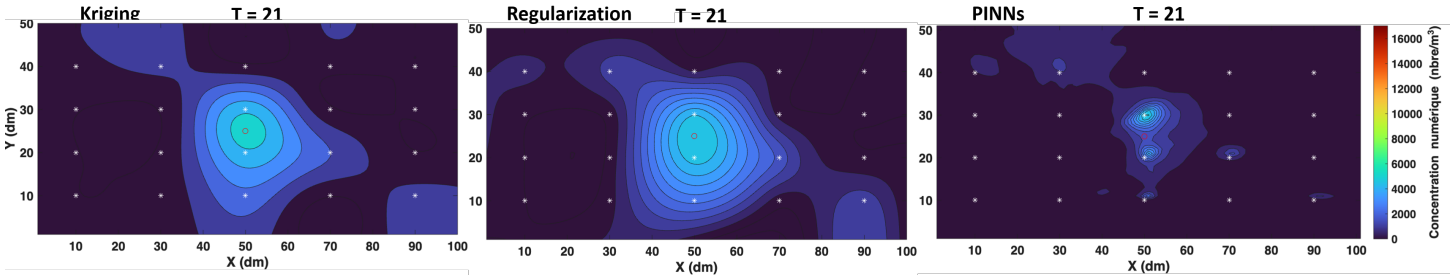
FIGURE 4.38 – Évolution des concentrations numérique et massique en particules de taille 10 µm (PM10) et 2,5 µm (PM 2,5) enregistrées par les modules capteurs du RFCp (C1 à C20) au cours de l'essai S2-E1

tion ; l'emploi du modèle PINNs conduit à des cartes assez différentes comparées aux deux autres modèles (Voir figures 4.39-4.41). Sur toutes les cartes considérées, les zones de forte concentration apparaissent localisées à proximité immédiate de la source d'émission et les concentrations décroissent nettement à mesure que l'on s'éloigne cette dernière. Pour les PM10 (Voir figures 4.40a, 4.41a), les cartes montrent des concentrations maximales très localisées et très marquées autour de la source d'émission, avec une extension spatiale limitée. Pour les PM2,5 (Voir figures 4.40b, 4.41b), la distribution des concentrations est très similaire et légèrement plus diffuse et étendue dans l'espace, avec des niveaux de concentration plus homogènes.

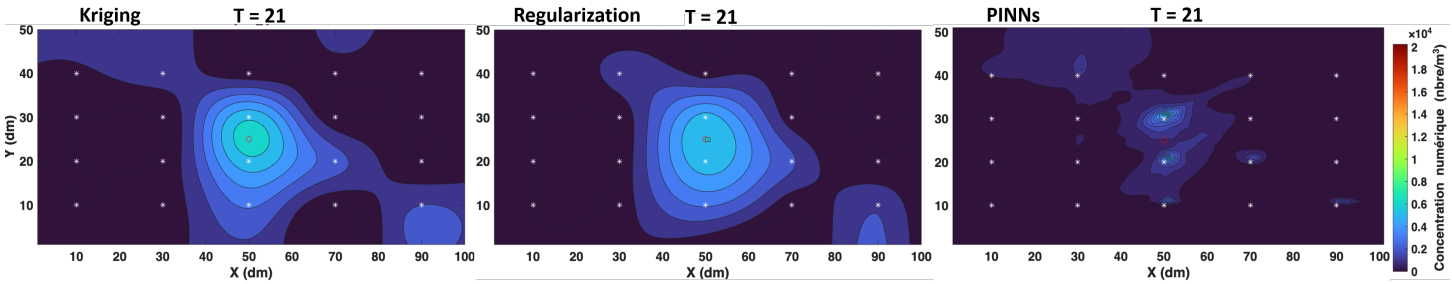
Les propriétés physico-chimiques propres à chaque polluant telles que la densité, la diffusivité, la réactivité chimique, la propension à s'adsorber sur les surfaces ou encore la taille des particules régissent leur dynamique dans l'air. Elles influencent directement les mécanismes de transport, de dépôt, de mélange et d'accumulation en milieu intérieur. Concernant les aérosols, la taille d'une particule d'aérosol est l'une des caractéristiques fondamentales qui détermine ses modalités de transport ([Renoux98]). Pour les besoins de nos essais, nous avons opté pour un aérosol modèle à dominante micrométrique généré par un nébuliseur assez bien caractérisé AGK 2000 de Palas. A l'échelle micrométrique, les mouvements Brownien dominent sur les phénomènes d'inertie entraînant une diffusion isotrope analogue à celle observée pour un polluant gazeux ([Vasilatou24]). Pour ces aérosols, le faible diamètre aérodynamique des particules leur confère une composante diffusionnelle qui est comparable à celles des gaz en termes de dispersion. Toutefois, d'autres mécanismes peuvent intervenir dans le transport des aérosols, notamment lors que la taille des particules est plus importante. Le modèle PINNs utilisé ici étant renseigné par les phénomènes physiques de diffusion uniquement, il est probable qu'il ne soit pas complètement adapté en l'état aux aérosols. L'intégration d'un composant "dépôt par inertie" permettraient sans doute de résoudre ce problème. Ainsi, contrairement au CO_2, les aérosols sont soumis à des phénomènes de dépôt sur les parois par inertie et à des interactions électrostatiques susceptibles de modifier localement leur concentration. Ces différences peuvent expliquer des structures spatio-temporelles distinctes.

La corrélation temporelle entre les concentrations massiques et numériques en aérosol confirme la cohérence du comportement aérodynamique des particules. L'évolution temporelle des mesures met en évidence l'effet prépondérant de la granulométrie sur la dynamique de dispersion : les particules les plus grosses se déposent rapidement par sédimentation, tandis que les plus fines demeurent plus longtemps en suspension dans l'air.

La comparaison de ces résultats avec les cartographies obtenues pour le CO_2 (essai S1-E9, voir figure 4.9) met en évidence des différences significatives dans la structure des champs de concentration reconstruits. Le CO_2, en tant que gaz, présente des distributions avec une certaine continuité temporelle, conduisant à des structures spatiales des concentrations plus aux proximité de la source. À l'inverse, les aérosols montrent des gradients de concentration plus marqués

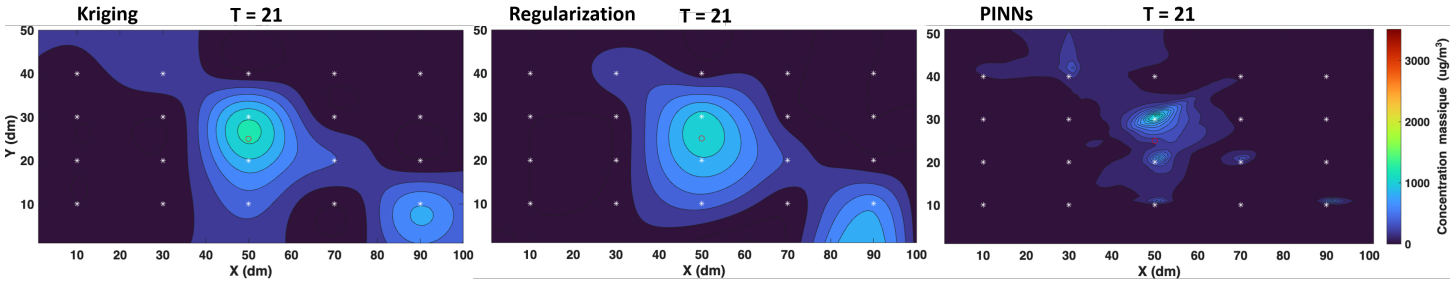


(a) $T = 21$ min | Essai S2-E1 | Concentration numérique 0.5 µm

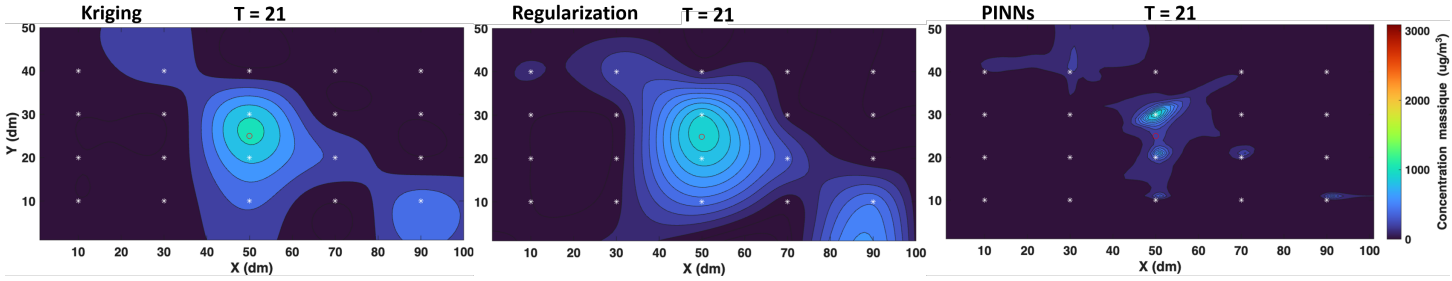


(b) $T = 21$ min | Essai S2-E1 | Concentration numérique 1 µm

FIGURE 4.39 – Cartographies spatio-temporelles des concentrations numérique des aérosols de taille 0.5 µm (a) et 1 µm (b) à $T = 21$ min de l'essai S2-E1

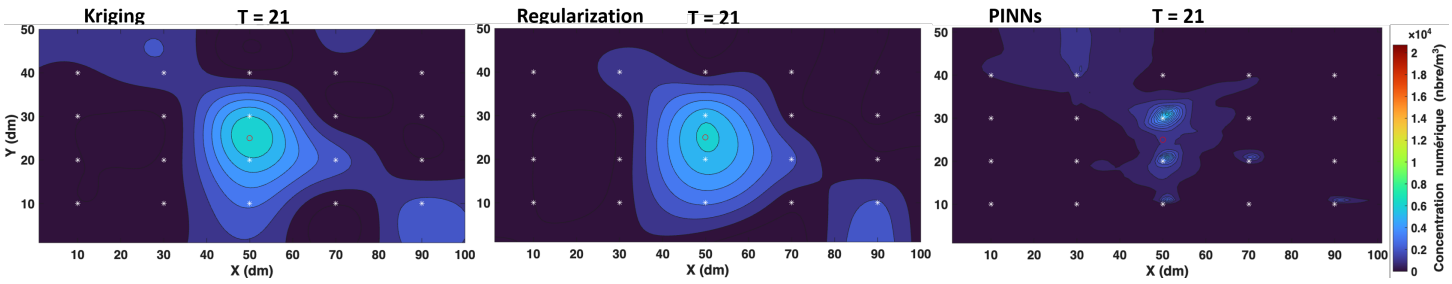


(a) $T = 21$ min | Essai S2-E1 | Concentration massique 10 µm

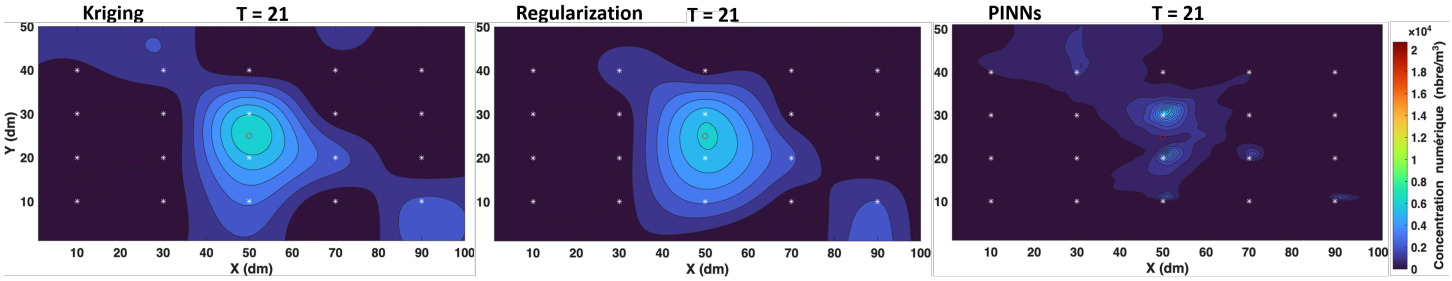


(b) $T = 21$ min | Essai S2-E1 | Concentration massique 2.5 µm

FIGURE 4.40 – Cartographies spatio-temporelles des concentrations massiques de nanoaréosols de taille 10 µm (a) et 2.5 µm (b) à $T = 21$ min de l'essai S2-E1



(a) $T = 21$ min | Essai S2-E1 | Concentration numérique 10 µm



(b) $T = 21$ min | Essai S2-E1 | Concentration numérique 2.5 µm

FIGURE 4.41 – Cartographies spatio-temporelles des concentrations en nombre en particules de taille 10 µm (a) et 2.5 µm (b) à $T = 21$ min de l'essai S2-E1

et une dispersion plus lente, conséquence de leur comportement mixte diffusif–dépositionnel. Ainsi, contrairement au CO_2, les aérosols sont soumis à des phénomènes de dépôt sur les parois par inertie et à des interactions électrostatiques susceptibles de modifier localement leur concentration. Ces différences peuvent expliquer des structures spatio-temporelles distinctes.

Le type de polluant apparaît comme un paramètre déterminant dans la formation et la structuration des champs de concentration en environnement confiné. Les nanoaérosols, bien que présentant à petite échelle un comportement proche de celui d'un gaz, sont soumis à des processus additionnels de dépôt et d'interaction avec les surfaces, qui modifient sensiblement leur répartition spatiale et leur persistance dans le temps.

Les résultats présentés dans cette section indiquent donc des différences importantes entre le polluant gazeux (CO_2) et particulaire (aérosol), à la fois concernant l'allure de l'évolution des concentrations dans le temps 4.37, 4.38 et les cartographies spatio-temporelles de ces concentrations 4.39- 4.41. Ainsi, le type de polluant apparaît comme un paramètre déterminant dans la formation et la structuration des champs de concentration en environnement fermé et confiné. Les évolutions observées impliquent des phénomènes physiques attribuables aux possibilités de diffusion et de dépôts par inertie des polluants considérés ; les débits d'émission sont également différents pour les sources utilisées. Des essais complémentaires avec des détecteurs de référence restent nécessaires concernant les aérosols afin de pourvoir statuer sur la fiabilité des cartes construites (calcul du RMSE). Ces essais pourront permettre de vérifier si les modèles utilisés pour la cartographie spatio-temporelle du CO_2 peuvent être utilisés en l'état ou si leur paramétrage doit être adapté. Il est fort probable que les modèles comme le PINNS doivent intégrer

d'autres renseignements physiques comme le dépôt par inertie et interférences électrostatiques.

4.4 Fusion de données : Représentation 3D de la chambre expérimentale avec les niveaux de concentrations

Cette section présente la méthodologie de fusion entre les données géométriques issues du scanner 3D (Voir Section 3.6.5) et les concentrations de CO_2 reconstruites par cartographie spatio-temporelle. L'objectif est de générer une représentation 3D réaliste de la chambre expérimentale intégrant la distribution du polluant dans l'espace. Cette intégration permet d'associer la géométrie réelle du lieu aux champs de concentration simulés, fournissant ainsi un support d'analyse visuelle et quantitative pour l'évaluation de la qualité de l'air intérieur. Une telle visualisation facilite l'identification des zones de forte concentration, susceptibles de présenter un risque sanitaire accru, et constitue un outil d'aide à la décision pour l'optimisation des dispositifs de ventilation ou la reconfiguration des espaces de travail.

Démarche expérimentale

La procédure de fusion repose sur l'intégration de données hétérogènes issues de diverses sources que sont :

— Des données géométriques 3D obtenues par le scanner 3D (voir section 3.6.5)

— Des données de concentrations en CO_2 issues de la cartographie spatio-temporelle.

Elle s'articule autour des étapes suivantes :

1. **Extraction et structuration des données géométriques :** La modélisation 3D de la chambre expérimentale a été exportée sous la forme d'un nuage de points $\mathcal{P}$ au format « `.ptx` » défini par :

$$\mathcal{P} = \{\, p_i = (x_i, y_i, z_i, R_i, G_i, B_i) \mid i = 1, \ldots, N \,\}, \tag{4.1}$$

   où (x_i, y_i, z_i) représentent les coordonnées spatiales des points et (R_i, G_i, B_i) les composantes colorimétriques associées.

2. **Référencement géométrique et recalage spatial :** Afin d'assurer une cohérence géométrique entre le modèle 3D et la carte de concentration, une analyse en composantes principales a été appliquée à la couche supérieure du nuage de points, correspondant au toit de la chambre expérimentale. Cette analyse permet de déterminer les vecteurs directeurs et le centre du plan de référence, assurant ainsi une orientation et un alignement cohérent entre la géométrie issue du scanner et les repères de la cartographie.

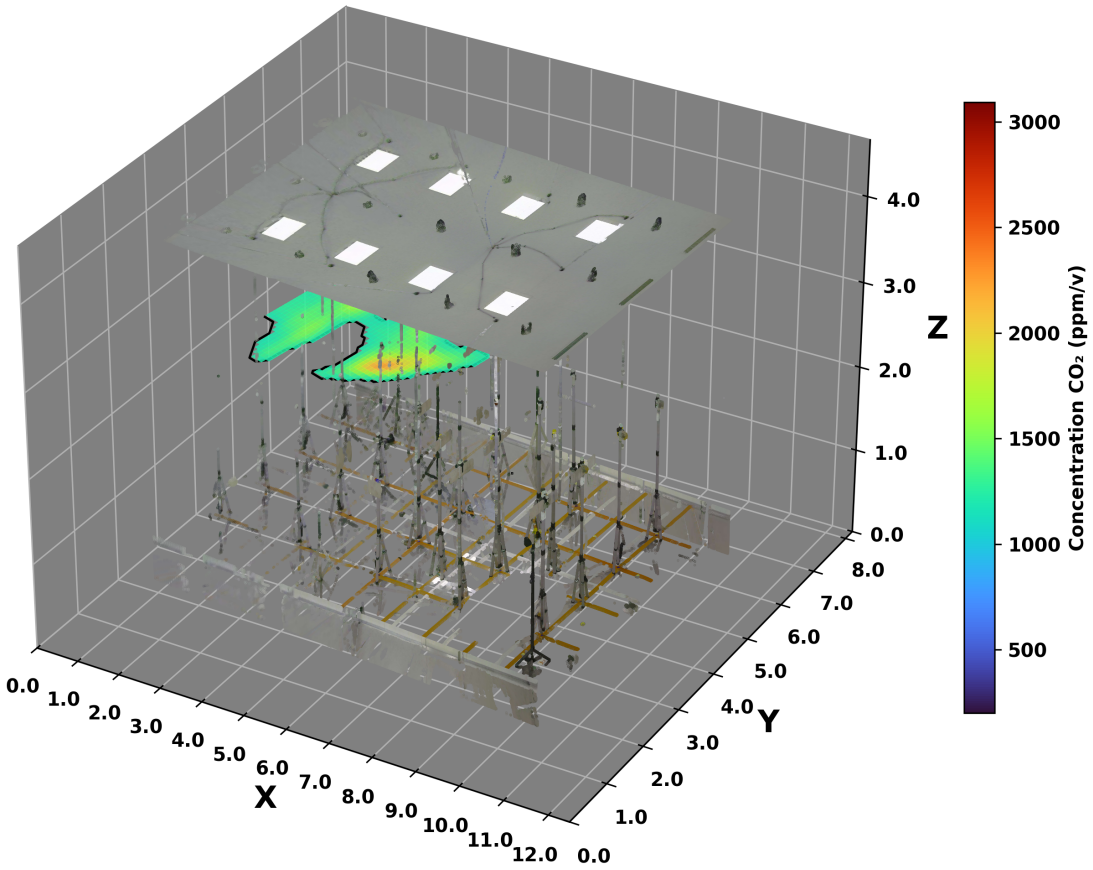
3. **Importation et projection des champs de concentration :** Les champs de concentration estimés à partir de la cartographie spatio-temporelle sont importés sous forme de triplets (x, y, t) associés aux valeurs de concentration correspondantes $C(x, y, t)$. Pour

chaque instant t, la carte 2D des concentrations est projetée dans le repère tridimensionnel défini par le plan du toit, à une hauteur fixe $z = 1,6$ m correspondant au plan de mesure des capteurs.

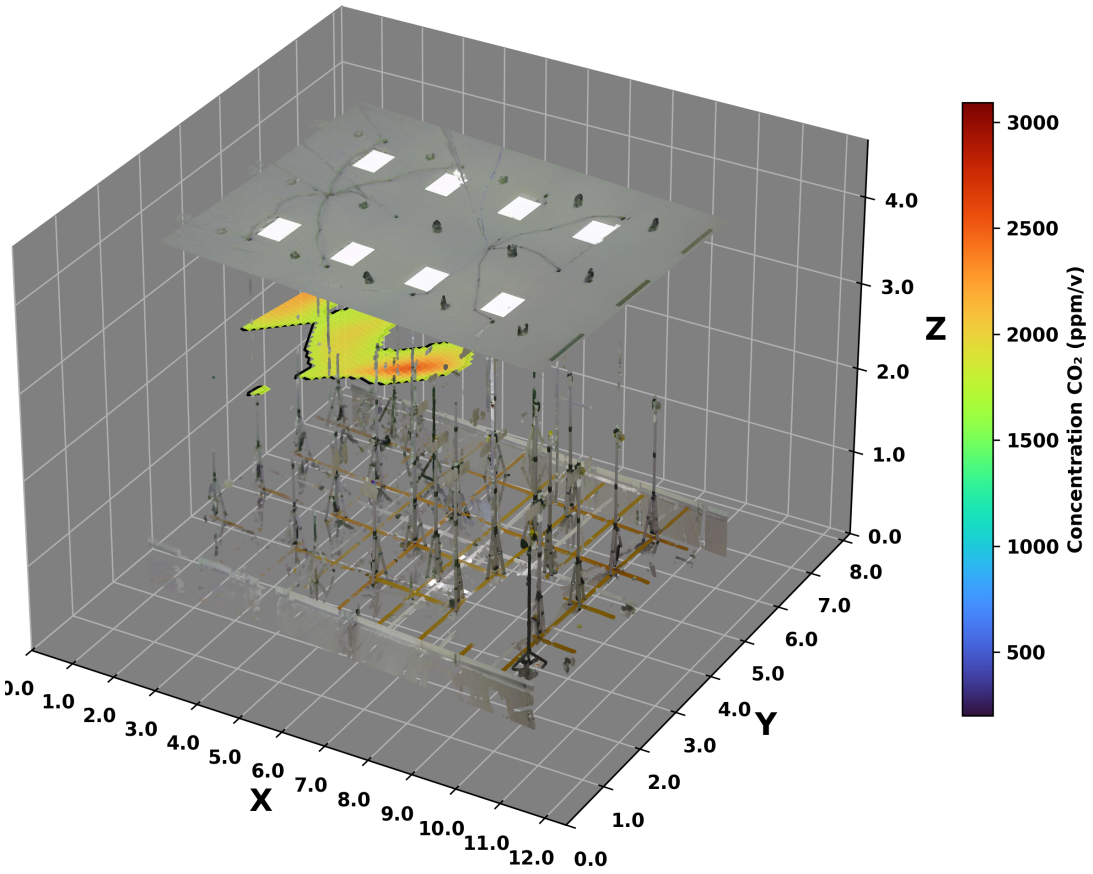
4. **Visualisation :** Afin de faciliter l'interprétation visuelle, un filtre est appliqué aux concentrations pour ne représenter que les zones présentant les concentrations les plus élevées à chaque instant t. Cette étape permet de mettre en évidence les zones à forte concentration, facilitant l'analyse des gradients spatiaux de CO_2 et la détection de zones d'exposition au polluant. Les champs projetés sont ensuite superposés à la modélisation 3D pour produire une visualisation dynamique de la distribution du CO_2 dans la chambre expérimentale.

Résultats et discussions

La Figure 4.42 illustre la représentation 3D résultante de cette procédure de fusion de la chambre expérimentale intégrant la distribution spatiale du CO_2 aux instants T = 45 min (Figure 4.42a) et T = 75 min (Figure 4.42b) de l'essai S1–E2. Les zones colorées correspondent aux zones de forte concentration détectées dans le volume, superposées à la géométrie reconstruite à partir du nuage de points issu du scanner 3D.



(a) T = 45 min



(b) T = 75 min

FIGURE 4.42 – Représentation de l'intérieur de la chambre expérimentale en 3D avec les niveaux de concentrations en CO_2 les plus élevées aux instants $T = 45$ min (a) , 75 min (b) de l'essai S1-E2

Cette visualisation intégrée met en évidence la dynamique spatio-temporelle du CO_2 et permet une analyse fine des concentrations élevées dans la chambre expérimentale. Au-delà de son intérêt méthodologique, cette approche présente une portée applicative directe en santé au travail.

Le CO_2, bien qu'inerte à faibles concentrations, est un indicateur critique du confinement, son accumulation au-delà de 1000 ppm/v signale une ventilation insuffisante susceptible d'altérer les performances cognitives et au-delà de 5000 ppm/v, d'entraîner des effets physiologiques liés à l'hypercapnie.

Des travaux ont montré que des concentrations accrues de CO_2 dans des espaces confinés peuvent altérer la vigilance, réduire la performance cognitive et induire une sensation d'inconfort ou de somnolence, notamment dans les environnements de travail fermés [Pang21].

À plus long terme, le suivi des niveaux de CO_2 dans les lieux de travail est recommandé comme indicateur indirect de la qualité de l'air intérieur et de la ventilation, en complément des mesures d'exposition directe aux polluants [Myatt02].

La fusion des données géométriques issues du scanner 3D avec la cartographie spatio-temporelle des concentrations constitue une approche novatrice dans ce domaine. Elle permet une représentation 3D couplant la géométrie réelle du lieu et la distribution du polluant, rendant ainsi possible l'identification précise des zones critiques d'accumulation. Ces informations sont essentielles pour orienter les stratégies de prévention, qu'il s'agisse d'optimiser la ventilation générale, de reconfigurer les postes de travail ou d'ajuster les dispositifs de ventilation localisée.

En ce sens, la fusion 3D et la cartographie spatio-temporelle constituent une avancée significative pour la visualisation et la compréhension des expositions en milieu professionnel, en matérialisant de façon quantitative et immersive la dynamique d'un polluant invisible et en soutenant la mise en œuvre de stratégies de prévention fondées sur les données.

4.5 Fusion de données : Reconstruction des expositions individuelles

4.5.1 Objectifs et principes des essais

L'objectif des essais de la série 3 est d'étudier la faisabilité de reconstituer l'exposition individuelle d'un opérateur évoluant dans un environnement de travail pollué avec du CO_2, à partir de la cartographie spatio-temporelle des concentrations de CO_2 ambiantes et de la position dynamique de l'opérateur dans cet environnement. L'approche consiste à croiser les données de positionnement et les estimations spatio-temporelles issues d'un réseau de capteurs afin d'obtenir une estimation la plus fiable possible des expositions individuelles. Les essais ont été menés dans la salle aménagée de l'INRS (Voir figure 3.3).

4.5.2 Méthodologie

4.5.2.1 Description de la démarche adoptée

La démarche consiste à comparer les niveaux d'exposition individuelle mesurés sur un vecteur mobile avec les concentrations estimées sur la trajectoire de ce dernier à partir d'une carte spatio-temporelle, elle même issue des mesures avec un réseau de capteurs sans fil.. Elle repose sur des essais expérimentaux comprenant la séquence d'étapes complémentaires suivante :

— **Déploiement du dispositif de mesure** : un réseau de capteurs sans fil est installé à des positions fixes dans l'environnement expérimental pour assurer l'acquisition continue des concentrations en CO_2, servant de données d'entrée pour la modélisation spatio-temporelle. En parallèle, des instruments de référence sont positionnés à des points intermédiaires pour constituer un jeu de données de validation. Les capteurs, sources de pollution et instruments sont répartis suivant la grille d'échantillonnage définie (voir Figure 4.43).

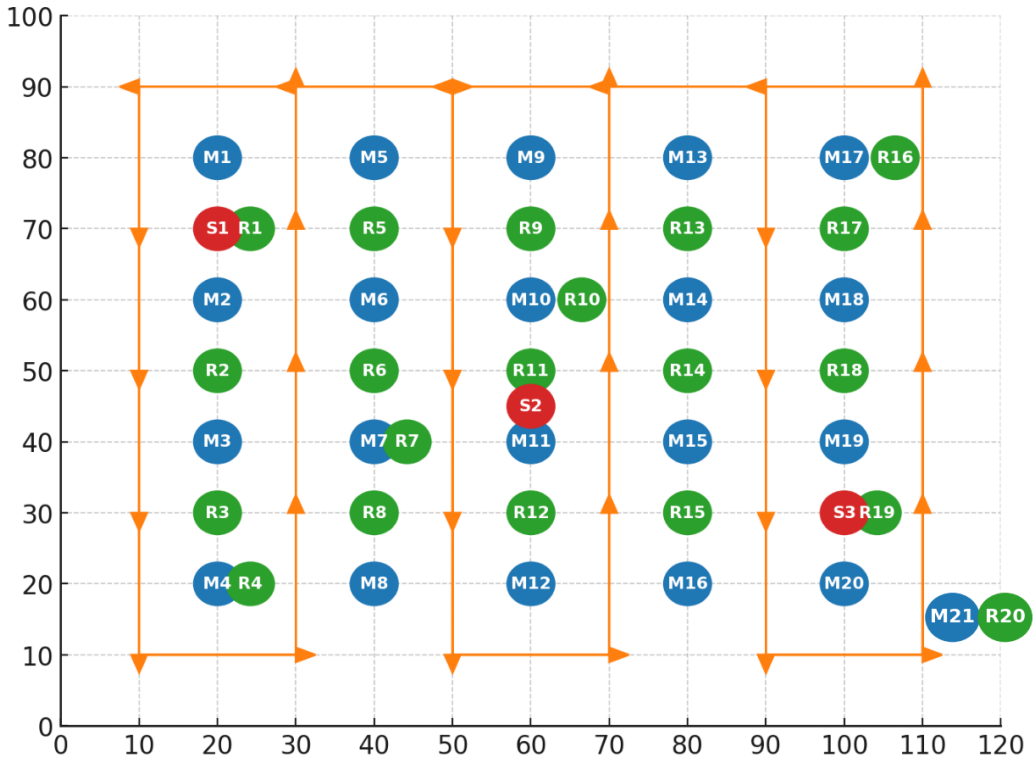


FIGURE 4.43 – Grille d'échantillonnage de l'environnement expérimental. Les points de mesure sont indiqués en bleu (M1 à M20), les points de référence en vert (R1 à R19), les points de sources de pollution en rouge (S1 à S3), en orange, la trajectoire à suivre par le robot équipé avec le capteur M21 et l'instrument de référence R20

— **Simulation de déplacement à l'aide d'un robot mobile** : Le robot SHERPA décrit dans la section 3.6.6 est équipé d'instruments de mesures (module du réseau de capteurs

sans fil et appareil de référence) et programmé pour suivre des trajectoires simulant le déplacement d'un opérateur. Ce dispositif permet de collecter des données de concentration le long du parcours défini. L'ensemble des fonctionnalités du robot, couplé à un système de journalisation (logs), permet de cartographier l'environnement expérimental et de définir une trajectoire prédéterminée, segmentée en points de passage successifs. Durant l'exécution du parcours, l'état du robot ainsi que sa position spatio-temporelle sont enregistrés en continu dans les fichiers de logs. Les données brutes enregistrées sont ensuite traitées par un script Python qui extrait les informations pertinentes et génère un fichier de données au format « .csv ». Ce fichier contient notamment, pour chaque point de passage, la position, l'heure d'arrivée et l'heure de départ. Ce fichier structuré constitue une base de données exploitable pour la suite du traitement, en particulier pour les opérations de fusion de données spatio-temporelles entre les mesures du module capteur Aranet , le détecteur de référence et les trajectoires du robot. Dans un souci de conserver le nombre de modules (20) de la grille spatiale, le robot n'a pas été équipé du module du Rcsfg, nous utiliserons donc le détecteur de référence pour la référence mobile des cartographies réalisées avec les données des modules du Rcsfg.



FIGURE 4.44 – Robot mobile SHERPA équipé des instruments de mesures au sein de l'environnement expérimental

— **Analyse spatio-temporelle et fusion de données** : les mesures de CO_2 à points fixes issues du réseau de capteurs sans fil sont traitées pour produire une cartographie spatio-temporelle des concentrations ambiantes à l'aide les méthodes d'analyse spatio-temporelle présentés dans le Chapitre 2. Cette cartographie est couplée aux trajectoires

du robot afin d'estimer l'exposition individuelle ponctuelle au cours du déplacement.

4.5.2.2 Détail du plan d'expérience

Le plan d'expérience correspond à celui des essais de la série 3 suivant la grille d'échantillonnage décrite dans la section précédente (4.43). Cette série suit le protocole défini par le plan d'expérience B, structuré en trois phases successives : une phase de pollution, une phase de stabilisation et une phase d'épuration. Compte tenu du fait que la salle expérimentale n'est pas hermétique, un léger courant d'air naturel est présent en continu. Ce phénomène est considéré, pour l'ensemble des essais, comme un système de ventilation naturelle de faible intensité, influençant modérément la dispersion du polluant. Pour la réalisation des essais, la phase de pollution a été mise en œuvre selon deux scénarios distincts. Pour le premier scénario, le CO_2 est généré à partir d'un unique point source, désigné S2 (voir Figure 4.43). Ce scénario a été répété 10 fois (Essais S3 E1 à E10 du tableau 4.17). Dans le second scénario, la pollution est produite de manière alternée à partir de trois points sources, S1, S2 et S3, introduisant une variabilité spatiale dans la dynamique de diffusion du gaz. Ce scenario a été répété 5 fois (essais S3 E11 - E15 du tableau 4.17). La phase de pollution a duré 86 min. La phase de stabilisation consiste à arrêter la génération et à laisser le polluant CO_2 s'homogénéiser dans l'espace expérimental. Elle a duré 30 min. La phase d'épuration consiste à ventiler l'espace expérimental pour évacuer le polluant CO_2 et atteindre une concentration de base 400 ppm/v. Elle a été réalisée avec un système de ventilation manuelle (en ouvrant les fenêtres). En fonction des conditions météorologiques, elle dure entre 20 min et 45 min.

Tout au long des essais, le robot mobile a suivi une trajectoire prédéfinie et identique pour l'ensemble des essais (voir figure 4.43). Suivant cette stratégie de mesure, l'ensemble des données collectées a été organisé en trois catégories :

- **Mesures fixes :** Mesures de concentration en CO_2 provenant du réseau de capteurs fixes, utilisés pour générer la cartographie spatio-temporelle.
- **Mesures mobiles :** Mesures de concentration en CO_2 et de position acquises par les capteurs positionnés sur le robot mobile.
- **Mesures de référence :** Mesures de concentration en CO_2 effectuées avec des appareils de référence installés en des points fixes intermédiaires.

Des paramètres environnementaux tels que la température, l'humidité relative et la vitesse de l'air ambiant sont également enregistrés, afin d'identifier d'éventuelles sources de variabilité dans les mesures. Deux types de réseau de capteurs sans fil ont été déployés pour les mesures de CO_2, aux emplacements fixes. Le RCSFg d'une part, et le réseau ARANET d'autre part. Les instruments de mesures ToxiRAE ont été installés aux points de référence. Pour les mesures mobiles, un ToxiRAE a été utilisé pour servir de référence mobile pour la cartographie réalisée avec les données du RCSFg, et un capteur Aranet pour celle générée par les données du réseau

Aranet. L'ensemble des dispositifs de mesure (fixes, mobiles et de référence) a été synchronisé de manière à effectuer des acquisitions à une fréquence d'une mesure par minute, assurant ainsi la cohérence temporelle des données utilisées pour l'analyse spatio-temporelle.

4.5.3 Résultats et discussions

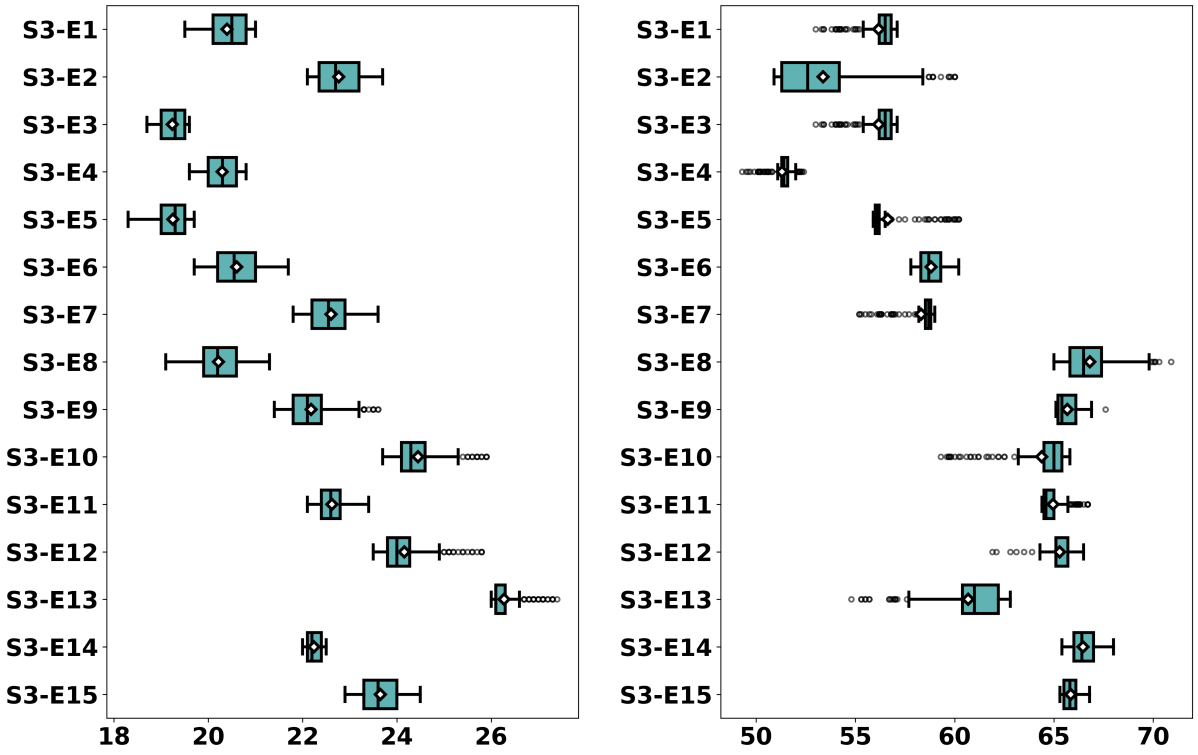
4.5.3.1 Déroulement des essais et conditions expérimentales

L'ensemble des essais planifiés a pu être réalisé comme prévu, entre le 23 et le 30 août 2024.

Essai N°	**Date**	**Grille d'échantillonnage**	**Nombre de sources**
$S3-E1$	23 Août	Grille 1	1
$S3-E2$	23 Août	Grille 1	1
$S3-E3$	26 Août	Grille 1	1
$S3-E4$	26 Août	Grille 1	1
$S3-E5$	27 Août	Grille 1	1
$S3-E6$	27 Août	Grille 1	1
$S3-E7$	27 Août	Grille 1	1
$S3-E8$	28 Août	Grille 1	1
$S3-E9$	28 Août	Grille 1	1
$S3-E10$	29 Août	Grille 1	1
$S3-E11$	28 Août	Grille 2	3
$S3-E12$	29 Août	Grille 2	3
$S3-E13$	29 Août	Grille 2	3
$S3-E14$	30 Août	Grille 2	3
$S3-E15$	30 Août	Grille 2	3

TABLEAU 4.17 – Liste des essais de la série 3 réalisés

Au cours de ces essais, les conditions moyennes, en termes de température et d'humidité relative de l'air, étaient respectivement de 23 ^{o}C et 60 % (Voir Figure 4.45).



(a) Température (°C)

(b) Humidité (%))

FIGURE 4.45 – Température (°C) et Humidité relative de l'air (%) mesurées au point 10 sur la grille d'échantillonnage 4.43 pour les essais de la série 3.

Ces essais ont été menés sur la base des scénarios de pollution définis en amont (voir Figure 4.46, représentant les grilles d'échantillonnage synthétiques associées à ces scénarios).

Plusieurs dysfonctionnements ont été observés au cours de la série d'essais, affectant les capteurs de mesure, les appareils de référence, ainsi que le robot mobile SHERPA. Plus précisément :

— Certains capteurs ont enregistré des concentrations en CO_2 anormalement élevées, probablement en raison de leur vieillissement et/ou des effets de la cause de vieillissement sur leur fonctionnement.

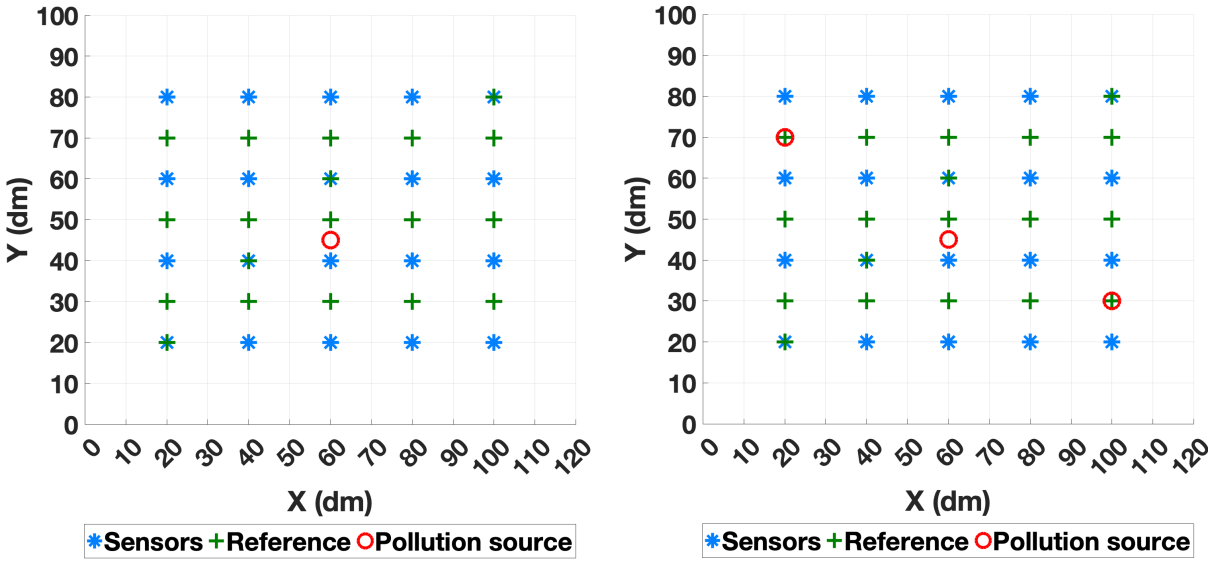
— Des interruptions ont été constatées en lien avec la décharge de la batterie de certains appareils de référence.

— Le robot SHERPA a rencontré des blocages lors de ses déplacements, entraînant une perte partielle ou totale des données de localisation pour certains essais.

— Pour l'essai S3-E8, l'enregistrement des données de déplacement a été interrompu pendant 7 minutes.

— Lors du premier essai S3-E1, une décharge prématurée de la batterie du robot est survenue. Cet incident a conduit à installer une zone dédiée à la recharge par induction pendant les phases d'inactivité du robot.

Les capteurs ayant émis de manière répétée des données incohérentes tout au long des essais,



(a) Grille du premier scénario de pollution (Grille 1)

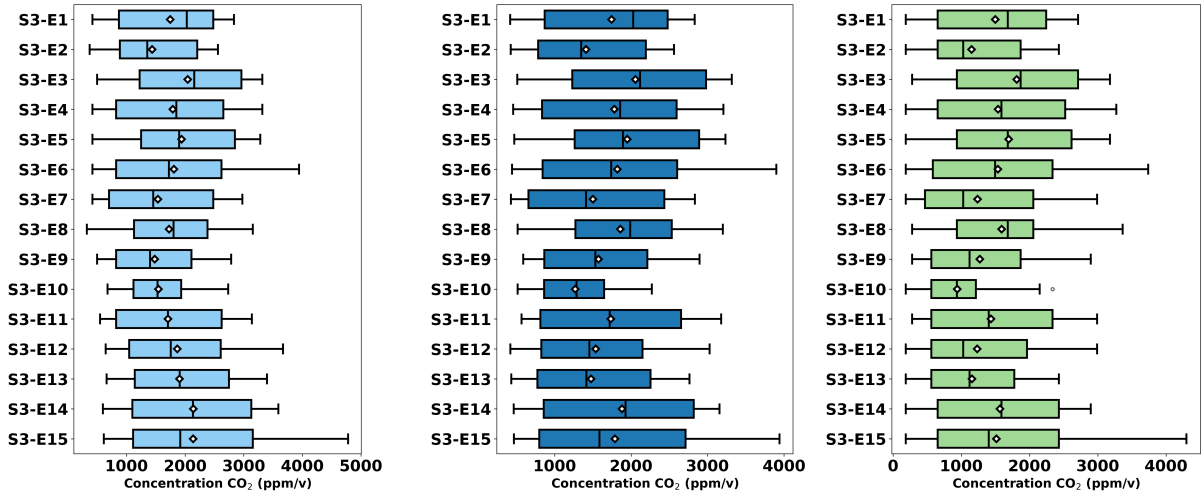
(b) Grille du second scénario de pollution (Grille 2)

FIGURE 4.46 – Représentation des grilles de mesure suivant les deux scénarios de pollution

ou ayant cessé de fonctionner suite à une décharge de batterie, ont été exclus de l'analyse. En revanche, les capteurs présentant uniquement quelques valeurs aberrantes ponctuelles ont été conservés, après application d'un processus de normalisation via interpolation linéaire. Les essais au cours desquels le robot SHERPA n'a pas respecté la trajectoire spécifiée dans le protocole ont été isolés de l'exploitation pour la fusion de données pour le processus de la reconstruction de l'exposition individuelle. Ainsi, sur les 15 essais initialement menés, quatre ont été écartés de l'analyse de fusion de données (essais S3-E1, S3-E8, S3-E11 et S3-E13). La génération de CO_2 au cours des essais s'est déroulée de manière satisfaisante, sans dysfonctionnement majeur.

4.5.3.2 Évolution des concentrations en CO_2 mesurées

Les concentrations mesurées par les capteurs ont varié de 200 à 4500 ppm/v (voir figure 4.47). Les concentrations ont été maintenues en deçà de la valeur limite d'exposition à court terme (VLCT) . Les concentrations médianes observées au cours des différents essais se situent entre 1100 et 2500 ppm/v. Bien que les essais aient été réalisés avec des protocoles identiques, il convient de souligner que l'absence d'étanchéité de la salle, ainsi que les fluctuations des conditions environnementales et météorologiques, peuvent avoir constitué une source de variabilité d'un essai à l'autre.



(a) Capteur RCSFg (b) Capteur Aranet (c) Détecteur de Référence

FIGURE 4.47 – Concentrations en CO_2 mesurées au cours des essais de la série 3 avec les capteurs du réseau Aranet (a), du RCSFg (b), détecteurs de référence (c) au point 10 de la grille d'échantillonnage 4.43.

Pour l'ensemble des essais, les concentrations de CO_2 mesurées présentent des profils temporels similaires, quel que soit le type de capteur utilisé (voir Figure 4.48). L'évolution des concentrations met ainsi en évidence trois phases successives : une augmentation progressive durant la phase de pollution, suivie d'une phase de stabilisation (caractérisée par un plateau ou une diminution plus ou moins lente selon les conditions météorologiques), puis une diminution graduelle des concentrations au cours de la phase d'épuration.

L'analyse des données repose sur l'exploitation conjointe des trois types de mesures : à points fixes, à points mobiles et à points de référence.

Par exemple, l'analyse des données issue des essais suivant les deux scénarios de pollution, respectivement les essais S3-E3 et S3-E14, a permis d'obtenir des données sur les points présentés sur les grilles (Figure 4.46).

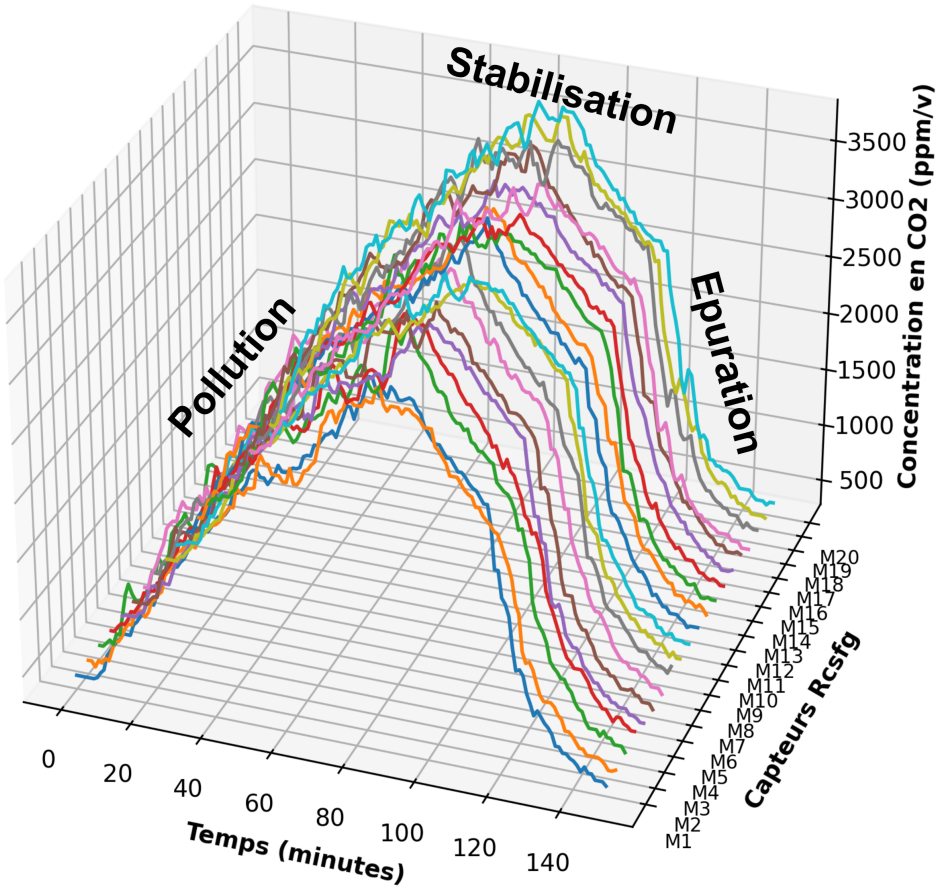


FIGURE 4.48 – Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) de l'essai S3-E3

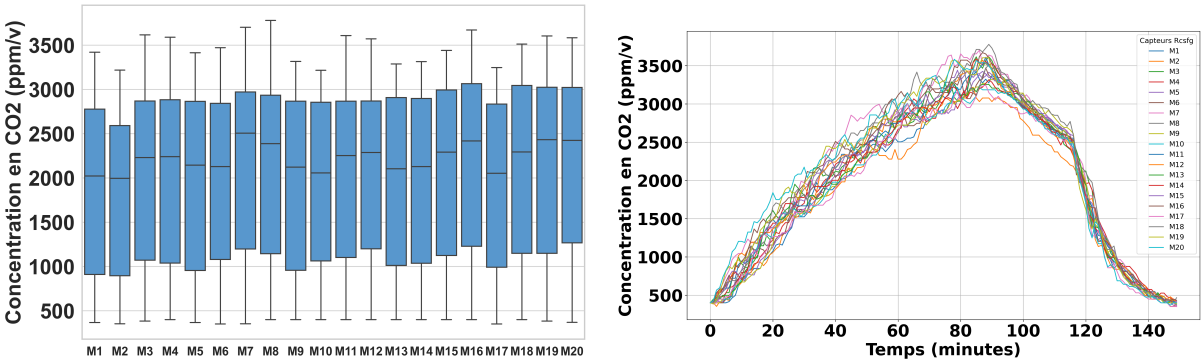


FIGURE 4.49 – Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) de l'essai S3-E3

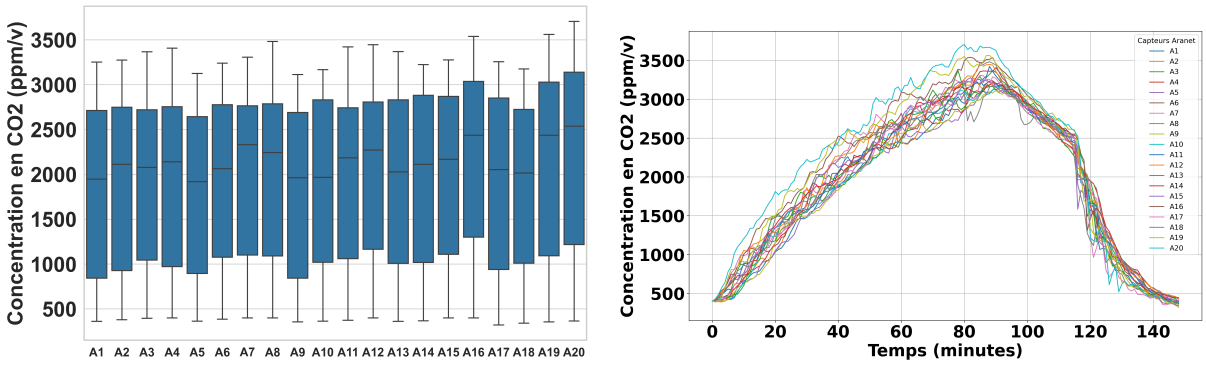


FIGURE 4.50 – Évolution des mesures de concentrations de CO_2 enregistrées par les capteurs du réseau Aranet (A1 à A20) de l'essai S3-E3

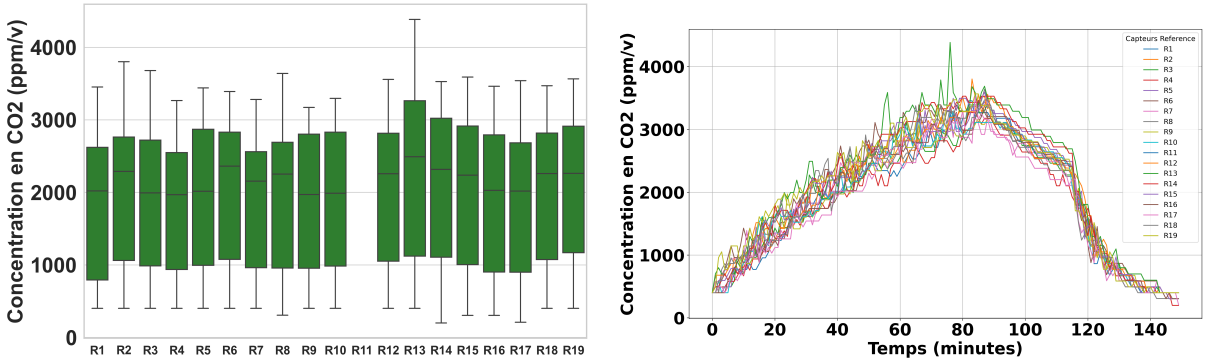


FIGURE 4.51 – Évolution des mesures de concentrations de CO_2 enregistrées par les capteurs de référence (R1 à R19) de l'essai S3-E3

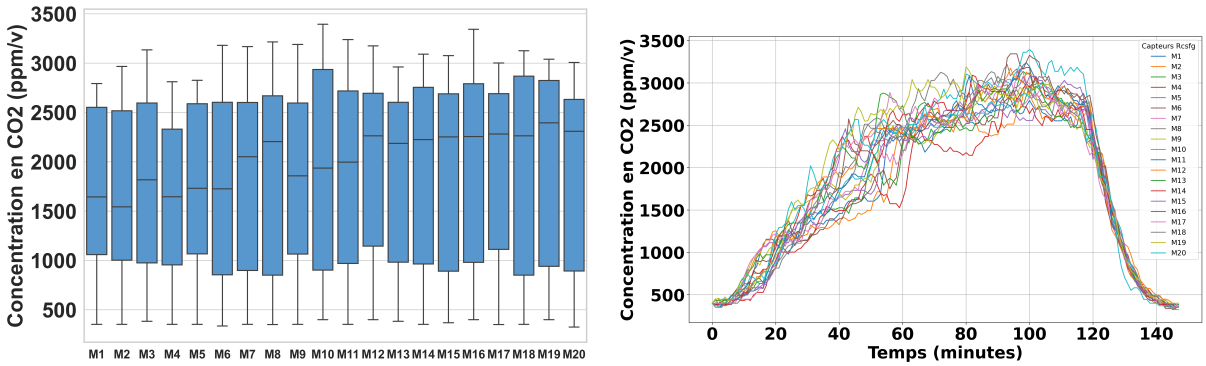


FIGURE 4.52 – Évolution des mesures de concentrations de CO_2 enregistrées par les modules capteurs du RCSFg (M1 à M20) de l'essai S3-E14

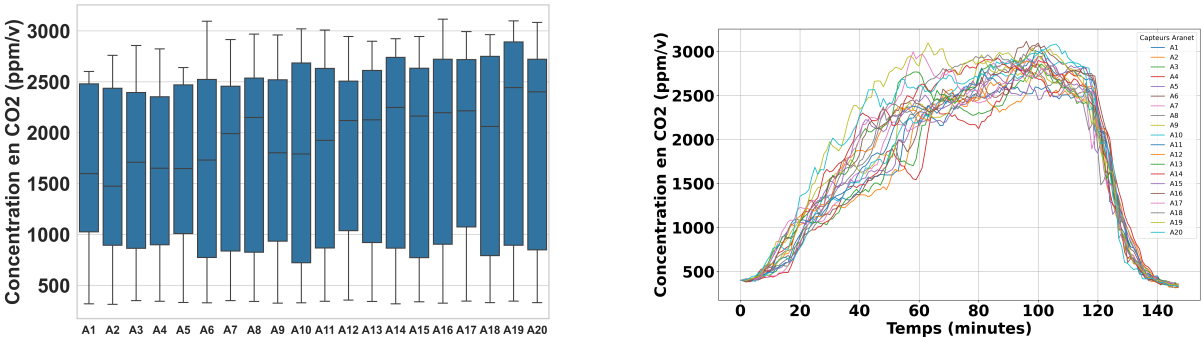


FIGURE 4.53 – Évolution des mesures de concentrations de CO_2 enregistrées par les capteurs du réseau Aranet (A1 à A20) de l'essai S3-E14

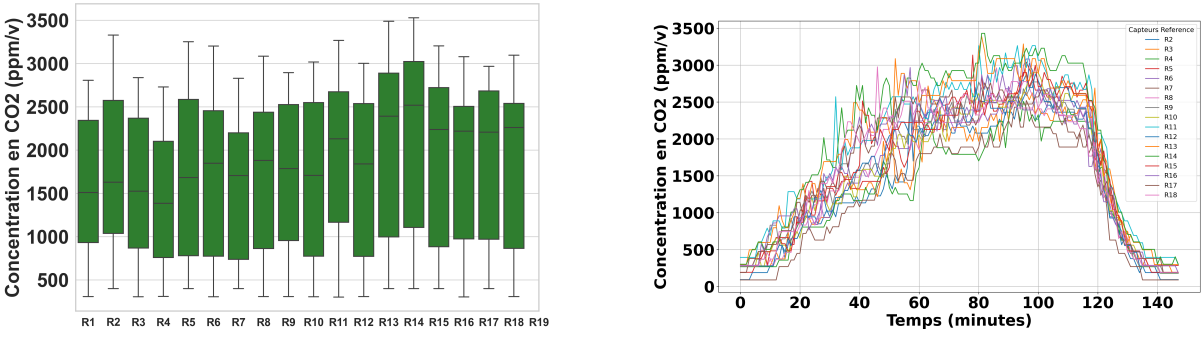
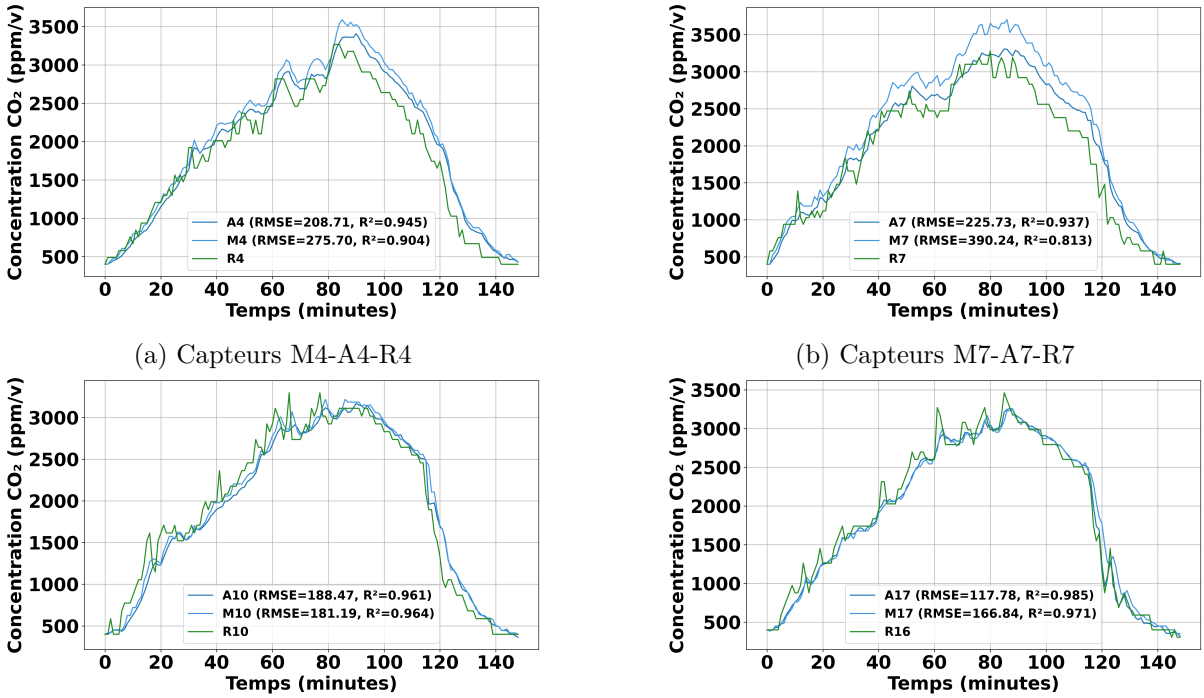


FIGURE 4.54 – Évolution des mesures de concentrations de CO_2 enregistrées par les détecteurs de référence (R2 à R19) de l'essai S3-E14

4.5.3.3 Validité des données de référence

La seconde étape vise à exploiter les données collectées par les réseaux de capteurs Aranet (Voir Figure 4.50, 4.53) et RCSFg (Voir Figure 4.49, 4.52) pour générer une cartographie spatio-temporelle des concentrations en CO_2. Avant d'engager le processus de fusion des données, il convient de vérifier la cohérence des cartes produites en les comparant, pour validation, aux données fournies par les détecteurs de référence (Voir Figure 4.51, 4.54). À cette fin, une première évaluation de l'écart entre les mesures issues des modules capteurs et celles des détecteurs de référence est réalisée aux points de co-localisation des instruments, à savoir : M4–R4, M7–R7, M10–R10 et M17–R16 de la grille d'échantillonnage 4.43 (voir Figures 4.55-4.56).



(a) Capteurs M4-A4-R4

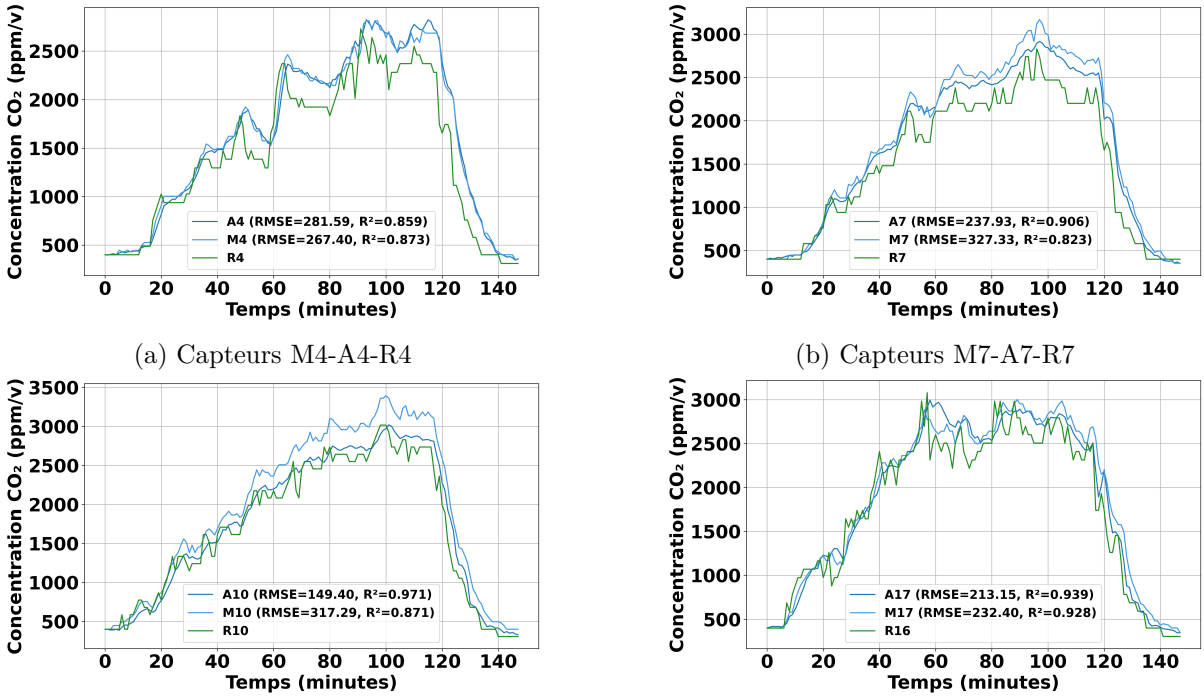
(b) Capteurs M7-A7-R7

(c) Capteurs M10-A10-R10

(d) Capteurs M17-A17-R16

FIGURE 4.55 – Comparaison des mesures de concentrations de CO_2 enregistrées au cours de l'essai S3-E3 aux points de co-localisation des capteurs des réseaux RCSFg et Aranet et les capteurs de référence ; M4-A4-R4 (a), M7-A7-R7 (b), M10-A10-R10 (c), M17-A17-R16 (d).

L'ensemble des modules capteurs a été préalablement étalonnés selon le protocole de calibration B (cf. Section 3.7.1). Toutefois les performances observées lors de la comparaison aux mesures de référence varient sensiblement d'un capteur à l'autre et d'un essai à l'autre. L'évaluation de ces performances repose sur le calcul des métriques classiques d'erreur, notamment la RMSE et le coefficient R^2 aux points de co-localisation entre modules capteurs « low-cost » et détecteurs de référence. En moyenne, les valeurs obtenues pour ces métriques sont généralement satisfaisantes, avec un coefficient R^2 globalement supérieur à 0.75 , ce qui traduit une bonne corrélation entre les données de référence et les données issues des réseaux de capteurs. L'évaluation de la validité des cartographies générées sur l'ensemble de la zone étudiée, s'est appuyée



(a) Capteurs M4-A4-R4

(b) Capteurs M7-A7-R7

(c) Capteurs M10-A10-R10

(d) Capteurs M17-A17-R16

FIGURE 4.56 – Comparaison des mesures de concentrations de CO_2 enregistrées au cours de l'essai S3-E14 aux points de co-localisation des capteurs des réseaux RCSFg et Aranet et les détecteurs de référence ; M4-A4-R4 (a), M7-A7-R7 (b), M10-A10-R10 (c), M17-A17-R16 (d).

sur les 19 points de référence répartis sur la grille d'échantillonnage (Figure 4.43). Leurs performances sont à l'image de l'échantillon des 4 points de co-localisation que nous avons examinés. Les tableaux récapitulatifs des valeurs de RMSE et de R^2 obtenues pour l'ensemble des essais de la série 3, aux quatre points de co-localisation des réseaux RCSFg et Aranet, sont présentés en annexes (Voir tableaux 22, 23).

4.5.3.4 Cartographie spatio-temporelles des concentrations de CO_2 au sein de salle expérimentale

Les méthodes de cartographie spatio-temporelle décrites dans la section 2.7 ont été appliquées aux données produites par les essais expérimentaux décrits dans la présente section (voir Figures 4.49–4.54). (Voir figures 4.57-4.62). Les cartographies présentées illustrent l'état spatio-temporel des concentrations de CO_2 au sein de la salle expérimentale pour trois instants représentatifs de l'expérience : $T = 25$ min pour la phase de pollution, $T = 90$ min pour la phase de stabilisation et $T = 115$ min pour la phase d'épuration.

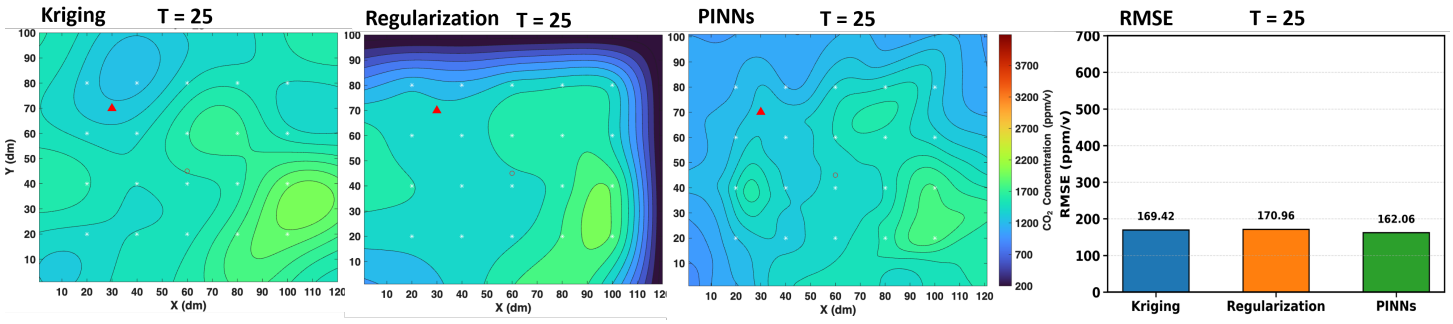


FIGURE 4.57 – Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase de pollution à $T = 25$ min de l'essai S3 E3 avec le réseau Aranet

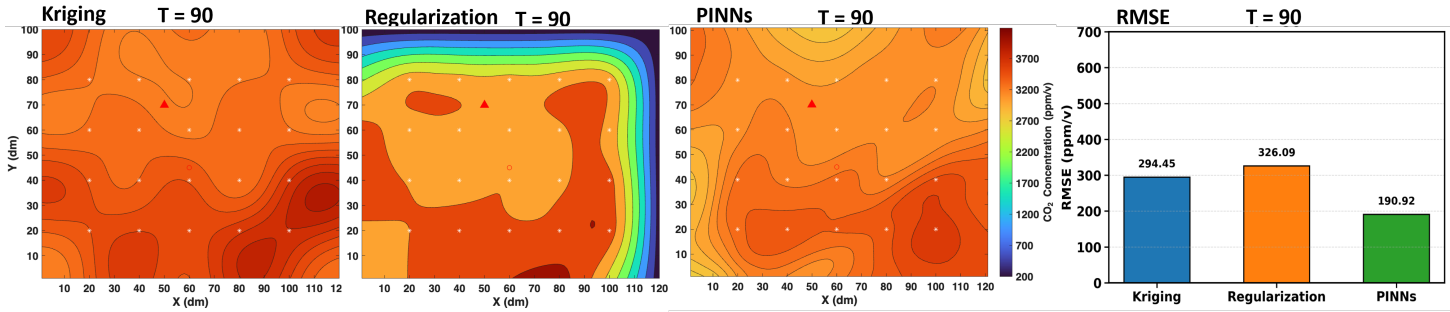


FIGURE 4.58 – Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase de stabilisation à $T = 90$ min de l'essai S3 E3 avec le réseau Aranet

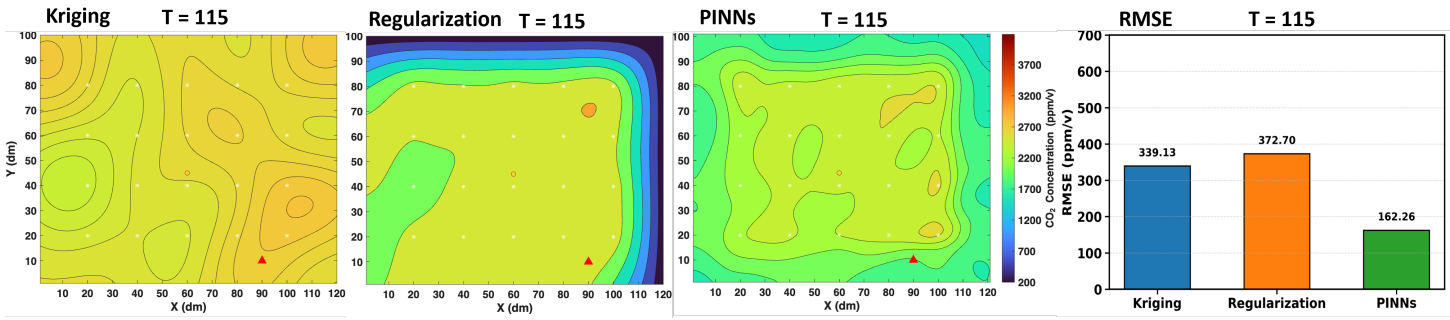


FIGURE 4.59 – Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase d'épuration à $T = 115$ min de l'essai S3 E3 avec le réseau Aranet

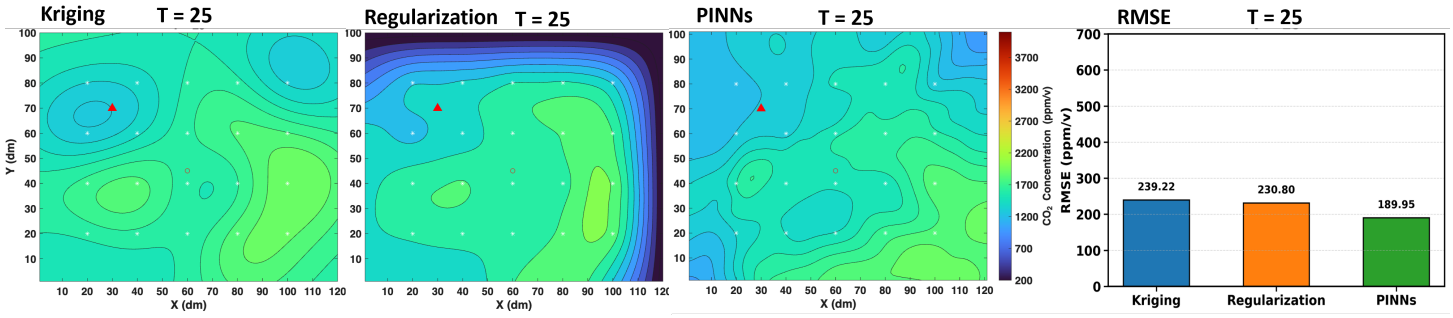


FIGURE 4.60 – Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase de pollution à $T = 25$ min de l'essai S3 E3 avec le réseau RCSFg

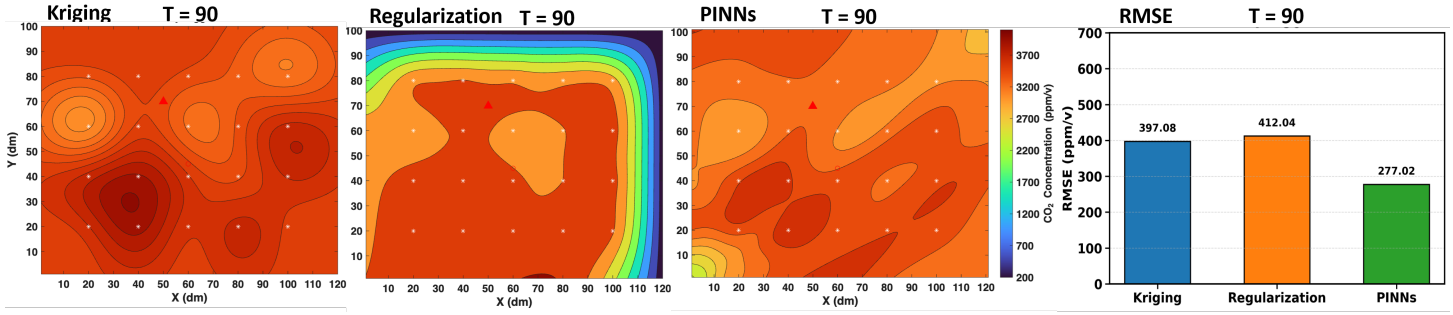


FIGURE 4.61 – Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase de stabilisation à $T = 90$ min de l'essai S3 E3 avec le réseau RCSFg

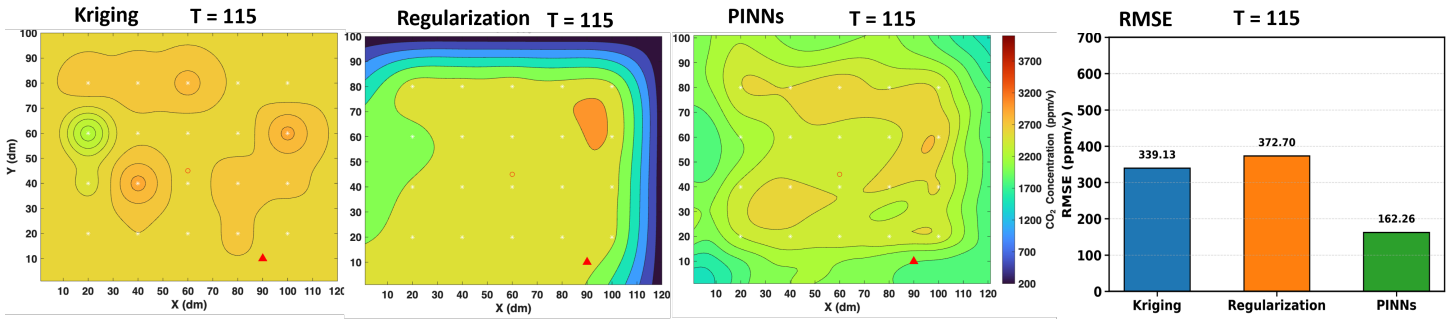


FIGURE 4.62 – Cartographie spatio-temporelle des concentrations de CO_2 et évaluation du RMSE pendant la phase d'épuration à $T = 115$ min de l'essai S3 E3 avec le réseau RCSFg

Afin d'évaluer la fiabilité des cartes générées à partir des données expérimentales, une comparaison systématique a été réalisée entre les différentes méthodes de cartographie spatio-temporelle considérées (Krigeage, Régularisation tensorielle, PINNs). Comme discuté dans la section 4.3, la densité et la régularité de notre grille d'échantillonnage assurent une représentativité satisfaisante des concentrations de CO_2, permettant une bonne évaluation des reconstructions.

La qualité des estimations a été évaluée à chaque instant à l'aide du calcul du RMSE. À l'instant $T = 25$ (Figures 4.57,4.60) : les différentes méthodes génèrent des champs continus et globalement conformes à la distribution attendue. Une zone de concentration accrue est obser-

vée depuis le centre vers la partie droite du domaine, ce qui peut s'expliquer par la position de la source de pollution : un tube de diffusion de CO_2 situé au centre de la salle, dont l'émission est influencée par les caractéristiques du dispositif de confinement (bande adhésive utilisée) et, potentiellement, par une pression d'injection légèrement excessive induisant une diffusion suivant un sens préférentiel. À l'instant $T = 90$ (Figures 4.58, 4.61) : Les concentrations tendent vers une distribution plus homogène à l'échelle spatiale. Les cartes reconstruites restituent correctement cette dynamique. Les PINNs se distinguent par leur capacité à restituer les hétérogénéités spatio-temporelles avec un RMSE plus faible. À l'instant $T = 115$ (Figures 4.59, 4.62), les cartes obtenues rendent compte de manière satisfaisante de la décroissance progressive des concentrations. Le krigeage restitue une structure globale cohérente, mais tend à générer des artefacts, en particulier dans les zones de forte concentration. Les PINNs offrent une représentation beaucoup plus fidèle de la dynamique expérimentale, en mettant en évidence l'effet de la ventilation à travers les ouvertures situées aux extrémités gauche et droite. Cette dynamique se traduit visuellement par une décroissance graduelle des concentrations depuis les bords vers la zone centrale.

Des vidéos complémentaires illustrant les cartographies spatio-temporelles obtenues pour l'ensemble des méthodes considérées sont disponibles en ligne.

- Vidéo de la cartographie spatio-temporelle de l'essai S3-E3 (Avec les données du réseau Aranet) avec la méthode de krigeage
- Vidéo de la cartographie spatio-temporelle de l'essai S3-E3 (Avec les données du réseau Aranet) avec la méthode de régularisation tensorielle
- Vidéo de la cartographie spatio-temporelle de l'essai S3-E3 (Avec les données du réseau Aranet) avec la méthode PINNs
- Vidéo de la cartographie spatio-temporelle de l'essai S3-E3 (Avec les données du réseau RCSFg) avec la méthode de krigeage
- Vidéo de la cartographie spatio-temporelle de l'essai S3-E3 (Avec les données du réseau RCSFg) avec la méthode de régularisation tensorielle
- Vidéo de la cartographie spatio-temporelle de l'essai S3-E3 (Avec les données du réseau RCSFg) avec la méthode PINNs

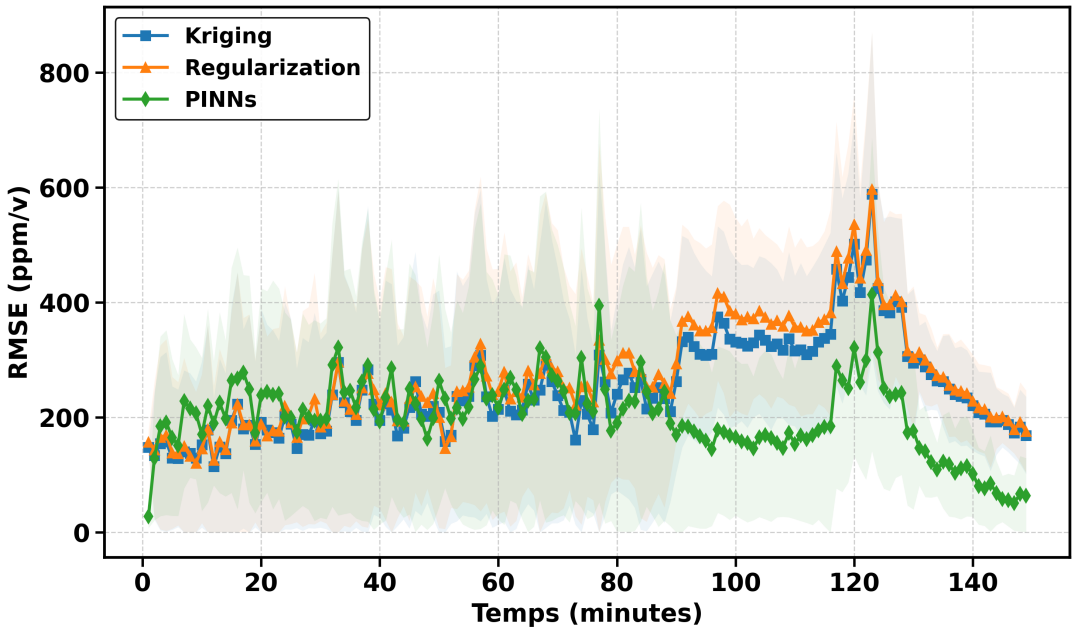


FIGURE 4.63 – Évolution temporelle du RMSE obtenues avec les méthodes de krigeage, de régularisation et de PINNs pour la reconstruction spatio-temporelle des concentrations de CO_2 | Essai S3 E3 avec le réseau Aranet

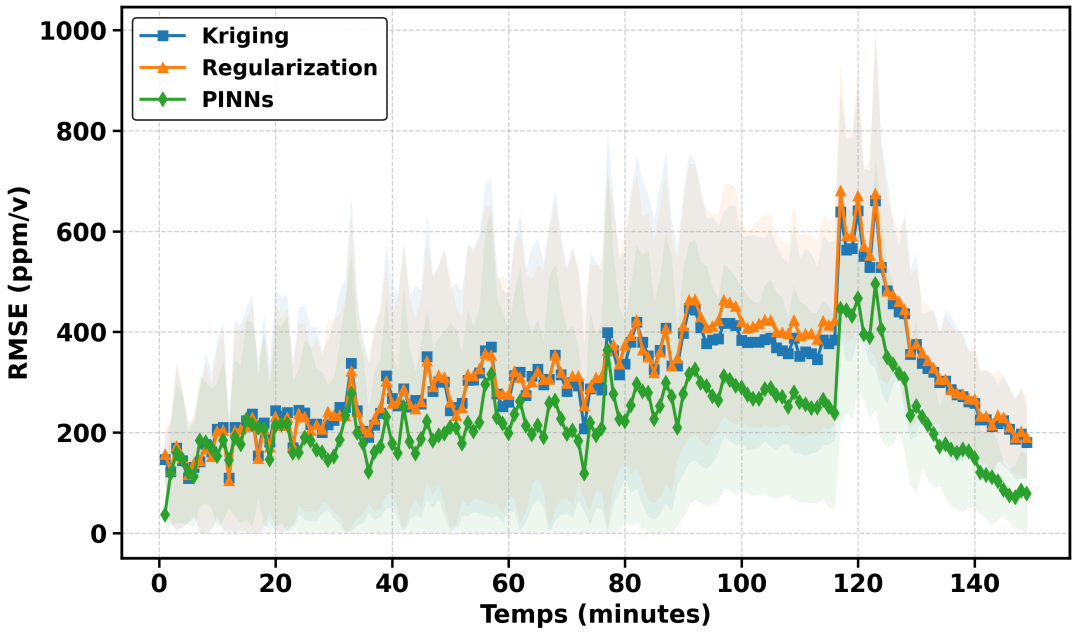


FIGURE 4.64 – Évolution temporelle du RMSE obtenues avec les méthodes de krigeage, de régularisation et de PINNs pour la reconstruction spatio-temporelle des concentrations de CO_2 | Essai S3 E3 avec le réseau RCSFg

Au-delà de ces instants représentatifs, les figures 4.63, 4.64 présentent l'évolution temporelle du RMSE sur l'ensemble de la période expérimentale. Le krigeage et la régularisation présentent des erreurs comparables et relativement stables, avec des fluctuations accrues lors des phases de forte variabilité des concentrations. Les PINNs, en revanche, mettent en évidence une amélioration notable de la précision, en particulier dans les régimes complexes où les données à elles seules ne suffisent pas pour comprendre les dynamiques spatio-temporelles de l'expérience. Dans ces conditions, l'intégration de contraintes physiques permet de mieux suivre les variations spatio-temporelles.

Dans l'ensemble, ces résultats confirment la capacité des PINNs à fournir des reconstructions spatio-temporelles plus précises que les approches classiques, en particulier lors des phases transitoires. Le krigeage, bien que compétitif, tend à atténuer les variations locales et à lisser les hétérogénéités spatiales, limitant ainsi sa capacité à restituer fidèlement la dynamique réelle. Quant à la régularisation, elle se révèle globalement moins adaptée à la représentation précise des phénomènes, en particulier dans un environnement expérimental moins contrôlé que la chambre expérimentale de l'INRS.

4.5.3.5 Fusion de données : Reconstruction de l'exposition individuelle

La fusion de données pour la reconstruction de l'exposition individuelle consiste à combiner les estimations issues des cartographies spatio-temporelles de concentration avec les coordonnées de la trajectoire d'un individu (dans notre contexte le robot mobile) dans l'espace étudié.

Soit $\hat{\boldsymbol{X}}(i,j,t)$ la concentration reconstruite avec une des méthodes de cartographies spatio-temporelles (krigeage, régularisation, PINNs) et $\theta(t) = [i(t), j(t)]$ la trajectoire spatio-temporelle discrétisée de l'individu, avec $i(t)$ et $j(t)$ les coordonnées spatiales dans le plan $x-y$, définies aux instants $t \in [0, T]$. L'exposition instantanée au cours du temps est donnée par :

$$\hat{\boldsymbol{X}}_\theta(t) = \hat{\boldsymbol{X}}(\theta(t), t). \tag{4.2}$$

Dans notre étude, elle correspond à la valeur reconstruite par la cartographie au point (i,j) de la trajectoire considérée, à un instant t donné. Conformément au plan expérimental, cette concentration est également mesurée simultanément par des détecteurs de référence, permettant ainsi une comparaison directe entre les valeurs estimées et les mesures de référence.

À partir de cette exposition individuelle reconstruite, l'exposition cumulée sur la durée T est définie par :

$$\hat{E}(T) \approx \sum_{k=1}^{N_t} \hat{\boldsymbol{X}}(\theta(t_k), t_k)\, \Delta t \tag{4.3}$$

où $E(T)$ représente la dose totale de polluant inhalée jusqu'à l'instant T, Δt le pas temporel d'acquisition et N_t le nombre d'échantillons temporels ($N_t = \lfloor \frac{T}{\Delta t} \rfloor$). Nous pouvons donc quan-

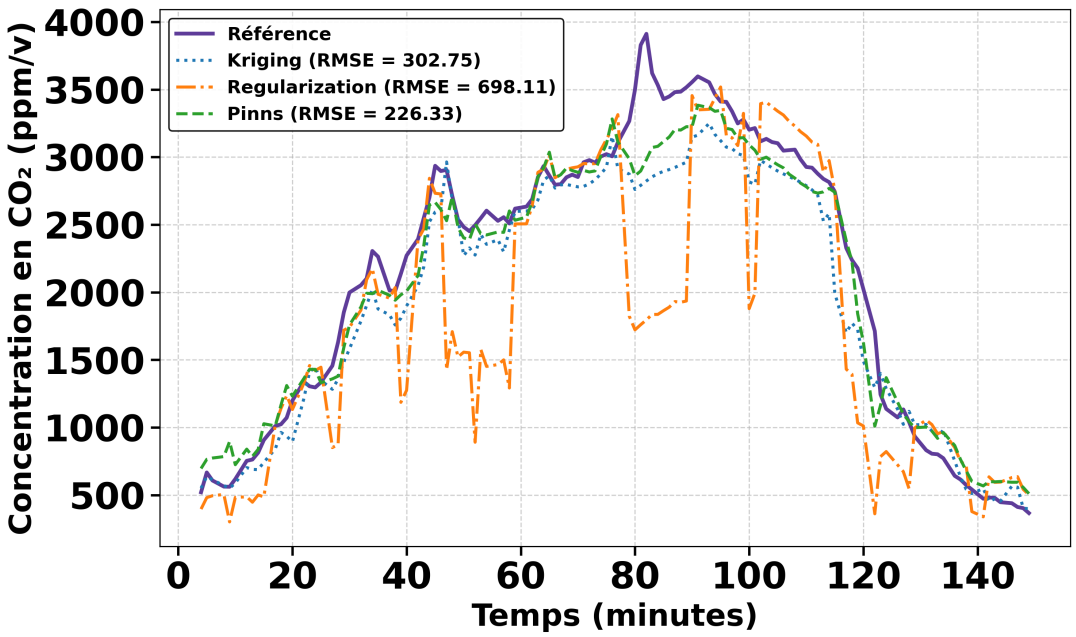


FIGURE 4.65 – Évolution temporelle de l'exposition individuelle au CO_2 reconstruite $\hat{\boldsymbol{X}}_\gamma(t)$ avec les méthodes de krigeage, de régularisation et de PINNs, comparée aux mesures de référence | Essai S3 E3 avec le réseau Aranet

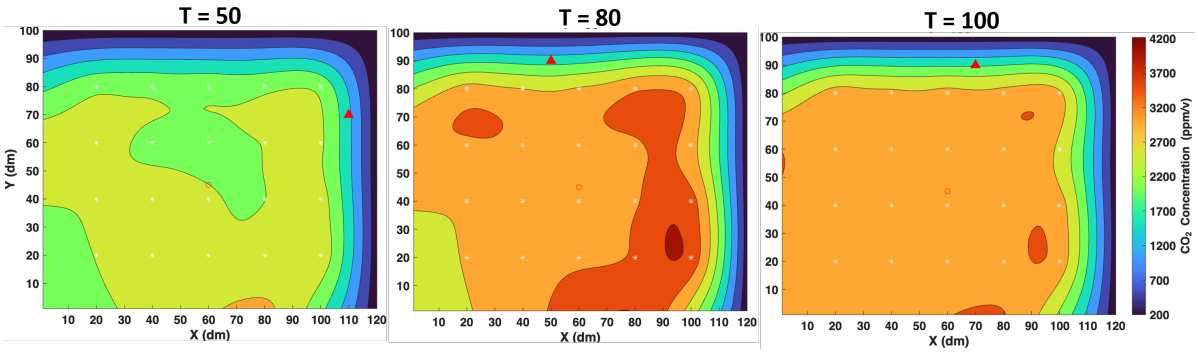


FIGURE 4.66 – Cartographie spatio-temporelle des concentrations de CO_2 avec le méthode de régularisation tensorielle pendant l'essai S3-E3 avec le réseau Aranet à $T = 50$ min, $T = 80$ min, $T = 100$ min représentant des instants où le vecteur mobile ▲ se trouve en bordure de l'espace étudié

tifier l'écart entre cette valeur estimée et l'exposition de référence en utilisant des métriques telles que le RMSE et le R^2.

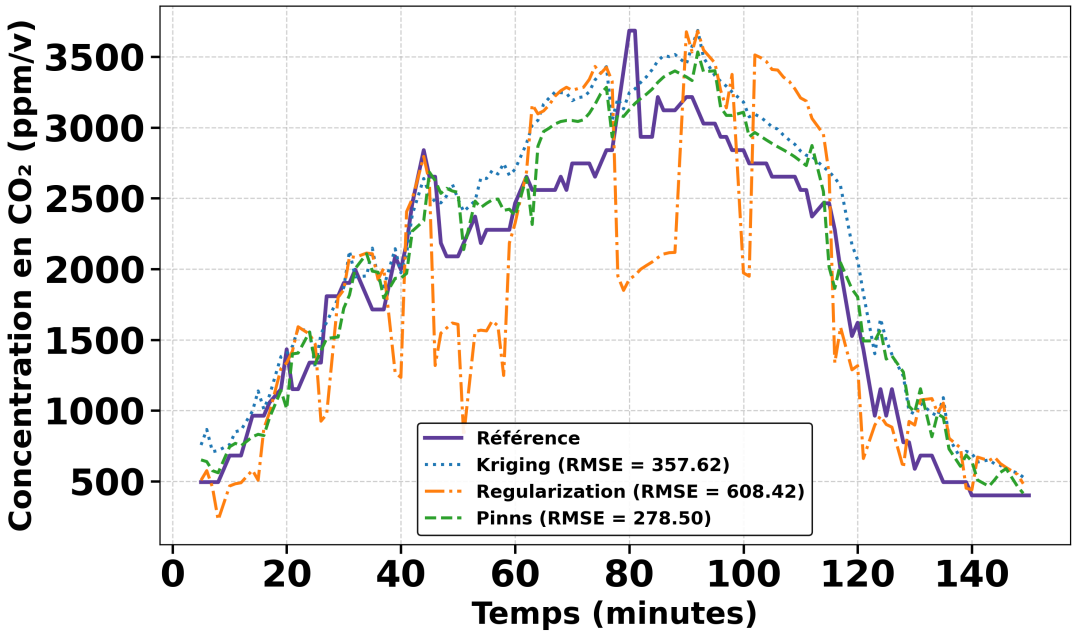


FIGURE 4.67 – Évolution temporelle de l'exposition individuelle au CO_2 reconstruite $\hat{\boldsymbol{X}}_\theta(t)$ avec les méthodes de krigeage, de régularisation et de PINNs, comparée aux mesures de référence | Essai S3 E3 avec le réseau RCSFg

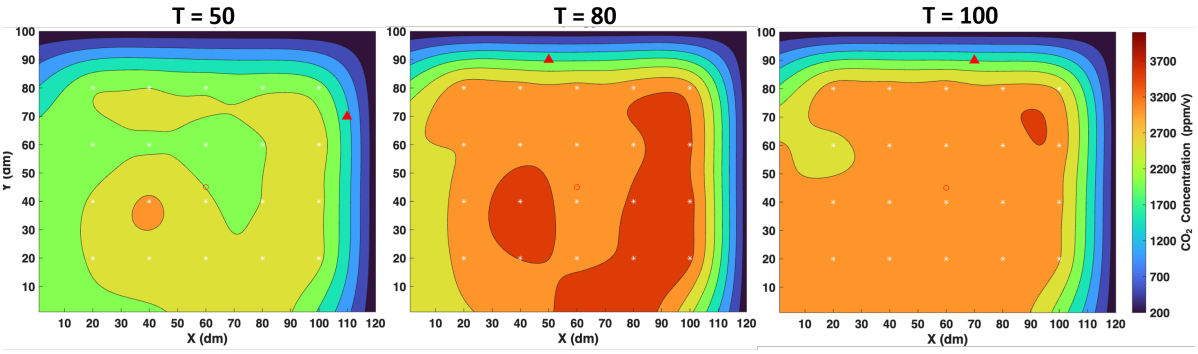


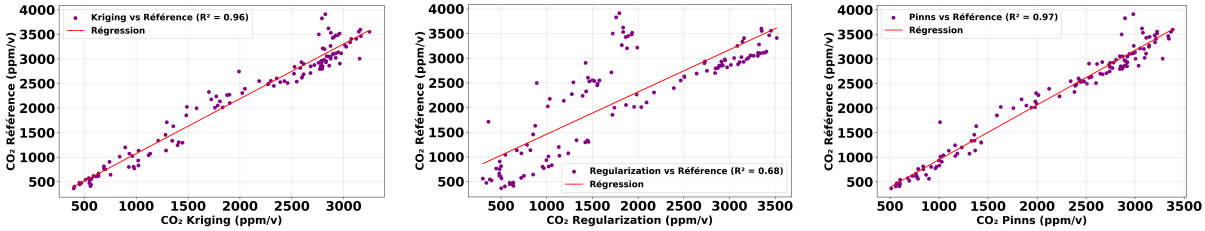
FIGURE 4.68 – Cartographie spatio-temporelle des concentrations de CO_2 avec le méthode de régularisation tensorielle pendant l'essai S3-E3 avec le réseau RCSFg à $t = 50$ min, $t = 80$ min, $t = 100$ min représentant des instants où le vecteur mobile ▲ se trouve en bordure de l'espace étudié

Les figures 4.65, 4.67 présentent cette comparaison temporelle respectivement pour le réseau Aranet et le RCSFg. Les PINNs présentent la plus faible erreur traduisant une capacité à

reproduire fidèlement la dynamique expérimentale. Le krigeage obtient une performance avec une tendance au lissage des extrêmes, tandis que la régularisation affiche une erreur plus élevée, en particulier lorsque le vecteur mobile se situe à proximité des bords de l'espace cartographié. Cette baisse de performance de la méthode de régularisation tensorielle est dû aux différentes positions du vecteur mobile pendant sa trajectoire. Aux instants où la valeur reconstruite est totalement décorrélée de la valeur de référence, le vecteur mobile se trouve aux bordures de l'espace cartographié (Voir Figures 4.66, 4.68). En effet, on observe sur ces cartes un effet de bord marqué sur les parties supérieure et droite en périphériques de l'espace, où les concentrations sont systématiquement mal estimées. Ces artefacts sont dus à la structure unilatérale des opérateurs de dérivation discrets utilisés dans la régularisation. Les matrices de différentiation $\boldsymbol{D}_x$ et $\boldsymbol{D}_y$ reposent sur des schémas de différences finies orientés vers le bas et vers la gauche : les valeurs situées sur la première ligne et sur la dernière colonne ne sont donc jamais régularisées par des différences « en amont ». Ces nœuds restent faiblement contraints par le terme de pénalisation, ce qui induit un champ reconstruit asymétrique aux frontières lorsque la régularisation est appliquée séparément sur les lignes et sur les colonnes. Lorsque la trajectoire $\theta(t)$ traverse ces zones (par exemple aux instants $t = 50$, 80 et 100 min en (Voir Figures 4.66, 4.68), l'exposition reconstruite hérite directement de ces biais locaux, ce qui explique l'erreur élevée observée pour la méthode de régularisation tensorielle dans les courbes temporelles.

Ces résultats confirment que la qualité des cartes $\hat{\boldsymbol{X}}(i, j, t)$ est déterminante pour pourvoir les utiliser pour estimer l'exposition individuelle : plus la méthode réduit le RMSE vis-à-vis de la référence, plus l'estimation $\hat{E}(T)$ se rapproche de l'exposition réelle mesurée $E(T)$.

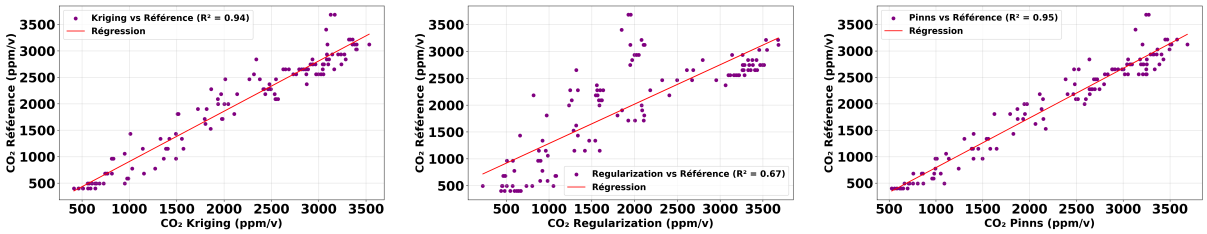
Afin d'évaluer plus finement la précision des reconstructions, les figures 4.70-4.69 présentent les relations entre les expositions individuelles reconstruites et celles mesurées par les détecteurs de référence. Les résultats mettent en évidence une corrélation satisfaisante pour les trois méthodes, mais avec des performances contrastées.



(a) Kriging (b) Régularisation (c) PINNs

FIGURE 4.69 – Corrélation entre exposition individuelle reconstruite $\hat{E}(T)$ et exposition de référence $E(T)$ | Essai S3 E3 avec le réseau Aranet

Les résultats des essais expérimentaux montrent qu'il est possible d'estimer l'exposition individuelle d'un opérateur à partir de la cartographie spatio-temporelle des concentrations en polluants et de la trajectoire de ce dernier. Des travaux antérieurs ont déjà abordé ce sujet. Par



(a) Kriging (b) Régularisation (c) PINNs

FIGURE 4.70 – Corrélation entre exposition individuelle reconstruite $\hat{E}(T)$ et exposition de référence $E(T)$ | Essai S3 E3 avec le réseau RCSFg

exemple [Ludwig17], a proposé une méthodologie de fusion de données issues de capteurs statiques et mobiles pour la cartographie spatio-temporelle des nuisances dans un environnement industriel. Leur modèle semi-paramétrique, basé sur une décomposition de Karhunen–Loève, permettait d'intégrer des mesures à résolutions spatiale et temporelle hétérogènes afin de produire des cartes dynamiques plus représentatives que les méthodes classiques de krigeage stationnaire. Toutefois, leurs travaux restaient centrés sur la représentation des distributions de concentration sans mise en relation explicite avec la trajectoire individuelle d'un opérateur.

De même, [Zuidema19a] a démontré qu'un réseau dense de capteurs multi-polluants (PM, CO, O_3 et bruit) pouvait être exploité pour estimer l'exposition personnelle à partir de la localisation simulée d'un opérateur. Les auteurs ont eu recours à une interpolation spatiale par pondération inverse de la distance (IDW) pour reconstituer les concentrations le long de trajectoires mobiles. Malgré une bonne corrélation entre les estimations du réseau et les mesures personnelles ($r^2 = 0,39 - 0,75$), la précision demeurait limitée par la calibration des capteurs et la résolution temporelle du réseau.

Plus récemment, [Ruiter24] ont approfondi cette approche en combinant capteurs personnels à bas coût et systèmes de localisation « indoor » pour modéliser les concentrations de particules fines en temps réel dans un atelier de menuiserie. L'utilisation du krigeage spatio-temporel a permis de générer des cartes dynamiques de $PM_{2.5}$ à partir de trajectoires réelles de travailleurs. Néanmoins, la validation externe a révélé des corrélations faibles à modérées, avec les capteurs fixes ($R^2 = 0,1 - 0,5$), traduisant encore une incertitude importante liée à la variabilité des mesures, aux limitations des capteurs « low-cost » mais également à la résolution temporelle trop grande des cartes (16 minutes en laboratoire et 45 minutes en entreprise) qui a été imposé pour permettre l'utilisation du krigeage spatio-temporel sur l'ensemble de l'environnement étudié.

Dans l'ensemble, ces travaux précurseurs confirment la faisabilité de la reconstruction de l'exposition individuelle par fusion de données spatiales, temporelles et de trajectoires, mais il est important de souligner la nécessité d'améliorer la calibration des capteurs, la robustesse des modèles de cartographie spatio-temporelle et l'intégration des itinéraires avec une bonne synchronisation spatio-temporelle pour faciliter la fusion de données.

L'acquisition synchronisée à une fréquence d'une mesure par minute a permis d'obtenir des données temporelles cohérentes, exploitables par les modèles de cartographie spatio-temporelle développés. Cette configuration assure une résolution temporelle fine, garantissant la fusion des champs de concentration avec la trajectoire $\theta(t)$ du vecteur mobile, et permettant ainsi la reconstruction de l'exposition instantanée $\hat{X}_\theta(t)$ ainsi que de l'exposition cumulée $\hat{E}(T)$.

Les méthodes proposées (Régularisation tensorielle, PINNs) et la méthode de krigeage ont été comparées via le RMSE et le R^2. Nous avons vérifié la conformité des champs reconstruits aux mesures de référence. La corrélation est globalement élevée lorsqu'on n'intègre pas les effets de bords générés par la méthode de régularisation tensorielle ($R^2 > 0,75$, dans la plupart des cas). Cependant, si nous tenons compte de l'entièreté de l'espace étudié nous avons $0,41 < R^2 < 0,94$. L'analyse temporelle du RMSE montre des écarts plus marqués lors des régimes transitoires (début/fin de pollution, ouverture/fermeture, phases d'épuration).

Les PINNs réduisent l'erreur en régime complexe en contraignant la dynamique par la physique (diffusion/ventilation), quand le krigeage a tendance à lisser les extrêmes et la régularisation à sous-représenter certaines hétérogénéités.

4.6 Applications en santé au travail : retombées indirectes des travaux de thèse

La cartographie spatio-temporelle des polluants en milieu professionnel s'est imposée ces dernières années comme un levier méthodologique complémentaire aux approches conventionnelles d'évaluation des expositions (mesure en différé à point fixe et en individuel) . En exploitant des instruments de mesure temps réel (IMTR) organisés en réseaux fixes, portés par un opérateur ou par un vecteur mobile, cette approche permet de représenter dynamiquement les concentrations en polluants, de localiser des « points chauds » de concentration, de relier les pics aux activités, et d'évaluer l'efficacité des mesures de prévention en conditions réelles.

Dans un contexte de santé au travail, cette cartographie peut permettre non seulement d'identifier des zones et des moments d'exposition élevée (ce qui oriente la prévention) mais aussi de communiquer envers les acteurs sur le risque de façon visuelle et de suivre l'impact des mesures correctives.

Ainsi, dans les environnements de travail, la variabilité spatiale (différents postes, zones de circulation, machines, etc.) et temporelle (cycles de production, fluctuations de ventilation, interruptions, etc.) est élevée, la cartographie spatio-temporelle offre une valeur ajoutée tangible pour la stratégie de prévention et la santé des travailleurs.

4.6.1 Apport sur les stratégies de mesure

Toute réalisation de mesures en environnement de travail nécessite de définir des objectifs précis pour les mesures (par exemple, mesurer les niveaux d'exposition, vérifier l'efficacité des moyens de prévention en place, etc.). Elle requiert aussi l'établissement d'une stratégie de mesure. La stratégie de mesure comprend tous les éléments techniques nécessaires à la réalisation des mesures et repose sur des informations contextuelles qui ont souvent déjà été collectées dans l'entreprise ou faisant l'objet de l'enquête. Il s'agit notamment de préciser comment (méthodes de mesures, mesures à point fixe ou individuel), ou et sur quels opérateurs, quand et à quelle fréquence, combien de fois, les mesures doivent être effectuées.

Positionnement d'un point fixe dans un espace de travail

La mesure à point fixe est souvent nécessaire à l'occasion d'une campagne de mesure. Elle accompagne aussi souvent les mesures effectuées en individuel soit pour mettre en évidence un phénomène particulier (une source d'émission) ou pour constituer une référence caractérisant l'ambiance de travail (concentration ambiante dans l'espace investigué). Si plusieurs articles ou normes ont été publiés pour aider à la construction d'une stratégie de mesures, il n'existe pas de recommandation explicite concernant le choix du positionnement d'un point fixe et celui-ci est souvent laissé à l'expert qui prépare les mesures.

Dans ce sens les travaux présentés dans ce manuscrit apportent des informations intéressantes qui peuvent aider l'expert dans son choix. En effet, la cartographie spatio-temporelle telle qu'expérimentée dans les travaux de thèse reposent sur le déploiement à point fixe d'IMTR sur une grille prédéfinie. Toutes les cartes présentées aux figures 4.8, 4.9, 4.10, 4.31 et courbes d'évolution des concentrations aux figures 4.5, 4.6, 4.7, 4.30 dans le manuscrit montrent une variation des concentrations dans l'espace et dans le temps. Autrement dit, il est clairement démontré que les résultats de mesure dépendent de la position du point mesure dans l'environnement étudié. Par exemple, les résultats montrent qu'une mesure au même instant, à un endroit donné n'est pas forcément représentative de celle effectuée à un autre endroit dans l'espace (voir Figure 4.6). Ainsi, si le positionnement d'un point de mesure fixe donné a pour objectif de donner une référence pour la concentration ambiante dans un espace de travail, il faudra relativiser la représentativité de la mesure effectuée à point fixe dans l'interprétation. De même, les résultats montrent que si la cartographie spatio- temporelle permet d'identifier la position d'une source dans certains cas (par exemple pour les aérosols voir Figure 4.39), ce qui n'est pas forcément le cas dans d'autres contextes (par exemple pour le CO_2 avec variation du débit de pollution voir Figure 4.31). Ainsi, si le positionnement d'un point de mesure fixe donné a pour objectif de situer une source dans un espace de travail, il sera nécessaire de considérer que le polluant ciblé (densité), les conditions d'émission (débit, orientation de la source) et les méthodes de mesures

(hauteurs des points de mesures) sont des facteurs intervenant dans la dynamique de polluant.

Intérêts d'intégrer les IMTR dans les stratégie de mesure

L'intégration d'un réseau d'IMTR peut compléter les stratégies de mesure des expositions en santé au travail. Traditionnellement, l'expert en prévention met en œuvre des méthodes de mesure en différé qui consistent à prélever des échantillons en entreprise qui sont ensuite transportés au laboratoire pour analyse ultérieure. Il met en œuvre des méthodes et matériels adaptés (filtre, pompe, appareils de mesures etc.) au poste de travail, souvent avec un nombre limité d'échantillons. Les résultats sont donnés en concentrations moyennes horaires ou journalières. La mesure en différé , bien qu'incontournable, présente plusieurs limites : faible résolution temporelle, peu de repères spatiaux (un ou deux points), et interpolation importante entre les mesures. Les travaux effectués dans le cadre de la thèse indiquent qu'un réseau d'IMTR peut présenter plusieurs avantages :

— **Acquérir des données à haute résolution temporelle** typiquement à des fréquences de l'ordre de 1 à 5 minutes ce qui permet de caractériser finement la dynamique des concentrations dans l'environnement de travail. Cette résolution offre la capacité de détecter et de quantifier les fluctuations rapides, notamment les pics d'exposition de courte durée ainsi que les phases de faible concentration, souvent invisibles dans les mesures en différé qui donnent qu'une moyenne.

— **Déployer plusieurs capteurs spatialement**, positionnés à points fixes à des emplacements stratégiques ou portés par les travailleurs, voire par un vecteur mobile (robots, chariots, etc.) ce qui permet de capter la variabilité spatiale intra-atelier ou intra-zone. Cet avantage est bien illustré par les résultats de la fusion de données présentés dans ce manuscrit.

— **Générer des cartographies spatio-temporelles des concentrations** avec l'évolution temporelle de la pollution dans l'espace de travail, ce qui permet de relier les observations et les tâches de travail (ex : soudage, découpe, ventilation arrêtée) à l'exposition. Ceci peut aussi permettre d'évaluer l'impact précis des mesures de contrôle (ventilation, cloisonnement, horaires).

— **Optimiser le placement des capteurs et la densité du réseau** : par exemple, après une cartographie initiale, on peut réduire le nombre d'instruments tout en gardant une performance prédictive acceptable, comme démontré par [Berman18] dans un atelier de fabrication et étudier dans la section l'influence du nombre de points de mesures sur la cartographie spatio-temporelle 4.3.3.

En résumé, l'IMTR en réseau peut complémenter les stratégies de mesure d'une approche ponctuelle et moyennée (mesure en différé) vers une approche continue, dense et dynamique (mesure en temps réel), adaptée à la prise en compte de la complexité des ambiances profes-

sionnelles modernes et aux enjeux de santé au travail.

4.6.2 Périmètre de « validité » autour d'un point de mesure

Comme précisé dans la section précédente, des mesures à point fixe peuvent être intégrées à une stratégie de mesure visant à caractériser une situation de travail donnée. S'il est admis que la mesure à un point fixe dans un espace de travail n'est pas forcément représentative des l'ensemble de l'espace, la question corollaire est : jusqu'à quelle distance du point fixe la mesure est elle fiable ? Dans cette section nous avons tenté d'apporter des éléments de réponse à cette question en exploitant les cartographies spatio-temporelles qui ont été établies lors des essais. L'objectif est d'estimer le périmètre dans lequel une mesure reflète la concentration autour de ce point.

L'un des apports de la cartographie spatio-temporelle réside dans sa capacité à renseigner sur la continuité spatiale et temporelle des concentrations mesurées. Les modèles utilisés peuvent donc potentiellement permettre de définir et de quantifier le périmètre de validité associé à chaque point de mesure. Ce périmètre correspond à la zone au sein de laquelle la concentration mesurée par un capteur peut être considérée comme représentative de la concentration réelle dans son voisinage immédiat. Alors que cette notion demeure généralement implicite dans le cadre des campagnes de mesure conventionnelles, elle devient explicite lorsque les données issues d'un réseau de capteurs sont exploitées dans un processus de cartographie spatio-temporelle.

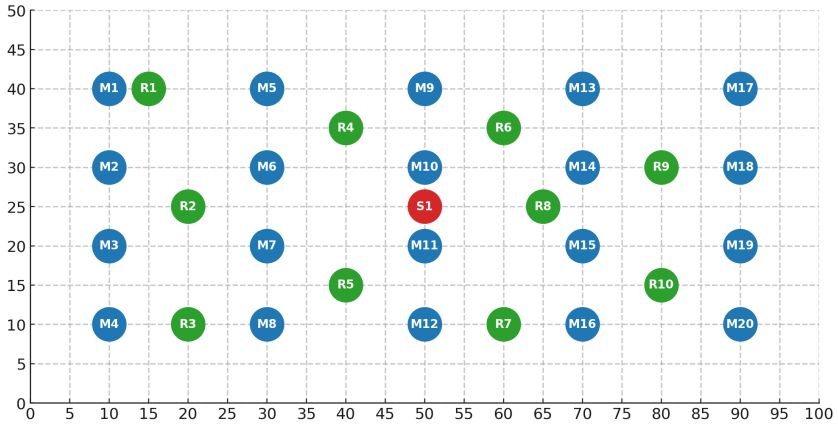


FIGURE 4.71 – Grille d'échantillonnage des points de mesures pour la cartographie spatio-temporelle avec M1-M20 les points de mesures du RCSFg, S1 la source de pollution, R1-R20 les points de référence

Pour traiter cette problématique, le grille d'échantillonnage correspondant à l'essai S1-E8 (Voir Figure 4.71) a été considérée. Conformément à la méthodologie présentée dans la section 4.3.1, une cartographie spatio-temporelle a été générée à partir des données recueillies par les modules du réseau de capteurs RCSFg (M1 à M20). L'application des méthodes de cartographie spatio-temporelle a permis de reconstituer l'évolution temporelle des concentrations en tout

point de la grille et sur l'ensemble de la durée de l'essai.

Pour déterminer le périmètre de validité, nous avons procédé par analyse du R^2 local. Pour chaque point de référence R1 à R10, la cohérence entre la concentration en CO_2 estimée et la mesure au point de référence est d'abord évaluée au moyen du R^2. Ensuite à partir des cartographies spatio-temporelles, la série temporelle associée à chaque point de l'espace est extraite, et un R^2 est calculé entre cette série et celle du point de référence. Cette opération permet d'obtenir, pour chaque référence, une carte de corrélation spatiale représentant la décroissance de la cohérence du signal avec la distance.

Les points de la carte sont ensuite classés en fonction de leur distance d au point de référence, puis les valeurs de R^2 sont analysées pour chaque distance croissante. Pour un rayon donné r, le quantile q des valeurs de R^2 à l'intérieur du disque de rayon r est calculé, permettant ainsi de définir une fonction décroissante décrivant la diminution de la corrélation spatiale avec l'éloignement.

Par exemple, les périmètres de validité de chaque point de référence ont été calculés pour l'essai S1-E9 avec la méthode PINNs pour la cartographie spatio-temporelle et sont ensuite définis à partir de seuils de performance fixés ($R^2 \geq 0.9$, $R^2 \geq 0.75$, $R^2 \geq 0.5$), correspondant respectivement à des zones de très forte, forte et faible représentativité spatiale. Ces seuils permettent de délimiter des *zones de validité* (zones verte, jaune, orange et rouge) centrées sur chaque point de référence, comme illustré dans les figures 4.72, 4.73.

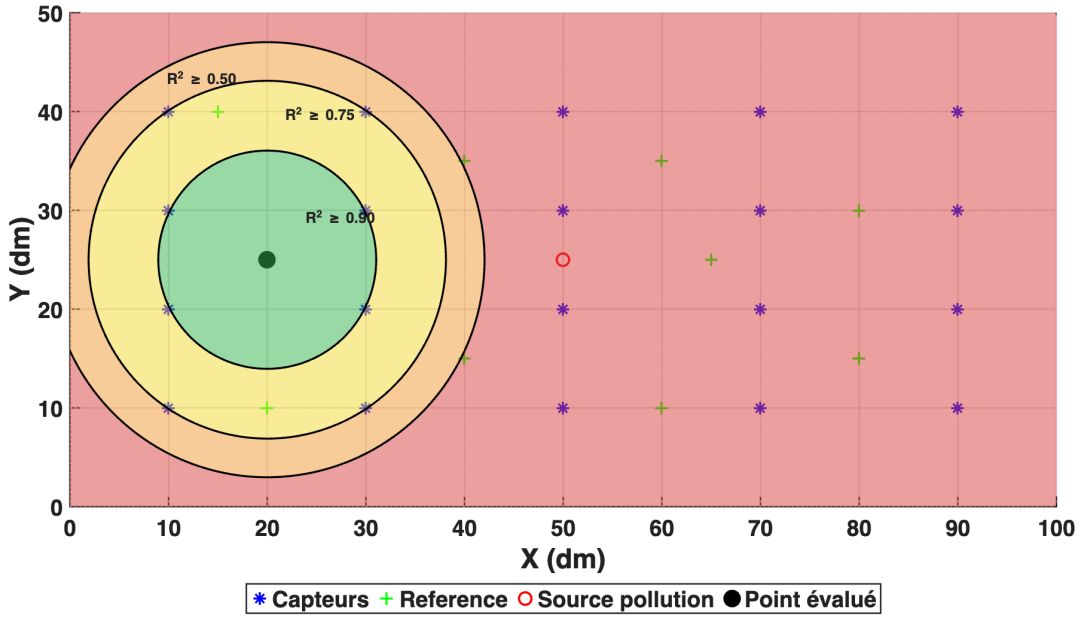


FIGURE 4.72 – Représentation du périmètre de validité du point de référence R2 en fonction du niveau de corrélation spatiale (R^2)

Ainsi, pour un point donné, le rayon de validité est défini comme la plus grande distance pour laquelle le coefficient R^2 reste supérieur à un seuil spécifié. Ce rayon traduit la distance

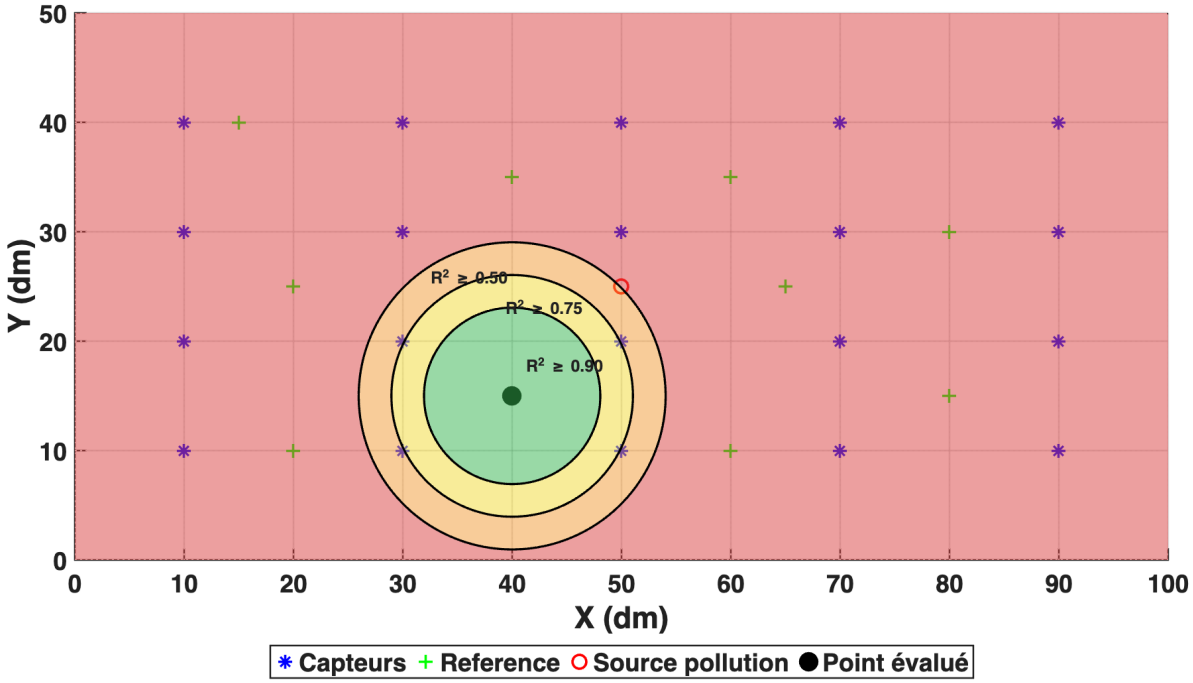


FIGURE 4.73 – Représentation du périmètre de validité du point de référence R5 en fonction du niveau de corrélation spatiale (R^2)

maximale au-delà de laquelle la mesure ne peut plus être considérée comme représentative de l'environnement local.

En pratique, cette approche permet de quantifier spatialement la portée d'une mesure ponctuelle, en intégrant les effets conjoints de la variabilité spatio-temporelle des concentrations et de la précision du modèle de cartographie. Le périmètre de validité constitue ainsi un indicateur pertinent pour interpréter la représentativité spatiale des mesures de polluants. Il peut aussi être utilisé pour dimensionner les réseaux de capteurs en fonction du degré de continuité souhaité.

Référence	**R**($R^2 \geq 0.9$)	**R**($R^2 \geq 0.75$)	**R**($R^2 \geq 0.5$)	**R(Validité)**
R1	3.61	5.66	7.62	5.66
R2	11.05	18.11	22.02	18.11
R3	5.00	7.21	37.20	7.21
R4	5.39	8.06	10.82	8.06
R5	8.06	11.05	14.04	11.05
R6	12.00	17.00	22.85	17.00
R7	9.06	20.22	25.46	20.22
R8	18.36	23.71	28.44	23.71
R9	20.12	38.42	43.93	38.42
R10	33.53	40.26	45.34	40.26

TABLEAU 4.18 – Rayons de validité associés à chaque point de référence (R1-R10) en dm calculés avec les données de cartographie spatio-temporelle de l'essai S1-E8

L'analyse des résultats présentés dans le tableau 4.18 met en évidence une variabilité du

rayon de validité selon la position du point de référence dans la grille d'échantillonnage. Cette variabilité traduit la dépendance spatiale du champ de concentration en CO_2, influencée à la fois par la position relative des capteurs, la proximité de la source d'émission et la dynamique de ventilation dans la chambre expérimentale.

Les points de référence situés à proximité immédiate de la source de pollution (comme R1 et R3) présentent des rayons de validité plus restreints, de l'ordre de 50 à 80 cm, indiquant une forte hétérogénéité locale des concentrations et une décroissance rapide de la corrélation spatiale. À l'inverse, les points plus éloignés de la source, tels que R8, R9 et R10, présentent des rayons de validité nettement supérieurs, dépassant 200 cm pour les seuils $R^2 \geq 0.75$, ce qui traduit une plus grande homogénéité spatiale dans ces zones du domaine.

En moyenne, pour l'ensemble des dix points de référence, le rayon de validité associé au seuil $R^2 \geq 0.75$ est de 17.07 dm. Cette valeur représente la distance moyenne à partir de laquelle une mesure ponctuelle de concentration peut être considérée comme représentative de son environnement local dans les conditions de l'essai S1-E8. Les rayons moyens pour les autres seuils de corrélation sont respectivement de 12.42 dm pour $R^2 \geq 0.9$ et de 25.07 dm pour $R^2 \geq 0.5$.

Ces résultats confirment que le périmètre de validité d'un point de mesure dépend étroitement du niveau de corrélation retenu, mais aussi des caractéristiques dynamiques du champ de concentration. Un seuil de $R^2 = 0.75$ apparaît ainsi comme un compromis pertinent entre fidélité de représentation et étendue spatiale. Il permet d'identifier la zone autour d'un capteur dans laquelle la mesure reste fiable sans surestimer sa portée.

En synthèse, les résultats indiquent que l'exploitation des cartographies spatio-temporelles permet de quantifier la représentativité spatiale des mesures issues d'un réseau de capteurs et d'objectiver la notion de proximité en cartographie spatio-temporelle. Le rayon de validité moyen ainsi obtenu constitue un indicateur opérationnel pour optimiser la densité de capteurs dans les futurs dispositifs de surveillance et améliorer la précision globale des modèles de reconstitution des concentrations. Ils montrent aussi que le périmètre de validité de la mesure autour d'un point de mesure n'a pas une valeur unique. En première approche, la distance minimale de validité de la mesure peut être estimée entre 5 et 8 dm. Cette estimation pourrait être confirmée en appliquant l'approche décrite dans la présente section à l'ensemble des essais réalisés au cours de la thèse. Il serait également nécessaire de reproduire cette analyse avec des données de terrain afin de valider les résultats obtenus.

Pour les experts en prévention, cela signifie que la cartographie aide à :

— Définir l'espacement maximal entre capteurs tout en conservant une précision acceptable (ce qui permet aussi d'optimiser les coûts de déploiement) ;

— Identifier les zones où un seul point de mesure n'est pas suffisant, par exemple dans les zones de forte hétérogénéité spatiale des concentrations, à proximité des sources d'émis-

sion ou dans les zones mal ventilées.

4.6.3 Identification des zones de danger

L'un des bénéfices majeurs de la cartographie spatio-temporelle en milieu professionnel réside dans sa capacité à identifier non seulement les niveaux de concentration, mais également les sources d'émission et les dynamiques spatio-temporelles associées. Cette approche permet une représentation continue et quantitative des gradients de concentration, ouvrant la voie à une analyse fine des mécanismes de dispersion et d'accumulation des polluants au sein de l'espace de travail.

La cartographie spatio-temporelle des concentrations peut aussi être exploitée pour identifier des zones de danger. En effet, plutôt que de cartographier les zones de concentration, il serait pertinent de représenter les zones dans lesquelles les concentrations dépassent des niveaux de référence au delà desquels la concentration représentent un danger pour la santé d'un travailleur qui y serait exposé. Dans le cadre de la prévention des risques chimiques, ces niveaux de référence existent pour certains agents chimiques avec les valeurs limites d'exposition professionnelle (VLEP) et des valeurs limites à court terme (VLCT). Les VLEP correspondent à la concentration moyenne pondérée sur une période de référence (généralement 8 h), en dessous de laquelle l'exposition quotidienne d'un travailleur est réputée sans effet nocif pour la santé. Les VLCT, quant à elles, définissent des seuils maximaux instantanés à ne pas dépasser, même sur une durée très brève (souvent 15 min), afin de prévenir les effets aigus ou irritatifs [INRS23].

Pour certaines situations de travail (l'opération de soudage, de nettoyage au solvant ou de maintenance, etc.), ces valeurs réglementaires, peuvent être dépassées localement et transitoirement sans être détectées par des méthodes de mesure en différées. L'exploitation de réseaux de capteurs à mesure temps réel (IMTR) couplés à des modèles de cartographie spatio-temporelle pourrait alors constituer une alternative intéressante et complémentaire en fournissant une vision dynamique et localisée du champ de danger.

Il serait ainsi possible de générer des cartes révèlant les zones où les concentrations approchent ou excèdent ponctuellement les VLEP ou VLCT, en représentant non pas les niveaux de concentration mais les pourcentages qu'ils représentent par rapport à la VLEP et la VLCT. Ces cartes permettraient d'objectiver les zones de danger caractérisées. De telles informations, difficilement accessibles par des mesures à point fixe, permettent d'identifier les sources primaires (points d'émission localisés) ainsi que les sources secondaires (zones de ré-accumulation dues à la recirculation de l'air ou à une ventilation inefficace).

4.6.4 Conclusion partielle

Sur le plan méthodologique, l'analyse spatio-temporelle des pics de concentration permet de corréler les événements de dépassement à des phases précises du processus de production ou à des conditions environnementales spécifiques (arrêt de ventilation, ouverture d'une enceinte, déplacement d'un opérateur). Ainsi la cartographie spatio-temporelle grâce aux travaux de fusion de données 3D pour la représentation de l'environnement expérimental avec les niveaux d'expositions (Voir section 4.4) facilite la détection des gradients maximaux et le repérage de trajectoires de propagation des polluants dans l'espace de travail.

En pratique, la cartographie spatio-temporelle constitue un outil d'aide à la décision pour la prévention en santé au travail. Elle permet :

— **D'identifier les zones d'émission ou d'accumulation** où la concentration dépasse les seuils de sécurité ;

— **De localiser et caractériser les dépassements des VLEP et VLCT** dans le temps et l'espace ;

— **D'hiérarchiser les zones de danger** selon leur niveau de risque et leur fréquence d'occurrence ;

— **D'évaluer le bon fonctionnement des systèmes de ventilation et de protection** (ventilation, captage à la source, cloisonnement).

Conclusion et perspectives

La caractérisation de l'exposition professionnelle aux polluants chimiques constitue un enjeu majeur en santé au travail, en raison de ses effets sur la santé des travailleurs et de la difficulté à caractériser finement les situations d'exposition. Elle repose sur la capacité à mesurer, à modéliser et à représenter les concentrations en polluants dans le temps et dans l'espace. Cependant, les méthodes traditionnelles d'interpolation spatiale, telles que le krigeage, ne permettent pas de capturer avec précision la dynamique temporelle ni les hétérogénéités fines des phénomènes de transferts de polluants observés dans les environnements réels.

Le travail présenté visait principalement à développer des approches de cartographie spatio-temporelle des concentrations en polluants intégrant explicitement la dimension temporelle et exploitant les mesures issues de réseaux de capteurs déployés en milieu de travail. Il s'agissait de reconstruire le champ spatio-temporel des concentrations à partir de données partielles et bruitées, en combinant des stratégies d'échantillonnage, de régularisation et de modélisation physique.

Les travaux de recherche ont été organisés suivant deux étapes successives : (i) un travail sur des données simulées visant à développer et à évaluer des modèles de cartographie spatio-temporelle intégrant explicitement la dimension temporelle ; (ii) un travail d'application de ces modèles à des données acquises par le bais d'expérimentations menées en conditions de laboratoire avec des réseaux de capteurs et des instruments de mesure en temps réel.

Deux nouvelles méthodes de cartographie spatio-temporelle ont ainsi été développées. La première est fondée sur un modèle tensoriel et formulée comme un problème d'optimisation intégrant des contraintes de douceur spatio-temporelle, permettant de restituer la continuité des champs de diffusion. La seconde correspond à un modèle de réseaux de neurones informés par la physique (« Physics-Informed Neural Networks », PINNs), qui intègre les équations différentielles gouvernant le phénomène de *diffusion-réaction* des polluants au sein de la fonction coût, afin d'améliorer la cohérence physique de la reconstruction. La comparaison des méthodes développées avec une approche classique, le krigeage, montre que les méthodes développées (i) conduisent à une meilleure cohérence temporelle des cartes crées et, (ii) permettent de mieux représenter les hétérogénéités fines des phénomènes de diffusion observés. De plus, la régularisation tensorielle et les PINNs se distinguent du krigeage, en offrant une meilleure stabilité et

précision des cartes reconstruites.

Ces méthodes ont été appliquées à des données expérimentales acquises en environnement contrôlé en mesurant les concentrations ambiantes en gaz (CO_2) et en particules avec des réseaux de capteurs et des instruments de mesure en temps réel, déployés dans un environnement d'essai. Ces essais, ainsi que les résultats qui y sont associés, permettent de formuler plusieurs conclusions. Premièrement, l'application des approches développées est opérationnelle et la modélisation apparaît indissociable des méthodes et stratégies de mesures employées pour l'acquisition des données de concentration dans le processus de cartographie spatio-temporelle. Deuxièmement, le recours à des indicateurs objectifs est essentiel pour juger de la fiabilité des cartes dans l'espace et le temps. Ceux qui ont été expérimentés dans le cadre des travaux rapportés dans le présent manuscrit (RMSE, R^2) ont montré leur pertinence. Ils pourraient être intégrés dans la cartographie de manière à donner une indication visuelle de la robustesse des cartes. Par ailleurs, le calcul de ces indicateurs nécessite le déploiement de détecteurs de référence en même temps que le réseau de capteurs ; aussi, les résultats de nos essais ouvrent clairement un champ d'investigation sur les stratégies de déploiement de ces détecteurs de référence à mettre en œuvre pour tirer les meilleurs profits de ces indicateurs. Troisièmement, les essais ont permis d'identifier les principales caractéristiques de la cartographie spatio-temporelle (choix de la méthode de cartographie, grille de positionnement et nombre des points de mesures, type de polluant, nombre et orientation des sources, etc.) qui déterminent sa performance. Quatrièmement, les travaux ont également permis de fusionner la cartographie des concentrations en polluants avec une modélisation 3D de l'environnement acquise par le bais d'un scanner. Ce résultat montre que la fusion de données hétérogènes est techniquement réalisable dans un contexte « santé au travail ». De plus, l'intégration des cartes construites, directement dans l'espace qui a été cartographié, ouvre des perspectives d'application intéressantes pour objectiver des situations de travail. La méthode doit encore être expérimentée sur des environnements de travail réels. Des travaux complémentaires pourraient être envisagés en vue de (i) produire un démonstrateur opérationnel dont la mise en œuvre ne soit pas trop chronophage et, (ii) étendre la fusion à une cartographie 3D. Cinquièmement, les travaux ont aussi montré que l'exposition individuelle d'un opérateur se déplaçant dans un environnement pollué peut être estimée en fusionnant la cartographie spatio-temporelle des concentrations avec la trajectoire de l'opérateur dans cet environnement. Cette application de la cartographie constitue un résultat très important pour la santé au travail et elle pourrait constituer une approche nouvelle dans ce contexte. D'un côté, elle permettrait d'estimer l'exposition individuelle des travailleurs à partir des concentrations ambiantes mesurées dans leur environnement. D'un autre côté, elle pourrait compléter les modèles de simulations numériques qui sont mis en œuvres, par exemple, pour simuler l'effet d'une ventilation sur les concentrations en polluants.

Les travaux confirment que plusieurs facteurs influencent la robustesse des cartographies

spatio-temporelles. En particulier, ils montrent que la capacité des méthodes à représenter la dynamique spatio-temporelle des polluants repose la prise en compte de différents paramètres spatiaux, temporels et physiques dans les algorithmes utilisés. Par ailleurs, les méthodes et stratégies d'acquisition des données utilisées pour réaliser les cartographies spatio-temporelles ont également un rôle déterminant. Les premières applications sur des données acquises en conditions de laboratoire montrent que l'utilisation d'un réseau de capteurs couplée à une méthode de cartographie peut-être utile dans un contexte de prévention (localisation de sources, estimation des expositions individuelles, etc.). Elles montrent également que la cartographie spatio-temporelle présente des limites propres aux modèles utilisés, dont il convient de tenir compte lors de l'interprétation des cartes et de l'estimation de l'exposition aux polluants.

Ces réalisations représentent une contribution significative à l'application de la cartographie spatio-temporelle aux problématiques de santé au travail. Elles offrent la possibilité de mieux appréhender les émissions de polluants et de reconstituer les expositions individuelles aux polluants en environnement de travail . Elles ouvrent également des perspectives de travaux de recherches qui pourraient s'articuler autour deux axes méthodologiques (Voir Figure 74) :

— un axe concernant l'amélioration des modèles de cartographie spatio-temporelle des polluants ;

— un axe plus appliqué, en lien avec l'instrumentation et la mesure, le développement de stratégies d'acquisition et visant une amélioration de la qualité d'estimation de l'exposition des travailleurs aux polluants.

Le premier axe concerne les méthodes de cartographie développées. Pour la méthode tensorielle, il peut être envisagé d'exploiter la structure de rang faible du tenseurs de données. L'un des principaux avantages des modèles tensoriels est qu'ils admettent des représentations de rang faible (*e.g.,* CP, Tucker [Kolda09]) avec des garanties d'unicité sous des faibles contraintes. L'utilisation de ces représentations de rang faible permettraient de régulariser d'avantage les cartographies spatio-temporelles tout en réduisant le coût de calcul des méthodes (réduction du nombre des paramètres). Une autre piste d'amélioration de la méthode est la prise en compte des géométries plus complexes (et pas forcement régulières) des environnements à cartographier. Il s'agit ici de prendre explicitement en compte, dans les algorithmes d'estimation, les régions où la propagation des polluants ne peut pas se faire, ce qui donnerait des représentations plus réalistes de la distribution des polluants.

Pour la méthode fondée sur le réseau de neurones (PINNs), une première piste serait l'amélioration de l'identifiabilité et de l'interprétabilité. L'estimation simultanée des paramètres $k(t)$ et $f(x, y, z)$ peut conduire à des solutions non-uniques ou instables. Introduire des contraintes de lissage ou des pénalités sur ces deux termes permettraient de garantir la plausibilité physique et limiter le surapprentissage. Une autre solution serait de représenter la fonction $f(x, y, z)$ par

une combinaison de fonctions de base (*e.g,* polynômes ou noyaux gaussiens) dont les coefficients sont appris par le réseau, ce qui renforcerait la stabilité et la robustesse de l'estimation. Une deuxième piste concerne l'extension multiphysique du réseau PINNs. En effet, le transport de polluants est souvent couplé à des phénomènes de convection d'air, de température et d'humidité. Un solution pour la prise en compte de ces phénomènes consisterait à utiliser un modèle *advection–diffusion–réaction* incluant un champ de vitesse, décrit par une EDP de type :

$$\frac{\partial \boldsymbol{X}(x,y,z)}{\partial t} + \langle \mathbf{u}(x,y,t), \nabla \boldsymbol{X}(x,y,z) \rangle = k(t)\,\nabla^2 \boldsymbol{X}(x,y,z) + f(x,y,t), \tag{4}$$

avec $\boldsymbol{u}(x,y,t)$ un champ de vitesse vectoriel. Autrement dit, $\boldsymbol{u}(x,y,t)$ est une fonction qui, en chaque point de l'espace et à chaque instant, indique la direction et la vitesse du flux d'air. Dans ce cas, développer des PINNs multi-sorties permettrait d'inférer simultanément le champ de concentration $\boldsymbol{X}$ et le champ de vitesse $\boldsymbol{u}$ tout en respectant les équations de conservation de la masse et de la quantité de mouvement.

Le deuxième axe vise à :

— Appliquer la cartographie spatio-temporelle, fondée sur des réseaux de capteurs, à des situations réelles en entreprises. C'est une étape qui était initialement prévue dans le programme de la thèse mais qui n'a pas pu être menée.

— Étudier et définir des stratégies de déploiement des instruments de mesure de référence en vue de mieux appréhender l'évaluation de la robustesse des cartes.

— Mener des travaux complémentaires permettant de réaliser des cartographies spatio-temporelles en 3D (x, y, z).

— Poursuivre le développement d'un démonstrateur opérationnel permettant de fusionner des données hétérogènes intégrant l'espace cartographié.

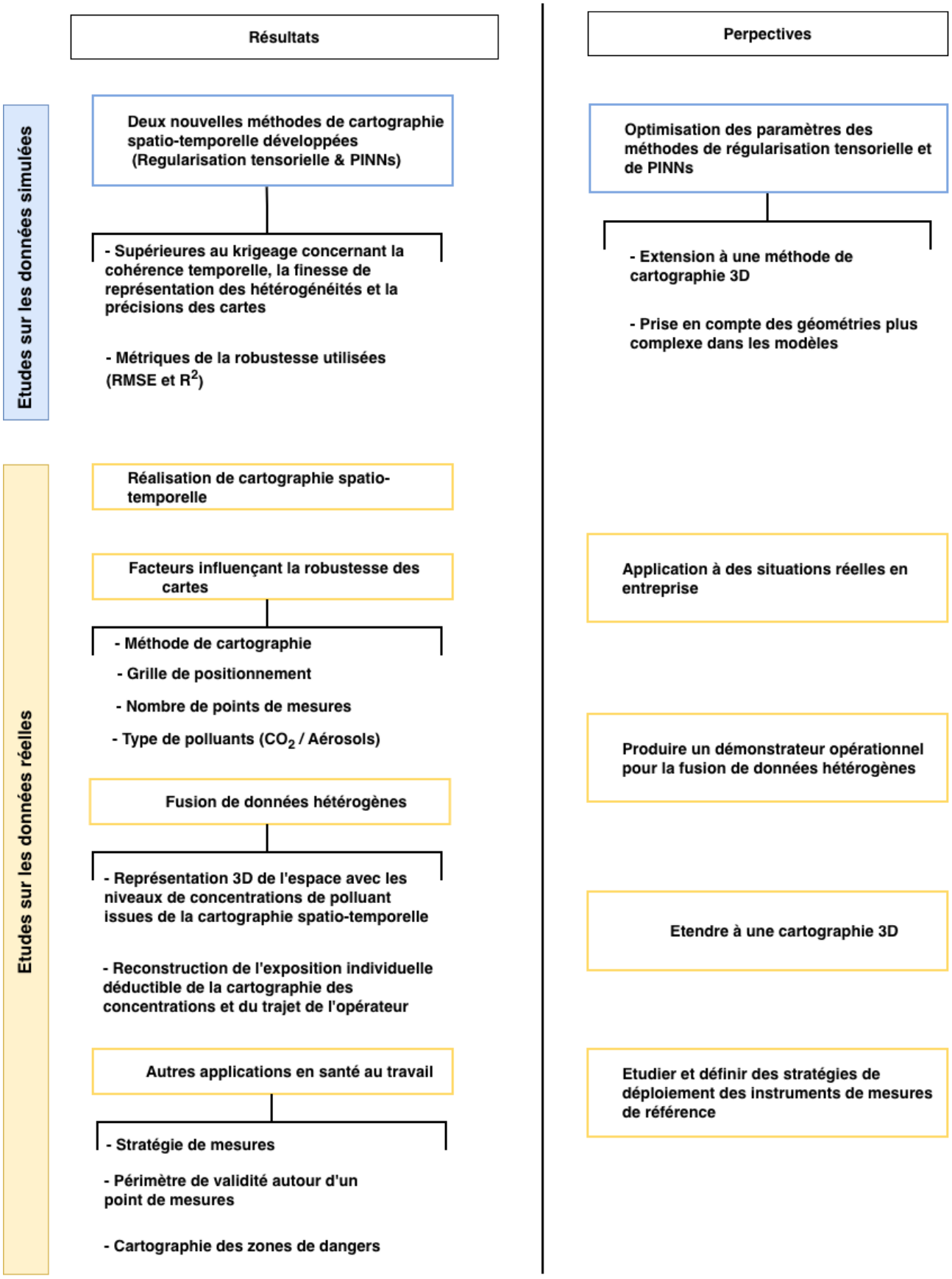


FIGURE 74 – Synthèse des résultats obtenus et perspectives des travaux de recherche

Annexes

Tableau 19 – Erreurs quadratiques moyennes (RMSE) et coefficients (R^2) entre les mesures des capteurs du réseau RCSFg acquises au cours de l'essai S1-E8 et les données CFD.

Modules capteurs	RMSE (ppm/v)	R^2
M1	240.89	0.873
M2	243.33	0.730
M3	248.14	0.876
M4	106.73	0.972
M5	143.71	0.946
M6	154.68	0.943
M7	247.13	0.879
M8	155.51	0.936
M9	105.38	0.962
M10	1155.40	0.207
M11	1120.75	-2.151
M12	543.45	-0.152
M13	108.61	0.958
M14	357.75	-0.104
M15	125.43	0.943
M16	216.83	0.814
M17	82.52	0.974
M18	115.34	0.952
M19	90.79	0.964
M20	126.26	0.940

TABLEAU 20 – Erreurs quadratiques moyennes (RMSE) et coefficients (R^2) entre les mesures des modules capteurs du RCSFg acquises au cours de l'essai S1-E9 et les données CFD.

Modules capteurs	RMSE (ppm/v)	R^2
M1	192.80	0.918
M2	287.15	0.617
M3	129.01	0.963
M4	94.83	0.974
M5	145.29	0.953
M6	111.38	0.969
M7	252.60	0.820
M8	133.61	0.953
M9	334.51	0.713
M10	842.11	-1.880
M11	96.51	0.967
M12	124.35	0.936
M13	399.19	0.407
M14	382.08	-0.236
M15	156.65	0.910
M16	99.63	0.960
M17	97.03	0.965
M18	164.56	0.900
M19	114.64	0.943
M20	83.11	0.976

TABLEAU 21 – Erreurs quadratiques moyennes (RMSE) et coefficients (R^2) entre les mesures de référence et les données CFD.

Référence	RMSE (ppm/v)	R^2
R1	250.79	0.862
R2	313.96	0.819
R3	277.73	0.844
R4	202.97	0.906
R5	1866.12	0.236
R6	659.36	-0.729
R7	170.11	0.901
R8	133.43	0.936
R9	142.18	0.926
R10	93.76	0.967

Tableau 22 – Erreurs quadratiques moyennes (RMSE) et coefficients (R^2) entre les mesures des modules capteurs du réseau Aranet acquises au cours des essais de la série 3 et les données des détecteurs de référence

Essai	A4–R4		A7–R7		A10–R10		A17–R16	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
S3-E0	207.347	0.9189	206.617	0.9266	201.851	0.9584	2459.580	-16.8456
S3-E1	204.686	0.9047	209.161	0.9151	107.910	0.9826	167.720	0.9457
S3-E2	238.839	0.8750	245.524	0.8253	144.577	0.9528	158.426	0.9453
S3-E3	208.709	0.9450	225.728	0.9374	188.472	0.9607	153.693	0.9736
S3-E4	231.087	0.9046	237.636	0.9193	179.700	0.9639	375.939	0.7775
S3-E5	188.115	0.9508	211.381	0.9401	157.676	0.9712	1748.477	0.0000
S3-E6	182.321	0.9346	268.255	0.8917	272.005	0.9279	1429.198	0.0000
S3-E7	229.022	0.9239	281.060	0.8426	178.553	0.9545	–	–
S3-E8	196.774	0.8801	230.757	0.8549	262.731	0.8944	283.714	0.8462
S3-E10	201.921	0.8948	225.522	0.8861	167.461	0.9640	140.415	0.9711
S3-E11	533.544	-0.2449	197.650	0.7462	173.599	0.8527	276.717	0.5784
S3-E12	1263.149	0.0000	278.111	0.8191	201.935	0.9388	232.313	0.9111
S3-E13	1067.685	0.0000	279.410	0.6790	177.465	0.9397	138.811	0.9739
S3-E14	281.588	0.8588	237.934	0.9065	149.396	0.9713	213.152	0.9394
S3-E15	496.248	0.4830	362.741	0.7548	260.091	0.9384	165.884	0.9558

Tableau 23 – Erreurs quadratiques moyennes (RMSE) et coefficients (R^2) entre les mesures des modules capteurs du réseau RCSFg acquises au cours des essais de la série 3 et les données des détecteurs de référence

Essai	M4–R4		M7–R7		M10–R10		M17–R16	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
S3-E0	222.635	0.9065	321.269	0.8226	206.678	0.9564	2478.298	-17.1183
S3-E1	226.292	0.8835	320.687	0.8005	103.461	0.9840	201.614	0.9215
S3-E2	299.995	0.8028	370.714	0.6018	151.540	0.9481	190.564	0.9208
S3-E3	275.704	0.9041	390.245	0.8130	181.190	0.9636	191.689	0.9589
S3-E4	231.132	0.9046	332.191	0.8422	156.443	0.9726	208.321	0.9317
S3-E5	243.275	0.9177	366.123	0.8202	153.179	0.9728	1803.373	0.0000
S3-E6	169.232	0.9436	391.342	0.7694	208.325	0.9577	1422.500	0.0000
S3-E7	235.714	0.9194	351.681	0.7535	169.863	0.9588	–	–
S3-E8	222.198	0.8471	317.162	0.7259	251.596	0.9032	233.732	0.8956
S3-E10	237.356	0.8546	301.681	0.7961	158.266	0.9678	154.940	0.9649
S3-E11	507.881	-0.1281	226.809	0.6657	253.798	0.6852	196.158	0.7881
S3-E12	1192.298	0.0000	283.221	0.8124	302.939	0.8623	258.616	0.8898
S3-E13	1103.224	0.0000	294.701	0.6429	320.291	0.8037	227.956	0.9297
S3-E14	267.400	0.8726	327.328	0.8230	317.289	0.8707	232.402	0.9279
S3-E15	301.497	0.8092	338.527	0.7864	372.609	0.8736	193.205	0.9400

Publications

Article de journal (soumis) :

— Houessou, N., Miron, S., Gerardin, K., Galland, B. and Duquenne, P. Accuracy of sensor networks applied for spatiotemporal mapping of airborne CO2 concentration for occupational purposes ; a study under laboratory conditions. *International Journal of Hygiene and Environmental Health*

Conférences internationales

— Houessou, N., Miron, S., Gerardin, K., Galland, B. and Duquenne, P. A tensor regularized approach to spatio-temporal pollutant cartography applied to working environments. In *IEEE ICFSP 9th International Conference on Frontiers of Signal Processing*, Paris, France, 2024, pp. 63-67, https://doi.org/10.1109/ICFSP62546.2024.10785370

— Houessou, N., Miron, S., Gerardin, K., Galland, B. and Duquenne, P. Spatiotemporal mapping of occupational exposures to airborne chemical using a physics-informed neural network method with sparse observations. In *The 11th International Symposium on Modern Principles of Air Monitoring and Biomonitoring (AIRMON 2025)* https://hal.science/hal-05262108

— Houessou, N., Miron, S., Gerardin, K., Galland, B. and Duquenne, P. Design and characterisation of a methodology for spatio-temporal mapping of exposures at the workplace. In *5th PEROSH Research Conference, (PEROSH 2023)* https://hal.science/hal-05262040

— Houessou, N., Miron, S., Gerardin, K., Galland, B. and Duquenne, P. Machine learning and data fusion for environmental monitoring. Application to risk exposure prevention in working environments. In *7th EUROSHNET European conference, Artificial Intelligence meets safety and health at work, (EUROSHNET 2022)* https://hal.science/hal-05262104

Bibliographie

[Aerts18] Aerts S., Wiart J., Martens L., and Joseph W. Assessment of long-term spatio-temporal radiofrequency electromagnetic field exposure. *Environmental Research*, 161 :136–143, 2018.

[Akyildiz02] Akyildiz I.F., Su W., Sankarasubramaniam Y., and Cayirci E. Wireless sensor networks : A survey. *Computer Networks*, 38(4) :393–422, 2002. doi :10.1016/S1389-1286(01)00302-4.

[Andersson93] Andersson I.M., Niemelä R., Rosén G., and Säämänen A. Control of styrene exposure by horizontal displacement ventilation. *Applied Occupational and Environmental Hygiene*, 8(12) :1031–1037, 1993.

[Ash'aari20] Ash'aari Z.H., Aris A.Z., Ezani E.H., Kamal N.I.A., Jaafar N., Jahaya J.N., Manan S.A., and Saifuddin M.F.U. Spatiotemporal variations and contributing factors of air pollutant concentrations in malaysia during movement control order due to pandemic covid-19. *Aerosol and Air Quality Research*, 20(10) :2047–2061, 2020. doi :10.4209/aaqr.2020.06.0334.

[Beck17] Beck A. *First-order methods in optimization*. Society for Industrial and Applied Mathematics, 2017. doi :10.1137/1.9781611974997.

[Bémer24] Bémer D. and Gérardin F. Changes in aerosol concentration in a meeting room with a portable air cleaner - mapping fine particle concentrations. *Aerosol and Air Quality Research*, 24(11) :240087, 2024. doi :10.4209/aaqr.240087. Open access.

[Berman18] Berman J.D., Peters T.M., and Koehler K.A. Optimizing a sensor network with data from hazard mapping demonstrated in a heavy-vehicle manufacturing facility. *Annals of Work Exposures and Health*, 62(5) :547–558, 2018.

[Borgini20] Borgini A., Ricco M., Veronese C., Farina M., and Cattaneo A. Indoor air quality and co_2 concentration levels in schools and workplaces : Insights for exposure assessment. *Science of The Total Environment*, 742 :140573, 2020. ISSN 0048-9697. doi :10.1016/j.scitotenv.2020.140573.

[Brunton20] Brunton S.L. and Kutz J.N. *Data-Driven Science and Engineering : Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2020. doi :10.1017/9781108380690.

[Buschmann16] Buschmann S., Trapp M., and Döllner J. Animated visualization of spatial–temporal trajectory data for air-traffic analysis. *The Visual Computer*, 32 :371–381, 2016.

[Cai10] Cai J.F., Candès E.J., and Shen Z. A singular value thresholding algorithm for matrix completion. *SIAM Journal on optimization*, 20(4) :1956–1982, 2010.

[Cameletti13] Cameletti M., Lindgren F., Simpson D., and Rue H. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *AStA Advances in Statistical Analysis*, 97 :109–131, 2013. doi :10.1007/s10182-012-0196-3.

[Candes12] Candes E. and Recht B. Exact matrix completion via convex optimization. *Communications of the ACM*, 55(6) :111–119, 2012.

[Cantat05] Cantat O. Dynamique spatio-temporelle d'un événement météo-climatique extrême-la canicule de l'été 2003 en europe de l'ouest. *Annales de l'Association Internationale de Climatologie*, 2 :99–136, 2005.

[Chen17] Chen D., Wang X., Li Y., Lang J., Zhou Y., Guo X., and Zhao Y. High-spatiotemporal-resolution ship emission inventory of china based on ais data in 2014. *Science of the Total Environment*, 609 :776–787, 2017.

[Chen21] Chen Z., Han J., Yang H., and Weinan E. Physics-informed learning of governing equations from sparse data. *Nature Communications*, 12(1) :6136, 2021. doi :10.1038/s41467-021-26434-1.

[Cherrie18] Cherrie J.W., Maccalman L., Fransman W., Tielemans E., and Tischer M. Modelling workplace exposure to airborne contaminants : Importance of source multiplicity. *Annals of Work Exposures and Health*, 62(10) :1221–1233, 2018. ISSN 2398-7316. doi :10.1093/annweh/wxy045.

[Cressie15] Cressie N. *Statistics for Spatial Data*. John Wiley & Sons, Hoboken, NJ, revised edition edition, 2015. ISBN 9781119115151. doi :10.1002/9781119115151.

[Cuomo22] Cuomo S., Di Cola V.S., Giampaolo F., Rozza G., Raissi M., and Piccialli F. Scientific machine learning through physics-informed neural networks : Where we are and what's next. *Journal of Scientific Computing*, 92(3) :88, 2022. doi :10.1007/s10915-022-01939-z.

[Davidson18] Davidson M.E., Schaeffer J., Clark M.L., Magzamen S., Brooks E.J., Keefe T.J., Bradford M., Roman-Muniz N., Mehaffy J., Dooley G., Poole J.A., Mitloehner F.M., Reed S., Schenker M.B., and Reynolds S.J. Personal exposure of dairy workers to dust, endotoxin, muramic acid, ergosterol, and ammonia on large-scale dairies in the high plains western united states. *Journal of Occupational and Environmental Hygiene*, 15(3) :182–193, 2018. doi :10.1080/15459624.2017.1403610.

[Daw22] Daw A., Buhrmester V., and Karniadakis G.E. Rethinking the importance of sampling in physics-informed neural networks. *Communications in Computational Physics*, 32(3) :838–866, 2022.

[De Iaco01] De Iaco S., Myers D.E., Posa D., and Wright R. Space-time analysis using a general product–sum model. *Statistics and Probability Letters*, 52(1) :21–28, 2001. doi :10.1016/S0167-7152(00)00200-6.

[Evans08] Evans D.E., Heitbrink W.A., Slavin T.J., and Peters T.M. Ultrafine and respirable particles in an automotive grey iron foundry. *Annals of Occupational Hygiene*, 52(1) :9–21, 2008.

[Felgueiras23] Felgueiras F., Mourão Z., Moreira A., and Gabriel M.F. Indoor environmental quality in offices and risk of health and productivity complaints at work : A literature review. *Journal of Hazardous Materials Advances*, 10 :100314, 2023. doi :10.1016/j.hazadv.2023.100314.

[Gao21] Gao H., Sun L., and Wang J.X. Phygeonet : Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state pdes on irregular domain. *Journal of Computational Physics*, 428 :110079, 2021. ISSN 0021-9991. doi :10.1016/j.jcp.2020.110079.

[Gelfand16] Gelfand A.E., Banerjee S., and Finley A.O. Spatial process modeling for univariate and multivariate data using gaussian processes. *Spatial Statistics*, 18 :37–51, 2016. doi :10.1016/j.spasta.2016.02.002.

[Giampaolo22] Giampaolo F., Venturi D., and Karniadakis G.E. A physics-informed neural network approach for solving reaction–diffusion equations. *Advances in Modeling and Simulation in Engineering Sciences*, 9(1) :5, 2022. doi : 10.1186/s40323-022-00219-7.

[Gneiting02] Gneiting T. Nonseparable, stationary covariance functions for space–time data. *Journal of the American Statistical Association*, 97(458) :590–600, 2002. doi :10.1198/016214502760047113.

[Goodfellow16] Goodfellow I., Bengio Y., and Courville A. *Deep Learning*. MIT Press, 2016.

[Heitbrink07] Heitbrink W.A., Evans D.E., Peters T.M., and Slavin T.J. Characterization and mapping of very fine particles in an engine machining and assembly facility. *Journal of Occupational and Environmental Hygiene*, 4(5) :341–351, 2007.

[Hu13] Hu Y., Zhou W., Wen Z., Sun Y., and Yin B. Efficient radio map construction based on low-rank approximation for indoor positioning. *Mathematical Problems in Engineering*, 2013(1) :461089, 2013.

[INRS23] INRS. Valeurs limites d'exposition professionnelle aux agents chimiques en france. actualisation 2023. *Hygiène et Sécurité du Travail*, 270 :5–31, 2023. doi :10.47446/HST.INRS.270.VLEP2023.

[Jumpponen13] Jumpponen M., Rönkkömäki H., Pasanen P., and Laitinen J. Occupational exposure to gases, polycyclic aromatic hydrocarbons and volatile organic compounds in biomass-fired power plants. *Chemosphere*, 90(6) :1289–1293, 2013. doi :10.1016/J.CHEMOSPHERE.2012.10.001.

[Karniadakis21] Karniadakis G.E., Kevrekidis I.G., Lu L., Perdikaris P., Wang S., and Yang L. Physics-informed machine learning. *Nature Reviews Physics*, 3(6) :422–440, 2021. doi :10.1038/s42254-021-00314-5.

[Koehler11] Koehler K.A. and Volckens J. Prospects and pitfalls of occupational hazard mapping : 'between these lines there be dragons'. *Annals of Occupational Hygiene*, 55(8) :829–840, 2011.

[Koehler13] Koehler K.A. and Peters T.M. Influence of analysis methods on interpretation of hazard maps. *Annals of Occupational Hygiene*, 57(5) :558–570, 2013.

[Kolda06] Kolda T.G. Multilinear operators for higher-order decompositions. Technical Report SAND2006-2081, Sandia National Laboratories (SNL), Albuquerque, NM and Livermore, CA, United States, 2006.

[Kolda09] Kolda T.G. and Bader B.W. Tensor decompositions and applications. *SIAM Review*, 51(3) :455–500, 2009. doi :10.1137/07070111X.

[Lake15] Lake K., Zhu J., Wang H., Volckens J., and Koehler K.A. Effects of data sparsity and spatiotemporal variability on hazard maps of workplace noise. *Journal of Occupational and Environmental Hygiene*, 12(4) :256–265, 2015.

[Léger11] Léger F., Yu G., and Sapiro G. Efficient matrix completion with gaussian models. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1113–1116. IEEE, 2011.

[Li14] Li L., Losser T., Yorke C., and Piltner R. Fast inverse distance weighting-based spatiotemporal interpolation : a web-based application of interpolating daily fine particulate matter pm2. 5 in the contiguous us using parallel programming and kd tree. *International journal of environmental research and public health*, 11(9) :9101–9141, 2014.

[Li18] Li J., Li H., Ma Y., Wang Y., Abokifa A.A., Lu C., and Biswas P. Spatiotemporal distribution of indoor particulate matter concentration with a low-cost sensor network. *Building and Environment*, 127 :138–147, 2018.

[Liu10] Liu S. and Hammond S.K. Mapping particulate matter at the body weld department in an automobile assembly plant. *Journal of Occupational and Environmental Hygiene*, 7(10) :593–604, 2010.

[Lu21] Lu L., Meng X., Mao Z., and Karniadakis G.E. Deepxde : A deep learning library for solving differential equations. *SIAM Review*, 63(1) :208–228, 2021.

[Ludwig17] Ludwig G., Chu T., Zhu J., Wang H., and Koehler K. Static and roving sensor data fusion for spatio-temporal hazard mapping with application to occupational exposure assessment. *The Annals of Applied Statistics*, 11(1) :139, 2017.

[Mandić-Rajčević16] Mandić-Rajčević S., Karanac M., Mihajlović M., and Jovanović M. Risk maps for industrial occupational health and safety. *Facta Universitatis, Series : Working and Living Environmental Protection*, pages 63–72, 2016.

[Mansouri10] Mansouri M., Sardouk A., Merghem-Boulahia L., Gaiti D., Snoussi H., Rahim-Amoud R., and Richard C. Factors that may influence the performance of wireless sensor networks. In Y.K. Tan, editor, *Smart Wireless Sensor Networks*, chapter 3. InTech, Rijeka, Croatia, 2010. doi : 10.5772/8688.

[Monsé14] Monsé C., Broding H.C., Sucker K., Berresheim H., Jettkant B., Hoffmeyer F., Merget R., Brüning T., and Bünger J. Exposure assessment of potash miners at elevated co_2 levels. *International Archives of Occupational and Environmental Health*, 87(5) :413–421, 2014. doi : 10.1007/s00420-013-0880-y.

[Mukwada18] Mukwada G. and Manatsa D. Spatiotemporal analysis of the effect of climate change on vegetation health in the drakensberg mountain region of south africa. *Environmental monitoring and assessment*, 190 :1–21, 2018.

[Myatt02] Myatt T.A., Minegishi T., Allen J.G., Macintosh D.L., Seppänen O., and Fisk W.J. A study of indoor carbon dioxide levels and sick leave among office workers. *Environmental Health*, 1(1) :3, 2002. doi :10.1186/1476-069X-1-3.

[O'Brien03] O'Brien D.M. Aerosol mapping of a facility with multiple cases of hypersensitivity pneumonitis : demonstration of mist reduction and a possible dose/response relationship. *Applied Occupational and Environmental Hygiene*, 18(11) :947–952, 2003.

[Pang21] Pang L., Zhang J., Wu X., Zhu H., Liu X., Xu S., and Luo J. The effects of carbon dioxide exposure concentrations on human vigilance and sentiment in an enclosed workplace environment. *Indoor Air*, 31(5) :1357–1367, 2021. doi :10.1111/ina.12746.

[Park10] Park J.Y., Ramachandran G., Raynor P.C., and Olson G.M. Determination of particle concentration rankings by spatial mapping of particle surface area, number, and mass concentrations in a restaurant and a die casting plant. *Journal of Occupational and Environmental Hygiene*, 7(8) :466–476, 2010.

[Park17] Park Y.M. and Kwan M.P. Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health & Place*, 43 :85–94, 2017.

[Patton22] Patton A.N., Medvedovsky K., Zuidema C., Peters T.M., and Koehler K. Probabilistic machine learning with low-cost sensor networks for occupational exposure assessment and industrial hygiene decision making. *Annals of Work Exposures and Health*, 66(5) :580–590, 2022.

[Pega22] Pega F., Hamzaoui H., Náfrádi B., and Momen N.C. Global, regional and national burden of disease attributable to 19 selected occupational risk factors for 183 countries, 2000–2016 : A systematic analysis from the who/ilo joint estimates of the work-related burden of disease and injury. *Scandinavian journal of work, environment & health*, 48(2) :158, 2022.

[Peng22] Peng W.Q., Pu J.C., and Chen Y. Pinn deep learning method for the chen–lee–liu equation : Rogue wave on the periodic background. *Communications in Nonlinear Science and Numerical Simulation*, 105 :106067, 2022. doi :10.1016/j.cnsns.2021.106067.

[Perchoux13] Perchoux C., Chaix B., Cummins S., and Kestens Y. Conceptualization and measurement of environmental exposure in epidemiology : accounting

for activity space related to daily mobility. *Health & Place*, 21 :86–93, 2013.

[Peters06] Peters T.M., Heitbrink W.A., Evans D.E., Slavin T.J., and Maynard A.D. The mapping of fine and ultrafine particle concentrations in an engine machining and assembly facility. *Annals of Occupational Hygiene*, 50(3) :249–257, 2006.

[Peters12] Peters T.M., Anthony T.R., Taylor C., Altmaier R., Anderson K., and O'Shaughnessy P.T. Distribution of particle and gas concentrations in swine gestation confined animal feeding operations. *Annals of Occupational Hygiene*, 56(9) :1080–1090, 2012.

[Rainham10] Rainham D., McDowell I., Krewski D., and Sawada M. Conceptualizing the healthscape : contributions of time geography, location technologies and spatial ecology to place and health research. *Social Science & Medicine*, 70(5) :668–676, 2010.

[Raissi19] Raissi M., Perdikaris P., and Karniadakis G.E. Physics-informed neural networks : A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378 :686–707, 2019. doi :10.1016/j.jcp.2018.10.045.

[Renoux98] Renoux A. and Boulaud D. *Les aérosols : Physique et métrologie*. Tec & Doc – Lavoisier, Paris, 1998. ISBN 978-2-7430-0208-4.

[Rivera-Gonzalez15] Rivera-Gonzalez L.O., Zhang Z., Sánchez B.N., Zhang K., Brown D.G., Rojas-Bracho L., Osornio-Vargas A., Vadillo-Ortega F., and O'Neill M.S. An assessment of air pollutant exposure methods in mexico city, mexico. *Journal of the Air & Waste Management Association*, 65(5) :581–591, 2015. doi :10.1080/10962247.2015.1020974.

[Rosen90] Rosen G., Andersson I.M., and Juringe L. Reduction of exposure to solvents and formaldehyde in surface-coating operations in the woodworking industry. *The Annals of Occupational Hygiene*, 34(3) :293–303, 1990.

[Ruiter24] Ruiter S., Franken R., Krone T., Le Feber M., Gunnink J., Kuijpers E., Peters S., Vermeulen R., and Pronk A. Spatiotemporal modeling of occupational particulate matter using personal low-cost sensor and indoor location tracking data. *Journal of Occupational and Environmental Hygiene*, 21(10) :696–708, 2024. doi :10.1080/15459624.2024.2389279.

[Schlünssen23] Schlünssen V., Mandrioli D., Pega F., Natalie N.M., Ádám B., Chen W., Cohen R.A., Godderis L., Göen T., Hadkhale K., Kunpuek W., Lou J.,

Mandic-Rajcevic S., Masci F., Nemery B., Popa M., Rajatanavin N., Sgargi D., Siriruttanapruk S., Sun X., Suphanchaimat R., Thammawijaya P., Ujita Y., van der Mierden S., Vangelova K., Ye M., Zungu M., and Scheepers P.P. The prevalences and levels of occupational exposure to dusts and/or fibres (silica, asbestos and coal) : A systematic review and meta-analysis from the who/ilo joint estimates of the work-related burden of disease and injury. *Environment International*, 178 :107980, 2023. doi : 10.1016/j.envint.2023.107980. Open access under CC BY 4.0.

[Shabat12] Shabat G. and Averbuch A. Interest zone matrix approximation. *The Electronic Journal of Linear Algebra*, 23 :678–702, 2012.

[Shen21] Shen H., Hou W., Zhu Y., Zheng S., Ainiwaer S., Shen G., and Tao S. Temporal and spatial variation of pm2.5 in indoor air monitored by low-cost sensors. *Science of The Total Environment*, 770 :145304, 2021.

[Shepard68] Shepard D. A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM national conference*, pages 517–524. 1968.

[Shin24] Shin J., Woo J., Choe Y., Min G., Kim D., Kim D., Lee S., and Yang W. Spatiotemporal exposure assessment of pm2.5 concentration using a sensor-based air monitoring system. *Atmosphere*, 15(6) :664, 2024. doi : 10.3390/atmos15060664.

[Sousan18] Sousan S., Gray A., Zuidema C., Stebounova L., Thomas G., Koehler K., and Peters T. Sensor selection to improve estimates of particulate matter concentration from a low-cost network. *Sensors*, 18(9) :3008, 2018.

[Stein05] Stein M.L. Space–time covariance functions. *Journal of the American Statistical Association*, 100(469) :310–321, 2005. doi :10.1198/016214504000000854.

[Steinle13] Steinle S., Reis S., and Sabel C.E. Quantifying human exposure to air pollution—moving from static monitoring to spatio-temporally resolved personal exposure assessment. *Science of the Total Environment*, 443 :184–193, 2013.

[Sun12] Sun Y., Li B., and Genton M.G. Geostatistics for large datasets. *Advances in Statistical Analysis*, 96 :47–62, 2012. doi :10.1007/s10182-011-0166-3.

[Vasilatou24] Vasilatou K., Iida K., Kazemimanesh M., Olfert J., Sakurai H., Sipkens T.A., and Smallwood G.J. Aerosol physical characterization : A review on the current state of aerosol documentary standards and calibration

strategies. *Journal of Aerosol Science*, 183 :106483, 2024. doi :10.1016/j.jaerosci.2024.106483.

[Vosburgh11] Vosburgh D.J., Boysen D.A., Oleson J.J., and Peters T.M. Airborne nanoparticle concentrations in the manufacturing of polytetrafluoroethylene (ptfe) apparel. *Journal of Occupational and Environmental Hygiene*, 8(3) :139–146, 2011.

[Wackernagel03] Wackernagel H. *Multivariate Geostatistics : An Introduction with Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[Wang10] Wang Y., Resnick S.M., and Davatzikos C. Spatio-temporal analysis of brain mri images using hidden markov models. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010 : 13th International Conference, Beijing, China, September 20-24, 2010, Proceedings, Part II*, pages 160–168. Springer Berlin Heidelberg, 2010.

[Wang13] Wang C., Quddus M., and Ison S. A spatio-temporal analysis of the impact of congestion on traffic safety on major roads in the uk. *Transportmetrica A : Transport Science*, 9(2) :124–148, 2013.

[Wang21] Wang S., Teng Y.Y., and Perdikaris P. Understanding and mitigating gradient pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, 43(5) :A3055–A3081, 2021.

[Xiao16] Xiao Y., Gu X., Yin S., Shao J., Cui Y., Zhang Q., and Niu Y. Geostatistical interpolation model selection based on arcgis and spatio-temporal variability analysis of groundwater level in piedmont plains, northwest china. *Springerplus*, 5 :425, 2016.

[Xue19] Xue L., Wang J., Zhang L., Wei G., and Zhu B. Spatiotemporal analysis of ecological vulnerability and management in the tarim river basin, china. *Science of the Total Environment*, 649 :876–888, 2019.

[Yang21] Yang L., Meng X., and Karniadakis G.E. Bpinns : Bayesian physics-informed neural networks for forward and inverse pde problems with noisy data. *Journal of Computational Physics*, 425 :109913, 2021.

[Zikova17] Zikova N., Masiol M., Chalupa D.C., Rich D.Q., Ferro A.R., and Hopke P.K. Estimating hourly concentrations of pm2.5 across a metropolitan area using low-cost particle monitors. *Sensors*, 17(8) :1922, 2017.

[Zuidema19a] Zuidema C., Klepeis N.E., and Bennett D.H. Spatiotemporal modeling of occupational particulate matter using personal low-cost sensor and indoor location tracking data. *Building and Environment*, 159 :106153, 2019. doi : 10.1016/j.buildenv.2019.106153.

[Zuidema19b] Zuidema C., Sousan S., Stebounova L.V., Gray A., Liu X., Tatum M., and Koehler K. Mapping occupational hazards with a multi-sensor network in a heavy-vehicle manufacturing facility. *Annals of Work Exposures and Health*, 63(3) :280–293, 2019.

Abstract

Occupational exposure to pollutants is a major concern in workplace health due to its potentially severe impacts on workers and the challenges associated with accurately characterizing exposure scenarios. Effective risk prevention depends on the ability to measure, model, and represent pollutant concentrations across both space and time. Traditional spatial interpolation methods (e.g., kriging), however, fail to capture temporal dynamics and the fine-scale heterogeneities of diffusion processes observed in real-world environments. This study aims to develop spatio-temporal cartography approaches for pollutant concentrations that explicitly account for the temporal dimension and exploit data from sensor networks deployed in occupational settings. The primary objective is to reconstruct the spatio-temporal concentration field from partial and noisy measurements by combining sampling and interpolation strategies with physical modeling. A secondary objective is to evaluate the feasibility of estimating individual workers' exposure based on the reconstructed maps and their spatio-temporal trajectories. Two spatio-temporal mapping approaches were developed : (1) a tensor-based modeling method incorporating spatio-temporal regularity constraints to ensure continuity in the diffusion fields ; and (2) a Physics-Informed Neural Network (PINN) approach, which integrates the governing reaction-diffusion equations of pollutants. These methods were evaluated using experimental data obtained in a controlled environment at INRS (Vandoeuvre-lès-Nancy) and integrated with spatio-temporal trajectories to reconstruct individual exposures. A 3D model of the workplace environment was also incorporated to generate realistic representations of the studied setting. The results demonstrate the capability of these approaches to accurately model the spatio-temporal dynamics of pollutants and highlight their potential as decision-support tools for chemical risk prevention in occupational environments.

Keywords: Spatio-temporal cartography ; Tensor-based modeling ; Pollution sensor networks ; Airborne pollutants ; Spatio-temporal regularization ; Physics-Informed Neural Networks (PINNs) ; Individual exposure ; Data fusion ; 3D modeling

Résumé

L'exposition professionnelle aux polluants constitue un enjeu majeur en santé au travail, en raison de ses effets potentiellement graves sur la santé des travailleurs et de la difficulté à caractériser précisément les situations d'exposition. La prévention de ces risques repose sur la capacité à mesurer, modéliser et représenter les concentrations de polluants dans le temps et dans l'espace. Cependant, les méthodes traditionnelles d'interpolation spatiale (ex. krigeage), ne permettent pas de capturer la dynamique temporelle ni les hétérogénéités fines des phénomènes de diffusion observés dans les environnements réels. Ce travail vise à développer des approches de cartographie spatio-temporelle des concentrations de polluants intégrant explicitement la dimension temporelle et exploitant les mesures issues de réseaux de capteurs déployés en milieu de travail. Le problème étudié consiste à reconstruire le champ spatio-temporel des concentrations à partir de données partielles et bruitées, en combinant des stratégies d'échantillonnage, d'interpolation avec la modélisation physique. Un second objectif est d'évaluer la faisabilité d'estimer l'exposition individuelle des travailleurs à partir des cartes reconstruites et de leurs trajectoires spatio-temporelles. Deux approches de cartographie spatio-temporelle ont été développées : (1) une méthode fondée sur une modélisation tensorielle intégrant des contraintes de régularité spatio-temporelle, permettant de restituer la continuité des champs de diffusion ; et (2) une approche basée sur les réseaux de neurones informés par la physique (Physics-Informed Neural Networks, PINNs), qui intègre les équations différentielles gouvernant la diffusion-réaction des polluants. Ces méthodes ont été évaluées à partir de données expérimentales acquises en environnement contrôlé à l'INRS (Vandoeuvre-lès-Nancy) et fusionnées avec des trajectoires spatio-temporelles pour la reconstruction de l'exposition individuelle. Une modélisation 3D de l'environnement de travail a également été intégrée afin de générer des représentations réalistes de l'environnement étudié. Les résultats obtenus démontrent la capacité de ces approches à modéliser la dynamique spatio-temporelle des polluants et ouvrent la voie à de nouveaux outils d'aide à la décision pour la prévention des risques chimiques dans les environnement de travail.

Mots-clés: Cartographie spatio-temporelle ; modélisation tensorielle ; réseaux de capteurs de pollution ; polluants atmosphériques ; régularisation spatio-temporelle ; PINNs ; exposition individuelle ; fusion de données ; modélisation 3D